

Carine Eskes

## **Foutenanalyses volgens globale of fijnmazige categorieën. Wat is de beste methode om spellingvaardigheid en –ontwikkeling in kaart te brengen?\***

Een onderzoek naar hoe gedetailleerd de categorieverdeling moet zijn om de spellingvaardigheid en –ontwikkeling te beschrijven, geïllustreerd aan de hand van medeklinkerverdubbeling.

### **Abstract**

What is the best method to map out spelling competence and development? That question will be answered by means of consonant duplication at the primary school in grades 3 to 8 (Dutch system). According to previous research, describing spelling competence on the basis of global category 4.2 ('inheemse verdubbelingswoorden') in the *Van Dale Basisspellinggids* does not give enough information. Spelling development, on the other hand, seems to be described quite well with this global distinction as regardless of the possible differences in development of the 'inheemse verdubbelingswoorden', the 'goedpercentage' in the 8<sup>th</sup> grade almost reaches 100%. The more detailed breakdown in subcategories 4.2a, 4.2b, 4.2c and 4.2d show that 'pudding-words' (subcategory 4.2c: words ending in -ing, -ig, -lijk, a schwa that is not written as <e>) and 'herrie-words' (subcategory 4.2d: native 'verdubbelingswoorden' that look like other categories, but differ from them because they have more than one 'volle vocaal') hardly occur in the corpus. The 'tamme words' (subcategory 4.2a: words with a suffix which contain a schwa, written as <e>) and 'korrel words' (subcategory 4.2b: words ending in -el, -er which are part of the base, no flecional suffixes, schwa written as <e>) give a valuable description of the spelling competence and development. The detailed

distinction in triplet clusters gives a clear picture and that is mainly because of the limited number of observations made in these categories.

*Keywords:* spelling, spelling development, spelling competence, consonant duplication, triplets

## 1 Inleiding

Begin op een verjaardag over *spelling* en het eerste wat je hoort, is: “Moeilijk hè, ik weet ook nooit of het nou met *dt* is of niet.” Dit voorbeeld illustreert dat ‘de gewone Nederlander’ de neiging heeft om te denken dat werkwoordspelling het moeilijkst is. Toch blijken er ook andere spellingonderwerpen te zijn die een groot aandeel hebben in spellingsfouten. Een van die andere spellingonderwerpen is de *medeklinkerverdubbeling*.

### 1.1 Theoretisch kader

Voor spelling in het algemeen is veel aandacht in het wetenschappelijke veld, zoals de discussies laten zien rond het Groene Boekje (1991/2005) en Witte Boekje (2006) en de vereenvoudigde spelling die toch niet zo eenvoudig leek. Tien jaar eerder concludeerde Ghesquière (1997, p. 81) na zijn exploratieve onderzoek naar spellingfouten onder Vlaamse leerlingen van het secundair onderwijs dat het zeker nog de moeite waard is om aandacht te besteden aan spelling(onderwijs).

Het is onmogelijk om alle spellingsonderdelen tegelijkertijd te onderzoeken, dus moet er een keuze gemaakt worden. Welk onderdeel gaat vaak fout en verdient extra onderzoek? Intuïtief zal het antwoord op deze vraag zijn: de werkwoordspelling, maar uit verschillende onderzoeken is gebleken dat het verenkelen en verdubbelen van letters een zeer groot aandeel heeft in het aantal spelfouten. Dit blijkt onder andere uit Verhoeven (1979). Ook uit het onderzoek van Booij, Hamans, Van Minnen, Verhoeven en Balk-Smit Duyzentkunst (1979) blijkt dat niet de werkwoordspelling, zoals vaak gedacht wordt, maar het ten onrechte dubbel of enkel spellen van consonanten de nummer 1 fout is met 17,2% fout. Het ten onrechte dubbel of enkel spellen van vocalen haalt een percentage van 10,8% en eindigt daarmee op de derde plaats in Booij's fouten-top 10. Ruijssenaars, Claes en De Caluwe (1992) zijn na het afnemen van een zinnendictie tot de conclusie gekomen dat wel 15% van de fouten medeklinkerverdubbeling- of klinkerverenkelingfouten zijn.

Bovenstaande uitkomsten zijn niet één op één met elkaar te vergelijken, omdat er bijvoorbeeld niet dezelfde foutenclassificatiesystemen gebruikt zijn, maar ook omdat er verschillende soorten leerlingenmateriaal gebruikt is, zoals dictees, opstellen en toetsen. Desalniettemin is de overeenkomst duidelijk: het verenken en verdubbelen verdient aandacht.

Op de vraag waarom er in het Nederlands verenkend en verdubbeld wordt, geven De Schryver en Neijt in hun *Handboek spelling* (2005, p. 69) een simpel antwoord: omdat er maar vijf lettertekens zijn om dertien klinkers weer te geven. De klanken [i], [u] en [Ø] worden weergegeven door aparte tekencombinaties, te weten <ie>, <oe>, en <eu>, maar voor de overige klanken maakt men onderscheid door enerzijds het teken zelf aan te passen en anderzijds door de omgeving aan te passen.<sup>1</sup> Dus bijvoorbeeld door een enkele <o> te schrijven in *stok* en een dubbele <k> in *stokken*, of door een dubbele <o> te schrijven in *stook* en een enkele <k> in *stoken*. Gevorderde spellers hebben niet veel problemen met deze spellingsregels, maar dat betekent volgens De Schryver en Neijt niet dat de regels simpel zijn.

In tegenstelling tot gevorderde spellers leveren de verdubbelings- en verenkelingsregels voor jonge onervaren spellers moeilijkheden op. Hoe het leren spellen globaal verloopt, legt Ghesquière (1997, p. 71) overzichtelijk uit. Hij stelt allereerst dat het Nederlands een aantal basisprincipes kent. Bij het leren spellen beginnen de leerlingen met de klankzuivere woorden volgens het fonologisch principe, ook wel alfabetisch principe genoemd. Dit doen zij door middel van de auditieve strategie waarbij elke klank wordt gerepresenteerd door een letter: de letter-klankkoppeling is dus één-op-één. In deze fase leren de leerlingen dat [p] [O] [t] gespeld moet worden als <pot> en [p] [o] [t] als <poot>.<sup>2</sup> Wanneer zij dit onder de knie hebben, gaan ze verder met de regelstrategie. Hierbij gaat het bijvoorbeeld om *hond* met een <d> op het eind, omdat het meervoud *honden* ook met een <d> wordt geschreven, maar *hij wordt* moet met <dt> worden geschreven, omdat hier de regel ‘stam + t’ geldt. Tot slot gaat het etymologisch beginsel een rol spelen in de spellingontwikkeling. Dit beginsel gaat uit van de herkomst van woorden. Zo schrijven wij nu nog steeds *zei* en *zij* verschillend, omdat deze vroeger in het Nederlands een verschillende uitspraak hadden.

De eerste fase die Ghesquière (1997) noemt (fonologisch/alfabetisch principe), sluit aan bij het zogenaamde ‘single-phase model of alphabetic and phonemic awareness’ zoals Neijt, Peters en Zuidema (2012) die beschrijven. Dit fasemodel laat zien dat er eerst één-op-éénmapping plaats-

vindt en dat de kinderen leren dat dubbele klinkerletters voor lange klanken geschreven worden en zo dus bevestigd worden in het idee dat grafemen en fonemen één-op-één gekoppeld zijn. Dit laatste zou het maken van fouten bij verenkelling en verdubbeling juist in de hand kunnen werken, omdat de leerlingen bijvoorbeeld het woord *potten* in de klanken [p] [O] [t] [a] [n] opdelen en voor elke klank een letter schrijven en zo bij de verkeerde spelling <poten> uitkomen. 3 + 4

Er zijn nog meer factoren die ervoor kunnen zorgen dat het lijkt alsof leerlingen de regels niet beheersen als ze fouten maken, zoals een foutieve uitspraak van woorden. Zo hebben Kloots, De Schutter, Gillis en Swerts (2003) laten zien dat zelfs docenten de <a> in bijvoorbeeld *banaan*, *matroos* en *manier* uitspreken als een gedekte [ɑ], wat illustreert dat een foute schrijfwijze als *mannier* niet alleen een gevolg hoeft te zijn van foutieve regeltoepassing, maar dat het ook door een foute (of andere: dialectisch/regionaal bepaalde) uitspraak veroorzaakt kan zijn.<sup>5</sup> Daarnaast kan variatie in uitspraak ook voor fouten zorgen in de spelling. Zo hebben Schiffelers, Bosman en Van Hell (2002) gevonden dat bij het leren van bijvoorbeeld dicteewoordjes de uitspreken-wat-er-staat-methode (Ormrod & Jenkins, 1989) voor woorden met inconsistente foneem-grafeemrelaties (spelling komt niet overeen met de algemene Nederlandse uitspraak) effectiever is gebleken dan de lezen-zoals-het-hoort-methode. Weinig kinderen gebruiken uit zichzelf de uitspreken-wat-er-staat-methode, hoewel deze studiestrategie als enige significant positief gecorreleerd werd met de posttestscores. Hieruit kun je afleiden dat kinderen over het algemeen de woorden schrijven zoals ze uitgesproken zouden moeten worden, tenzij je anders aanleert. Kinderen schrijven dus de woorden zoals ze ze uitspreken, waarbij ze bij het leren lezen en schrijven hebben geleerd dat een enkele klinker een korte klank representeert en een dubbele klinker gelijk is aan een lange klank. Dit geeft de verwachting dat vooral de jongere kinderen veel spelfouten blijven maken in woorden waar de letter-klankkoppeling niet één op één is en dat dit onder andere voor problemen zorgt in woorden met verenkelling en verdubbeling.

Zoals de resultaten uit de eerdergenoemde onderzoeken (o.a. Verhoeven, 1979; Booij e.a., 1979) laten zien, is het aantal fouten bij medeklinkerverdubbeling niet altijd gelijk aan het aantal klinkerverenkelingsfouten. Zo hebben Vanderlocht en Ghesquière (1997, p. 92) gevonden dat er beduidend meer fouten gemaakt worden in de medeklinkerverdubbeling dan in de klinkerverenkelling. Bij deze resultaten is echter een aantal kanttekeningen te plaatsen, zoals dat dit onderzoek onder Vlaamse proef-

personen heeft plaatsgevonden en niet onder Nederlandse. Daarnaast is het niet duidelijk of deze uitkomsten significant zijn of niet. Toch zou dit een aanwijzing kunnen zijn dat het verenkelen en verdubbelen niet met elkaar hoeft samen te hangen.

### 1.1.1 Verenkeling en verdubbeling

Voor de beschrijving van de fouten in verenkeling en verdubbeling is het overzichtelijker om één term te gebruiken. Daartoe is er zowel naar Vanderlocht en Ghesquière (1997) als naar Ruijssenaars, Claes en Caluwe (1992) gekeken. Eerstgenoemden maken namelijk duidelijk wanneer er van verdubbeling gesproken wordt en wanneer van verenkeling. Dit onderscheid is niet altijd duidelijk, want het ligt eraan van welke vorm je uit gaat. Is *maat – maten* klinkerverenkeling of is *maten – maat* klinkerverdubbeling en is *put – putten* medeklinkerverdubbeling of is *putten – put* medeklinkerverenkeling? Vanderlocht en Ghesquière (1997, p. 84) geven duidelijk aan dat je van medeklinkerverdubbeling spreekt in de variant *put – putten* en van klinkerverenkeling in de variant *maat – maten*. Wanneer er over verdubbeling gesproken wordt, gaat het in het huidige onderzoek dus over medeklinkerverdubbeling (en niet *maten – maat*: dus klinkerverdubbeling) en bij verenkeling gaat het om de klinkerverenkeling (en niet *putten – put*: medeklinkerverenkeling).

Als er tegelijkertijd wordt gesproken van medeklinkerverdubbelingsfouten en klinkerverenkelfouten wordt de term mvkv-fouten (Ruijssenaars, Claes, & Caluwe, 1992) gebruikt (**medeklinkerverdubbeling-en-klinkerverenkeling-fouten**). Het gaat dan niet specifiek om verenkelen of verdubbelen van letters, maar om beide problemen tegelijk. Zo hebben Ruijssenaars e.a. (1992) een onderzoek gedaan naar mvkv-fouten en niet naar verenkelfouten of verdubbelingsfouten apart. Beide spellingsonderwerpen worden in dat geval dus samen genomen en als één onderwerp gezien: <pooten> voor *poten* en <poten> voor *potten* zijn dus mvkv-fouten, terwijl <pooten> op zichzelf een verenkelfout is en <poten> een verdubbelingsfout.

In het huidige onderzoek wordt aan de hand van alleen medeklinkerverdubbeling onderzocht wat de beste manier is om spellingvaardigheid en spellingontwikkeling in kaart te brengen. Naast de medeklinkerverdubbeling en klinkerverenkeling is er nog een tweetal verenkelfouten- en verdubbelingsgevallen. Er is namelijk sprake van klinkerverdubbeling in

*auto-autootje* en medeklinkerverenkeling in *hij eett – eet* (Neijt, 2002). Deze vormen worden in huidig onderzoek buiten beschouwing gelaten.

### **1.1.2 Spelfoutenclassificatiesysteem**

Om spelfouten te classificeren zijn er talloze classificatiesystemen bedacht door verschillende onderzoekers uit verschillende vakgebieden (Dumont, 1993; Ojeman, 1978; Horbach-Kleijnen, 1988; Bolle, Simoens & Vermiert, 1990). Gezien het beschikbare corpus is de categorieverdeling van Cranshoff en Zuidema (2010) de beste keuze, omdat er in het corpus beschrijvingen van woorden beschikbaar zijn die op die categorieën gebaseerd zijn. Daarnaast kan er op basis van het classificatiesysteem van Cranshoff en Zuidema (2010) van een globale naar een steeds fijnmazigere verdeling gewerkt worden. In de *Van Dale Basisspellinggids* beschrijven Cranshoff en Zuidema (2010) veertien hoofdcategorieën in spelling. Elke hoofdcategorie is onderverdeeld in kleinere categorieën, in totaal 75. In het huidige onderzoek wordt er ingezoomd op de categorie 4.2: ‘gesloten lettergreep, gevolgd door lettergreep met stomme e’, want deze categorie beschrijft de medeklinkerverdubbeling die in dit onderzoek ter illustratie gebruikt wordt voor de benodigde mate van detail in de foutenanalyses. Deze categorie zal in het vervolg *inheemse verdubbelingswoorden* genoemd worden.

Vervolgens zijn in de *Van Dale Basisspellinggids* sommige categorieën, waaronder de inheemse verdubbelingswoorden, onderverdeeld in subcategorieën die gelabeld zijn door letters (4.2a, 4.2b enz.). Het is niet duidelijk of een ‘hogere’ categorie (bijvoorbeeld 4.2 ten opzichte van 4.1 en 4.2a ten opzichte van 4.2b) als moeilijker wordt bestempeld door de auteurs van de *Basisspellinggids*, maar dit is wel aannemelijk. Zie ook figuur 1.

4	Open en gesloten lettergreep	
4.2	gesloten lettergreep, gevolgd door lettergreep met stomme e	<p>a gesloten: meervoud of werkwoord op <i>-en</i>, vormen op <i>-e</i>, verkleinwoord op <i>-etje</i></p> <ul style="list-style-type: none"> <li>• schrappen, tellen, klimmen, poppen, plukken, verpakken, bestellen, godinnen, felle, tamme, kannetje, bommetje, hindernissen</li> </ul> <p>b gesloten: geen meervoud</p> <ul style="list-style-type: none"> <li>• binnen, nummer, effen, korrel, gemiddelde, verbitterd, dwarrelen, babbelen, bibberen</li> </ul> <p>c gesloten: woorden op <i>-ig</i>, <i>-ing</i>, <i>-lijk</i></p> <ul style="list-style-type: none"> <li>• sappige, pittig, verschrikkelijk, vergissing</li> </ul>

Figuur 1. Spellingcategorieën 4.2 uit de Van Dale Basisspellinggids

Subcategorie 4.2a noemen we de zogenaamde *tamme woorden*. Het zijn meervoudsvormen en werkwoorden op *-en*, diminutief op *-etje* en vormen op *-e*. Deze woorden zijn de ‘makkelijke’ (tamme) woorden, waarmee bedoeld wordt dat ze geen uitzonderingen bevatten en dat het gaat om de meest voor de hand liggende verdubbelingen. De volgende subcategorie is 4.2b en deze bevat de woorden op *-el* en *-er* die bij de basis horen en geen onderdeel zijn van de flectionele uitgang. Deze woorden zullen in het vervolg *korrel-woorden* genoemd worden. De *pudding-woorden* zijn de woorden die onder subcategorie 4.2c vallen. Het zijn woorden op *-ing*, *-ig* en *-lijk*, waarbij de stomme *e* niet als <e> wordt geschreven. De laatste categorie staat (nog) niet in de *Van Dale Spellinggids*. Deze subcategorie is later toegevoegd, omdat het bij het analyseren van de woorden zinvol bleek te zijn om een vierde categorie te onderscheiden. Deze lijkt sterk op de andere verdubbelingscategorieën, maar verschilt daarvan doordat er meer dan één de klemtoon aanwezig is, zoals *luilakken* en *hindernissen*, en/of doordat de medeklinkers worden gevolgd door een volle vocaal in plaats van een sjwa, zoals *makkie* of *herrie* (Zuidema, p.c.). De woorden in deze subcategorie zullen voor het gemak *herrie-woorden* genoemd worden. In tabel 1 staat een uitgebreid overzicht van de beschrijving en benaming van de categorieën.

Categorie	Beschrijving Basisspellinggids	Voorbeelden	Benaming
4.2	Inheemse woorden met een gesloten lettergreep gevolgd door stomme e		Inheemse verdubbelingswoorden
4.2a	meervoud en werkwoord op <i>-en</i> , diminutief op <i>-etje</i> en vormen op <i>-e</i> , dus woorden met een suffix dat een stomme <i>e</i> bevat die je als <e> schrijft	tellen, tamme, bommetje	Tamme woorden
4.2b	<i>-el</i> , <i>-er</i> die bij de basis horen, geen flectionele suffixen, wel een stomme <i>e</i> die je als <e> schrijft	nummer, korrel, bibberen	Korrel-woorden
4.2c	Woorden op <i>-ing</i> , <i>-ig</i> , <i>-lijk</i> , dus een stomme <i>e</i> die niet als <e> geschreven wordt.	pittig, pudding, verschrikkelijk, vergissing	Pudding-woorden
4.2d	Inheemse verdubbelingswoorden die net even anders zijn dan bovenstaande woorden, omdat ze meer dan één volle vocaal bevatten.	makkie, herrie, luilakken, hindemissen	Herrie-woorden

Tabel 1. Overzicht beschrijving en benaming categorieën 4.2, 4.2a, 4.2b, 4.2c en 4.2d

De genoemde subcategorieën zijn vervolgens verder onder te verdelen in zogenaamde ‘tripletten’ (Zuidema & Neijt, 2012). Elk triplet beschrijft een onderdeel van een woord, waarbij zowel de orthografische weergave als de uitspraak wordt beschreven. Ook verwijst elk triplet naar de betreffende spellingregels en -categorieën uit de *Van Dale Spellinggids*.<sup>6</sup> Zuidema en Neijt (2012) hebben in hun artikel over de wenselijkheid en haalbaarheid van een verrijking van de Woordenlijst Nederlandse Taal de tripletten als nieuw concept geïntroduceerd. De auteurs zien deze tripletten als de ‘moleculen’ van de Nederlandse spelling en elk woord bestaat uit meerdere tripletten. Ze geven de relatie weer tussen spelling, uitspraak en regel. Schematisch ziet een triplet er zo uit: {1 ; 2 ; 3}, waarbij { } staan voor het begin/einde van het triplet en ; als scheiding dient tussen de spelling (1), uitspraak (2) en bijbehorende regel (3).

Het woord *koets* zou als volgt in een triplettenreeks weergegeven kunnen worden:

{ k ; k ; - } { oe ; u ; - } { t ; t ; - } { s ; s ; - }. In dit geval zijn alle letters klankzuiver, wat betekent dat de letter (spelling) op de eerste plaats en klank op de tweede plaats in het triplet overeenkomen. Het streepje op de derde plaats in het triplet geeft aan dat het hier om een spellingzuivere klank-letterverbinding gaat en er dus geen regel aan verbonden is. In



tabelvorm (tabel 2), van boven naar onder gelezen, is duidelijk te zien hoe de tripletten zijn gevormd. Deze tabel zou je dus voor elk woord kunnen invullen.

<i>koets</i>				
begin triplet	{	{	{	{
spelling	k	oe	t	s
scheidingsteken	:	:	:	:
uitspraak	k	u	t	s
scheidingsteken	:	:	:	:
regel	-	-	-	-
einde triplet	}	}	}	}
	{k;k;-}	{oe;u;-}	{t;t;-}	{s;s;-}

Tabel 2. Triplettenopbouw 'koets'

Wanneer je het woord *willen* en *kunnen* in tripletten schrijft, lijkt het op het eerste gezicht erg ingewikkeld: {w;[w;-]{il=1;I]=1;4.2a}{en;@n;2.2} en {k;k;-}{un=n;Y]=n;4.2a}{en;@n;2.2}. Bouw je deze tripletten echter op vanuit een tabel, zoals hierboven bij *koets*, dan kun je zien dat het juist heel overzichtelijk en duidelijk is. Let daarbij op de andere tekens die kunnen voorkomen in tripletten:

[...]klemtoon  
=syllabe- en lettergreepgrens

	<i>willen</i>			<i>kunnen</i>		
begin triplet	{	{	{	{	{	{
spelling	w	il=1	en	k	un=n	en
scheidingstek en	:	:	:	:	:	:
uitspraak	[w	I]=1	@n	k	Y]=n	@n
scheidingstek en	:	:	:	:	:	:
regel	-	4.2a	2.2	-	4.2a	2.2
einde triplet	}	}	}	}	}	}
	{w;[w;-}	{il=1;I]=1;4.2a}	{en;@n;2.2}	{k;k;-}	{un=n;Y]=n;4.2a}	{en;@n;2.2}

Tabel 3. Triplettenopbouw 'willen' en 'kunnen'

Door het gebruik van deze tripletten, wordt duidelijk dat deze twee woorden in spelling eigenlijk overeenkomen, ook al delen ze weinig letters en klanken. De tripletten zijn hetzelfde opgebouwd, al is het met verschil-

lende letters en klanken en ze verwijzen naar dezelfde regels uit de *Van Dale Basisspellinggids*.

In hoeverre dit steeds specifiekere wordende onderscheid nodig is, wordt onderzocht in dit artikel. Er wordt gekeken tot op welk detailniveau de spellingscategorieën opgesplitst moeten worden om nog zinvol te zijn voor het beschrijven van de spellingvaardigheid en spellingontwikkeling.

## 1.2 *Probleemstelling*

In dit onderzoek wordt er onderzocht in hoeverre het opdelen van de spelling in categorieën, subcategorieën en tripletten (clusters) bijdraagt aan een ‘goede’ beschrijving van de spellingvaardigheid en –ontwikkeling, met in het bijzonder aandacht voor de inheemse verdubbelingswoorden. Er wordt dus niet gekeken naar de soort categorie-indeling, maar naar de mate van detail. De hoofdvraag die hier bij hoort is:

*Hoeveel detailinformatie is er nodig om de spellingvaardigheid en –ontwikkeling van het verdubbelen te beschrijven?*

Om te onderzoeken hoeveel detailinformatie er nodig is, wordt aan de hand van de volgende deelvragen de categorieverdeling steeds gedetailleerder.

1. *Welke informatie geeft categorie 4.2 en is die (te) gedetailleerd (genoeg)?*
2. *Welke informatie geven de subcategorieën 4.2a, 4.2b, 4.2c en 4.2d en is die (te) gedetailleerd (genoeg)?*
3. *Welke informatie geven de triplettenclusters van type 4.2 en is die (te) gedetailleerd (genoeg)?*

Aangezien er in de gebruikte data een paar honderd verschillende tripletten zijn van de categorie 4.2, is het onmogelijk om alle tripletten apart te analyseren. Daarom is ervoor gekozen om de tripletten te analyseren in clusters. Er zijn verschillende manieren om de tripletten te clusteren, zoals het samennemen van tripletten met vormelijke overeenkomsten, maar in dit onderzoek wordt er gebruik gemaakt van clusters die gebaseerd zijn op frequentie. Dit leek een logische verdeling, omdat de verwachting is dat tripletten die vaak voorkomen vaker goed geschreven worden dan tripletten die niet frequent zijn.

## 2 Methode

### 2.1 Materiaal

In dit onderzoek is er gebruik gemaakt van de data van een onderzoeksproject dat reeds aan de Radboud Universiteit plaatsvindt.<sup>7</sup> In dit project wordt een databank aangelegd van teksten die geschreven zijn door basisschoolleerlingen. De databank is nog in ontwikkeling en zal uiteindelijk longitudinale data bevatten van leerlingen uit groep 4 tot en met 8 van reguliere en speciale scholen door heel Nederland. Op het moment dat dit onderzoek werd uitgevoerd, was de databank nog niet klaar, maar werd er een gedeelte van de verzamelde data ter beschikking gesteld. De teksten die zijn gebruikt voor dit onderzoek zijn de eerste teksten die zijn ingevoerd en geanalyseerd. Ze zijn geschreven in de eerste helft van het schooljaar, eind 2012, en zijn omgezet naar kolommen in Excel. Naast elk door een kind geschreven woord ('er staat'), is de spellingcorrecte versie ('moet zijn') van het woord gezet, zodat automatische analyse van de teksten mogelijk werd.

Aan de hand van de spellingcorrecte woorden zijn er achter het merendeel van de werkwoorden en nomina automatisch de bijbehorende tripletten gezet, want op het moment van onderzoek zijn er alleen tripletten voor deze twee woordsoorten beschikbaar. Woorden als *bakken*, *batterijen* en *appel* krijgen dan onderstaande tripletten, waarbij handmatig met een rode kleur is aangegeven welke woorden fout zijn geschreven en in welke triplet(ten) de fout zit.<sup>8</sup> In tabel 4 zijn voor de duidelijkheid de fouten aangegeven met grijze arcering.

'er staat'	'moet zijn'	triplet 1	triplet 2	triplet 3	triplet 4
batanje	battenjen	"b?[[b?-"	"at=t?A]=t?4.2b"	"e=nj=?@[r3]=?6.5&5.2"	"en?@n?2.2"
bakken	bakken	"b?[b?-"	"ak=k?A]=k?4.2a"	"en?@n?2.2"	
apol	appel	"ap=p?[A]=p?4.2b"	"el?@l?2.2"		

Tabel 4. Voorbeeld triplettoekenning

Het corpus bestaat uit 54017 items die geschreven zijn door in totaal 280 verschillende leerlingen. De leerlingen zijn verspreid over de groepen 4 tot en met 8. Iets minder dan de helft van alle items is voorzien van tripletten (25729 items). Een overzicht van het aantal leerlingen per leerjaar is te vinden in tabel 5 in de kolom 'totaal'.

**Tabel 5 – Per leerjaar het aantal leerlingen en het aantal leerlingen dat woorden schrijft van categorie 4.2.**

	aantal leerlingen	
	<i>totaal</i>	<i>categorie 4.2</i>
groep 4	30	25
groep 5	37	27
groep 6	82	79
groep 7	71	67
groep 8	60	60
<b>totaal</b>	<b>280</b>	<b>258</b>

*Tabel 5. Per leerjaar het aantal leerlingen en het aantal leerlingen dat woorden schrijft van categorie 4.2.*

Uit het totale bestand zijn alle, zowel goed als fout geschreven, items verzameld die onder de inheemse verdubbelingswoorden vallen (gesloten lettergreep gevolgd door lettergreep met stomme *e*). Bij deze selectie bleken er in totaal 22 leerlingen te zijn afgevallen, omdat zij geen enkel inheems verdubbelingswoord (categorie 4.2) hadden geschreven, dus er zijn 258 leerlingen overgebleven voor dit onderzoek. De leerlingen komen uit groep 4 tot en met 8. De verdeling van het aantal leerlingen per leerjaar dat woorden uit categorie 4.2 heeft geschreven, staat weergegeven in tabel 3 in de kolom ‘categorie 4.2’.

Groep 4 en 5 zijn ongeveer even groot, maar twee tot drie keer kleiner dan groep 6 tot en met 8. Groep 6 is de grootste groep met 79 leerlingen, gevolgd door groep 7 met 67 leerlingen, groep 8 met 60 leerlingen en tot slot groep 5 en 4 met respectievelijk 27 en 25 leerlingen.

## 2.2 *Goedpercentage*

In de literatuur wordt meestal enkel het aantal fouten geteld, maar in tegenstelling tot voorgaande onderzoeken wordt het aantal fouten in dit onderzoek omgezet in *goedpercentages*. Aangezien het de bedoeling is dat leerlingen een stijgende lijn laten zien in hun ontwikkeling, ligt de keuze voor goedpercentages in plaats van *foutpercentages* voor de hand. Per leerjaar wordt het goedpercentage berekend voor de verschillende categorieverdelingen.<sup>9</sup> Vervolgens worden de uitkomsten met elkaar vergeleken waardoor het duidelijk wordt hoe gedetailleerd de categorieverde-

ling moet zijn om de spellingvaardigheid en -ontwikkeling te kunnen beschrijven.

### **2.3 Procedure**

Om de eerste vraag te beantwoorden, worden alle items verzameld die onder de inheemse verdubbelingswoorden vallen. Vervolgens wordt het aantal gemaakte fouten geteld en het goedpercentage berekend. De uitkomsten (goed én fout) worden per jaarlaag gescheiden zodat er vergelijking mogelijk is en er een eventueel ontwikkelingspatroon verschijnt. Geeft het goedpercentage met dit globale onderscheid (categorie 4.2) genoeg informatie?

Vervolgens wordt er voor de tweede deelvraag een opdeling gemaakt binnen de categorie 4.2: tamme woorden (4.2a), korrel-woorden (4.2b), pudding-woorden (4.2c) en herrie-woorden (4.2d), in totaal dus vier stuks. De percentages goed en fout worden per subcategorie en jaarlaag vergeleken, waardoor er mogelijk een ontwikkelingspatroon te ontdekken is. Verschilt de ontwikkeling van percentages goed en fout per subcategorie? Wanneer er geen patroon ontdekt wordt, kan dit betekenen dat het niet zinvol is om de inheemse verdubbelingswoorden verder op te delen in subcategorieën. Is er genoeg informatie om een ontwikkeling te laten zien of is een globalere of fijnmazigere categorieverdeling (4.2a, b, c en d) beter?

De laatste deelvraag wordt beantwoord door de tripletten te verdelen in clusters op basis van frequentie. Aan de hand van de frequentiegegevens (beschikbaar gesteld door J. Zuidema) zijn er drie groepen gemaakt: hoogfrequente tripletten, middenfrequente tripletten en laagfrequente tripletten. De verdeling is zo gemaakt, dat de hoog- en laagfrequente tripletten met grote afstand van elkaar verschillen. De grens voor laagfrequente tripletten is gelegd op tien of minder voorkomens en de grens voor hoogfrequente tripletten op meer dan honderd voorkomens. Het aantal items binnen deze drie triplettenclusters in het corpus met inheemse verdubbelingswoorden is weergegeven in tabel 6.

	<i>Frequentie (lijst Zuidema)</i>	<i>Aantal items</i>
Laag	1 – 10	377
Midden	11 – 99	942
Hoog	100+	368

Tabel 6. Aantal items per frequentiegroep

De items in de hoog- en laagfrequente triplettenclusters zijn in aantal ongeveer even groot. De middenfrequente groep is bijna tweeënhalf keer zo groot.

Per leerjaar wordt het goedpercentage berekend voor elk van de drie triplettenclusters. Zo wordt het mogelijk om zowel de verschillende frequentiegroepen als de resultaten per leerjaar met elkaar te vergelijken. Is er genoeg of te veel informatie om een ontwikkeling te beschrijven?

### 3 Resultaten

Tijdens het analyseren van de data kwam er een aantal lastige gevallen naar voren. Aangezien in dit onderzoek niet gekeken wordt naar andere (spel)fouten dan die uit categorie 4.2, zijn woorden met fouten in andere categorieën goed gerekend. De fout zat in deze gevallen niet in het verdubbelen, dus is als fout niet relevant voor dit onderzoek. Woorden zoals *lopenm* (lopen), *exame* (examen), *voor#tel#len* (vertellen) tellen dus mee bij het goedpercentage. 10 Woorden die meerdere fouten bevatten, waaronder een verdubbelingsfout, zijn eveneens meegeteld als fout, bijvoorbeeld: *gelaazen* (geblazen) en *8#talen* (achttallen). Woorden als *gegeet* (gegeten) zijn eruit gelaten, omdat het onduidelijk is hoe ze geanalyseerd moeten worden, want daar lijkt niet iets fout te zijn gegaan bij het verdubbelen, maar lijkt er eerder sprake te zijn van een dialectvorm of iets anders en is het moeilijk om iets te zeggen over de verdubbeling. Gevallen als *hleren* (kleren) zijn er wel in gelaten en goed gerekend, want de twijfel zit niet in het relevante deel. Daarbij lijken de geschreven *h* en *k* veel op elkaar en kan het mogelijk verkeerd getranscribeerd zijn. Er zijn geen woorden gevonden waarin twee of meer verdubbelingsfouten zijn gemaakt in één woord.

Uiteindelijk zijn er 1687 items van inheemse verdubbelingswoorden overgebleven voor dit onderzoek. De verdeling van het aantal leerlingen en het aantal items over de leerjaren is weergegeven in tabel 7.

<i>groep</i>	<i>aantal leerlingen</i>	<i>items 4.2</i>
4	25	93
5	27	126
6	79	842
7	67	301
8	60	325
<b>totaal</b>	<b>258</b>	<b>1687</b>

Tabel 7. Verdeling van het aantal leerlingen en items per leerjaar voor leerlingen die 4.2-woorden schrijven

Tabel 7 laat zien dat groep 6 met grote afstand het hoogste aantal items heeft, maar dat er in groep 6 ook meer leerlingen zitten dan in de andere leerjaren. De uitkomsten behorend bij de verschillende deelvragen zullen nu achtereenvolgens behandeld worden.

### 3.1 Deelvraag 1

Om te kunnen onderzoeken welke informatie de analyse van inheemse verdubbelingswoorden oplevert, is er per leerjaar geteld hoeveel van deze woorden er gebruikt zijn. Vervolgens is aan de hand van het aantal fouten uitgerekend wat het goedpercentage is. De resultaten staan weergegeven in tabel 8.

	<i>totaal 4.2</i>	<i>aantal 4.2 fout</i>	<i>goedpercentage 4.2</i>
groep 4	93	47	49%
groep 5	126	68	46%
groep 6	842	88	90%
groep 7	301	25	92%
groep 8	325	8	98%
<b>totaal</b>	<b>1687</b>	<b>236</b>	<b>86%</b>

Tabel 8. Foutenanalyse inheemse verdubbelingswoorden (categorie 4.2)

Zowel in tabel 7 als in tabel 8 is te zien dat groep 6 veel meer items en leerlingen bevat dan de andere groepen. Om deze verschillen te kunnen duiden, staat in tabel 9 per leerjaar het gemiddelde totaal aantal geschreven woorden per leerling ('totaal woorden'), het gemiddelde

aantal geschreven inheemse verdubbelingswoorden per leerling ('type 4.2 woorden') en het percentage inheemse verdubbelingswoorden ten opzichte van het totaal aantal woorden ('% 4.2). De getallen onder 'totaal woorden' en 'type 4.2 woorden' zijn absolute getallen, onder '% 4.2' staan uiteraard percentages.

Gemiddeld per leerling			
	<i>Totaal woorden</i>	<i>Type 4.2 woorden</i>	<i>% 4.2</i>
groep 4	125	3,10	2,5%
groep 5	112	3,41	3,0%
groep 6	268	10,27	3,8%
groep 7	185	4,24	2,3%
groep 8	213	5,42	2,5%

Tabel 9. Gemiddeld aantal woorden (totaal en 4.2) per leerling

Tabel 9 laat zien dat er in groep 6 gemiddeld per leerling 10,27 inheemse verdubbelingswoorden worden gebruikt, terwijl het gemiddelde van groep 4 en 5 rond de drie woorden ligt en in groep 7 en 8 ligt het gemiddelde tussen de vier en ruim vijf woorden. In groep 6 schrijft de leerling gemiddeld ook meer woorden in totaal, maar dit verschilt niet significant met bijvoorbeeld groep 8. Aangezien de tekstomvang en dus het aantal gebruikte woorden per leerling varieert, zeggen de absolute getallen weinig. Daarom is het gemiddelde percentage inheemse verdubbelingswoorden ten opzichte van het totaal aantal geschreven woorden per leerling berekend. Het aantal geschreven inheemse verdubbelingswoorden in verhouding met het totaal aantal gebruikte woorden per leerling verschilt niet enorm: bij alle leerjaren ligt het gemiddelde tussen de 2,3% en 3,8%.

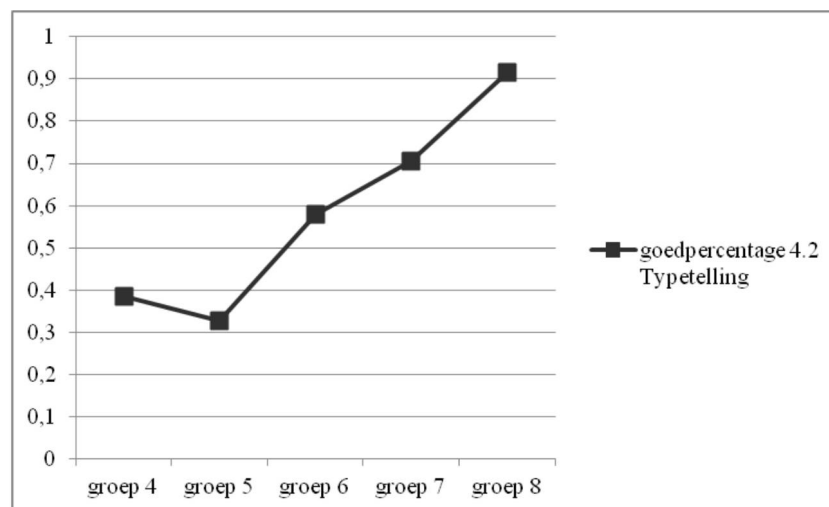
Het gemiddelde percentage inheemse verdubbelingswoorden ligt in groep 6 het hoogst en kan mogelijk verklaard worden doordat er in groep 6 heel vaak de woorden *hebben* en *kunnen* gebruikt zijn. Zo wordt *kunnen* 244 keer geschreven en *hebben* bijna 200 keer. Om de invloed van het gebruik van veel dezelfde woorden uit te sluiten, is er behalve een tokentelling, ook een typetelling uitgevoerd, zie tabel 10.



	Typetelling per leerling		
	<i>totaal 4.2</i>	<i>aantal 4.2 fout</i>	<i>goed% 4.2</i>
groep 4	39	24	38%
groep 5	55	37	33%
groep 6	112	47	58%
groep 7	78	23	71%
groep 8	84	7	92%

Tabel 10. Typetelling inheemse verdubbelingswoorden (4.2) per leerling

De uitschieter van groep 6 is door de typetelling een stuk kleiner geworden en de goedpercentages liggen lager dan die van de tokentelling. De type- en tokentelling verschillen echter zo weinig, dat alleen de resultaten van de typetelling gepresenteerd zullen worden. Ook omdat typetellingen in het algemeen het best reflecteren wat er in de hersenen gebeurt en we daar dus beter conclusies uit kunnen trekken. De gegevens uit tabel 10 zijn in grafiek 1 gezet.



Grafiek 1. Goedpercentage inheemse verdubbelingswoorden (4.2) per leerjaar

De grafiek geeft, behalve bij groep 5, een gelijkmatig stijgende lijn weer. Het verschil bij de typetelling tussen groep 4 en 5 is niet significant bevonden (t-toets) en er kan dus in geen geval gesproken worden van een U-bocht.

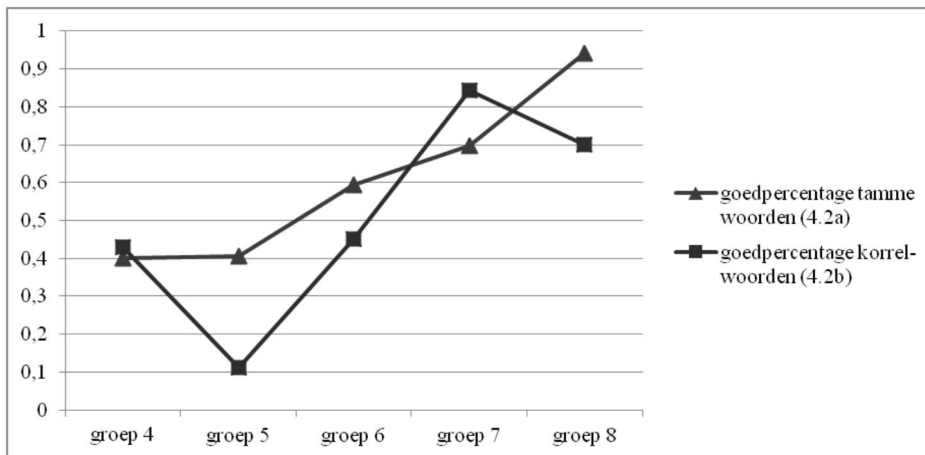
### 3.2 Deelvraag 2

Van de tamme (4.2a), korrel- (4.2b), pudding- (4.2c) en herrie-woorden (4.2d) is een typetelling gedaan.<sup>11</sup> De resultaten zijn weergegeven in tabel 12.

	totaal				aantal fout				goed%	
	4.2a	4.2b	4.2c	4.2d	4.2a	4.2b	4.2c	4.2d	4.2a	4.2b
groep 4	30	7	2		18	4	2		40%	43%
groep 5	42	9	2	2	25	8	2	2	40%	11%
groep 6	84	20	6	2	34	11	2	0	60%	45%
groep 7	53	19	2	4	16	3	2	2	70%	84%
groep 8	69	10	3	2	4	3	0	0	94%	70%

Tabel 12. Typetelling per leerjaar

Voor de pudding- en herrie-woorden (4.2c en d) waren te weinig waarnemingen om het percentage van te berekenen. De waarnemingen van tamme woorden (4.2a) en korrel-woorden (4.2b) zijn ook aan de lage kant, maar om een vergelijking te kunnen maken is er toch voor gekozen om percentages te berekenen. De gegevens uit tabel 12 zijn in grafiek 3 gezet.



Grafiek 3. Goedpercentage tamme woorden (4.2a) en korrel-woorden (4.2b), typetelling

In grafiek 3 zijn van de tamme woorden en korrel-woorden de goedpercentages van de typetelling te zien.<sup>12</sup> De daling tussen groep 7 en 8 is significant ( $t(159)=2,045$ ,  $p < .001$ ). Dat is een opvallend resultaat, aangezien er verwacht mag worden dat het goedpercentage in groep 8 het hoogst is en niet meer daalt.

De vergelijking tussen de inheemse verdubbelingswoorden en de onderverdeling in 4.2a, b, c en d is er nieuwe en nuttige informatie verkregen. Zo is het duidelijk geworden dat de pudding- en herrie-woorden (4.2c en d) te weinig voorkomen om iets over te zeggen en dat de tamme woorden en korrel-woorden (4.2a en b) een verschillende ontwikkelingslijn volgen. De tamme woorden worden over het algemeen beter gedaan dan de korrel-woorden.

### 3.3 Deelvraag 3

In tabel 14 staat het goedpercentage per leerjaar voor de laag-, midden- en hoogfrequente triplettenclusters.

	laagfrequent		middenfrequent		hoogfrequent	
	<i>totaal</i>	<i>aantal fout</i>	<i>totaal</i>	<i>aantal fout</i>	<i>totaal</i>	<i>aantal fout</i>
groep 4	3	1	16	10	19	12
groep 5	6	5	32	18	16	13
groep 6	6	3	68	25	37	20
groep 7	5	3	51	12	22	7
groep 8	7	0	41	6	35	2

Tabel 14. Typetelling frequenties

Allereerst valt het op dat bij deze analyse weinig waarnemingen zijn en dat heeft te maken met het (te) kleine corpus. In de tabel zijn dan ook geen goedpercentages te vinden. Het heeft geen toegevoegde waarde om de gegevens uit tabel 14 wel in percentages om te zetten of om de gegevens in een grafiek te presenteren.

## 4 Conclusie/Discussie

In dit onderzoek is er onderzocht hoeveel detailinformatie er nodig is om de spellingvaardigheid en –ontwikkeling te beschrijven. Aan de hand van één spellingonderwerp, namelijk medeklinkerverdubbeling, is dit onderzocht. Wat is de beste methode: een globale onderverdeling zoals categorie 4.2 (inheemse verdubbelingswoorden: gesloten lettergreep gevolgd door stomme *e*), een gedetailleerde verdeling in subcategorieën 4.2a (tamme woorden: meervoud en werkwoorden op –*en*, diminutief op –*etje* en vormen op –*e*), 4.2b (korrel-woorden: geen meervoud), 4.2c (pudding-woorden: woorden op –*ig*, –*lijk* en –*ing*) of 4.2d (herrie-woorden: lijkt op inheems, maar volle vocaal i.p.v. sjwa en/of mist klemtoon), of de fijnmazige verdeling in triplettencusters? Hieronder worden achtereenvolgens de deelvragen behandeld en wordt uiteindelijk antwoord gegeven op de centrale vraag.

### *1. Welke informatie geeft categorie 4.2 en is die (te) gedetailleerd (genoeg)?*

De informatie die verkregen wordt wanneer alleen de inheemse verdubbelingswoorden (4.2) worden onderscheiden, is overzichtelijk en geeft een ontwikkelingslijn weer, hoewel niet alle verschillen in goedpercentage tussen de leerjaren significant zijn. De ontwikkeling werd weergegeven in een lijngrafiek, maar er is geen sprake geweest van een longitudinaal onderzoek en per leerjaar is er maar één meetmoment geweest. Daarom is het belangrijk om vooral naar de meetpunten in de grafiek te kijken. Deze meetpunten laten zien dat er tussen groep 4 en 8 zeker sprake is van vooruitgang. In groep 8 hebben de leerlingen zelfs een goedpercentage van 98% gehaald.

Wat opvalt, is dat de inheemse verdubbelingswoorden maar een paar procent vormen van het totaal aantal woorden. De piek in het aantal inheemse verdubbelingswoorden in groep 6 is mogelijk te verklaren door de onderwerpen van de opstellen, want de woorden *kunnen* en *hebben* zijn elk goed voor ongeveer een kwart van het totaal aantal inheemse verdubbelingswoorden in groep 6. Dit wordt ondersteund door het feit dat bij de typetelling de piek is teruggebracht. De enorme stijging tussen groep 5 en 6 zou dan ook verklaard kunnen worden door het vele gebruik van die twee woorden en hoeft niet direct te maken te hebben met de inhoud van het spellingsonderwijs en het leren van medeklinkerverdubbeling.

Informatie van alleen de inheemse verdubbelingswoorden is niet genoeg, want het is niet duidelijk welk onderdeel van het verdubbelen (werkwoorden, meervoudsvormen, ‘moeilijke woorden’) er fout gaat, terwijl dit een belangrijk gegeven kan zijn, zeker voor het onderwijs. Daarnaast is er uit de informatie ook niet af te leiden of de leerlingen überhaupt een onderscheid maken in verschillende onderdelen van verdubbeling. Wanneer dus alleen categorie 4.2 (inheemse verdubbelingswoorden) wordt gebruikt om de spellingvaardigheid in kaart te brengen, lijkt er een ontoereikend beeld te worden gegeven. Categorie 4.2 is niet gedetailleerd genoeg.

Wat betreft de ontwikkeling in spelling lijkt er wel genoeg informatie te zijn, want uiteindelijk nadert het goedpercentage de 100%, wat betekent dat, ongeacht de verschillen binnen de inheemse verdubbelingswoorden, de verdubbeling bijna foutloos gaat in groep 8. Deze manier van analyseren brengt de ontwikkeling in zijn geheel, globaal, dus goed in kaart.

*2. Welke informatie geven de subcategorieën 4.2a, 4.2b, 4.2c en 4.2d en is die (te) gedetailleerd (genoeg)?*

De informatie die de subcategorieën 4.2a (tamme woorden), 4.2b (korrelwoorden), 4.2c (pudding-woorden) en 4.2d (herrie-woorden) geven is een stuk gedetailleerder dan de informatie van alleen de inheemse verdubbelingswoorden. Het is nu bijvoorbeeld duidelijk dat één subcategorie, de tamme woorden, het grootste aandeel heeft in het aantal verdubbelingsgevallen.<sup>13</sup> De korrelwoorden staan op de tweede plaats.<sup>14</sup> De pudding- en herrie-woorden komen in dit corpus nauwelijks voor.

De informatie die de onderverdeling in subcategorieën geeft, kan waardevol zijn, omdat hieruit af te leiden is dat het weinig rendement oplevert wanneer er in het onderwijs veel aandacht besteed wordt aan puddingwoorden en herriewoorden. Met name de verbetering van tamme woorden en korrelwoorden zal tot een hoger goedpercentage leiden. De tamme woorden gaan gemiddeld vaker goed dan korrelwoorden, wat ervoor pleit dat er aan de korrelwoorden extra aandacht besteed zou moeten worden om het goedpercentage van het verdubbelen nog meer te verhogen.

In eerste instantie leek de dalende lijn tussen groep 7 en 8 mogelijk een begin te zijn van een U-bocht die over de grenzen van het basis- en middelbaar onderwijs loopt of een gevolg van een plafondeffect, maar het

verschil tussen 7 en 8 is niet significant en ook de gemaakte fouten lijken niet te wijzen op een U-bochtontwikkeling. Mogelijk kan de aansluiting en overgang tussen basisonderwijs en middelbaar onderwijs in de toekomst onderzocht worden.

Wat betreft de vergelijking tussen inheemse verdubbelingswoorden (4.2) en de onderverdeling in de subcategorieën 4.2a, b, c en d, is er nieuwe en nuttige informatie verkregen. Geconcludeerd kan worden dat de kennis over tamme woorden niet automatisch gegeneraliseerd wordt naar korrel-woorden.

De spellingontwikkeling is goed te beschrijven door tamme woorden en korrel-woorden te onderscheiden. Ook de spellingvaardigheid is beter te beschrijven aan de hand van tamme woorden en korrel-woorden dan met pudding- en herrie-woorden, want die is gebaseerd op enkele waarnemingen en is niet betrouwbaar. Achter het gebruik van pudding-woorden en herrie-woorden lijkt geen logica te zitten. Dit type woorden lijkt namelijk relatief vaak fout te gaan en kan dus als ‘moeilijk’ gezien worden. Ondanks deze moeilijkheid is er aanwijzing gevonden dat deze woorden later in hogere klassen van de basisschool vaker gebruikt worden dan in de lagere klassen. Je zou namelijk verwachten dat moeilijke woorden pas later aangeboden worden en daardoor ook pas later steeds meer gebruikt zullen worden, maar dit is niet het geval in het corpus dat gebruikt is in dit onderzoek. Deze uitkomsten doen wel vermoeden dat de gekozen hiërarchie van Cranshoff en Zuidema (2010) in lijn ligt met de moeilijkheidsgraad of een zekere frequentie van het type woorden. De moeilijkheidsgraad die uit de categorie aanduiding volgt, klopt met de hier gepresenteerde bevindingen.

### *3. Welke informatie geven de tripletten (clusters) van type 4.2 en is die (te) gedetailleerd (genoeg)?*

De tripletten van de inheemse verdubbelingswoorden die op basis van frequentie geclusterd zijn, geven in grote lijnen dezelfde ontwikkeling weer als de verdeling op basis van de categorieën van Cranshoff en Zuidema (2010), maar er is een belangrijk verschil. De verwachting is namelijk dat hoogfrequente tripletten bekender zijn en daardoor op enige manier ‘makkelijker’ worden bevonden, maar uit dit onderzoek is gebleken dat de laagfrequente tripletten in groep 6 significant vaker goed geschreven worden. De hoogfrequente tripletten hebben in groep 6 en 7 een significant lager goedpercentage en dat is opvallend. Een duidelijke verklaring kan er niet gegeven worden, maar misschien is het aantal obser-

vaties te laag of is de verdeling op basis van Zuidema's frequenties geen goede categorisatie, in ieder geval niet voor het doel van dit onderzoek.

De informatie die de triplettenclusters van hoog-, midden- en laagfrequent geven, levert geen nieuwe inzichten op. Dit is vooral te wijten aan het feit dat het aantal waarnemingen aan de lage kant is, maar daar waren ook bij de andere categorieverdelingen problemen mee. In toekomstig onderzoek kunnen de tripletfrequenties nader onderzocht worden en bijvoorbeeld worden gekeken in hoeverre deze frequenties samenhangen met woordfrequenties.

De centrale vraag '*Hoeveel detailinformatie is er nodig om de spellingvaardigheid en –ontwikkeling van het verdubbelen te beschrijven?*' kan nu beantwoord worden. Het is duidelijk geworden dat alleen het onderscheiden van de inheemse verdubbelingswoorden (4.2) niet genoeg informatie geeft om de spellingvaardigheid van verdubbeling goed te kunnen beschrijven. Het vraagt enig voorbehoud om deze uitspraak ook te laten gelden voor de andere categorieën in de *Van Dale Basisspellinggids*, maar het is voor te stellen dat ook bij die andere categorieën een gedetailleerdere categorisering gewenst is voor het beschrijven de spellingvaardigheid en –ontwikkeling.

De subcategorieën geven een veel gedetailleerdere beschrijving dan alleen de verdubbelingswoorden en zijn met het oog op het onderwijs mogelijk voldoende onderverdeeld. Binnen de inheemse verdubbelingswoorden heeft de verdeling in tamme, korrel-, pudding- en herrie-woorden interessante informatie gegeven, want twee van de vier subcategorieën bleken nauwelijks gebruikt te worden in groep 4 tot en met 8. Het is duidelijk geworden welke onderdelen wezenlijk verschil kunnen maken in het goedpercentage. Daarbij zijn deze categorieën zodanig onderverdeeld, dat de onderwerpen per categorie zinvol en begrijpelijk zijn voor het onderwijs. De subcategorieën beslaan namelijk onderwerpen als 'meervoudsvormen op *-en*' en 'woorden op *-ig*, *-ing* en *-lijk*' en zijn makkelijk te vertalen naar het onderwijs en lesmethodes.

De meest gedetailleerde analysemethode, de verdeling in triplettenclusters, heeft een onduidelijk beeld gegeven. Door het (te) kleine corpus is niet van alle leerjaren betrouwbare informatie beschikbaar over de laagfrequente, middenfrequente en hoogfrequente triplettenclusters. De verwachting dat hoogfrequente tripletten eerder een hoger goedpercentage bereiken dan de andere twee clusters, is niet uitgekomen, want verrassend

genoeg bleken er in de hoogfrequente tripletten in een aantal gevallen juist meer fouten te worden gemaakt.

De eindconclusie is dat een meer gedetailleerde methode dan de onderverdeling in subcategorieën andere en nog waardevollere informatie kan opleveren, maar dan is een andere categorisatie waarschijnlijk beter dan de frequentie van triplettenclusters die gekozen zijn in dit onderzoek. Een onderverdeling op basis van woordfrequentie of een tripletfrequentie met een ander uitgangspunt geeft mogelijk andere (en betere?) informatie. Daarbij is het van belang om gebruik te maken van een veel groter corpus dan in dit onderzoek, omdat er in dit onderzoek voor bepaalde categorieën veel te weinig waarnemingen waren. Het is daarbij ook belangrijk om een verdeling aan te houden die op enige manier vertaald kan worden naar het onderwijs, want anders blijft het informatie die alleen betekenis heeft in het wetenschappelijke veld.

#### **4.1 Kritische noten**

Bij dit onderzoek bleek dat het corpus aan de kleine kant is, waardoor er voor een aantal variabelen te weinig waarnemingen zijn om conclusies aan te verbinden.<sup>15</sup> Ook zou een corpus met longitudinale data een uitkomst zijn, omdat een ontwikkeling dan pas echt te beschrijven is.

In hoeverre de in dit onderzoek verkregen informatie en conclusies te generaliseren zijn, is niet duidelijk. Toch geven de uitkomsten van dit onderzoek de verwachting dat een foutenanalyse op basis van de categorieën, zoals 4.2, 6.1 en 8.3, niet gedetailleerd genoeg is. Een analyse-methode die meer details geeft, zoals 4.2a, b, c en d, is waardevoller en geeft zinvollere informatie waar ook het onderwijs mee kan werken. Of de verdeling in tripletten(clusters) te gedetailleerd is, kan op basis van dit onderzoek niet met zekerheid gezegd worden.

#### **4.2 Vervolgonderzoek**

Voor vervolgonderzoek kan het interessant zijn om de uitkomsten van dit onderzoek, en de categorieverdeling van Cranshoff en Zuidema (2010) in het bijzonder, te vergelijken met verschillende schoolmethodes. Is de volgorde van de spellingonderwerpen in de schoolmethode hetzelfde als de categorieverdeling van Cranshoff en Zuidema? Zijn bijvoorbeeld de 'hoge' categorieën moeilijker en worden deze vaker of later behandeld?



In toekomstig onderzoek kan het ook interessant zijn om de inheemse verdubbelingswoorden te vergelijken met de andere categorieën uit hoofdcategorie 4 (4.1, 4.3 t/m 4.7). Deze verdeling van de categorieën zal zeker voor het onderwijs zinvolle informatie geven, omdat de categorieën ‘begrijpelijke’ onderwerpen beschrijven die eenvoudig te vertalen zijn naar het onderwijs en schoolmethodes. Daarnaast kan een onderzoek de verschillende categorieën van een hoger niveau uit de *Van Dale Basis-spellinggids* vergelijken (categorie 1 t/m 14).

Zoals eerder al werd aangegeven, kan een onderzoek over de grenzen van het basis- en middelbaar onderwijs waardevolle informatie opleveren. Het onderzoek wordt wel gecompliceerd doordat er op de basisschool over het algemeen geen onderscheid in niveau wordt gemaakt en op de middelbare school wel (vmbo, havo, vwo). Zeker gezien het hoge goedpercentage dat in dit onderzoek gehaald werd in groep 8 is het interessant om te zien hoe de niveaus zich verder ontwikkelen en of er überhaupt nog ontwikkeling zichtbaar is.

Bij de conclusie van deelvraag 3 werd al genoemd dat het nieuwe inzichten kan opleveren wanneer de tripletfrequenties en woordfrequenties worden vergeleken. Is er een verband tussen deze twee varianten of staan ze geheel los van elkaar? Betekent het bijvoorbeeld dat een triplet dat hoogfrequent is ook altijd in woorden zit met een hoge frequentie of kan het ook zijn dat hoogfrequente tripleten in laagfrequente woorden voorkomen?

#### Noten

\* Deze tekst is een voor de *Acta Neerlandica* aangepaste versie van mijn bachelorwerkstuk dat onder begeleiding van prof. dr. Anneke Neijt en dr. Johan Zuidema in 2013 is geschreven.

<sup>1</sup> IPA: [ɪ] [ʊ] [∅]

<sup>2</sup> IPA: [p] [ɔ] [t] en [p] [o] [t]

<sup>3</sup> IPA: [p] [ɔ] [t] [ə] [n]

<sup>4</sup> Eventueel geschreven als <potun>, maar het gaat hier om de enkel geschreven <ɒ>.

<sup>5</sup> Gedekte klinker: ook wel ‘korte’ klinker, deze wordt in dezelfde lettergreep gevolgd door een medeklinker.

<sup>6</sup> Zie methodesectie voor illustratie.

<sup>7</sup> In het BasiScript werkt de Radboud Universiteit samen met Tilburg University, de Universiteit Groningen en de Universiteit van Amsterdam. Het BasiScript project wordt gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onder-

zoek (NWO). Projectcoördinator is Dr. Agnes Tellings. (<http://www.basiscript.nl/content/>)

<sup>8</sup> Voor verdere uitleg over hoe tripletten in elkaar zitten: Zuidema & Neijt (2012).

<sup>9</sup> Goedpercentage =  $1 - (\text{aantal fout} / \text{totaal aantal woorden})$ .

<sup>10</sup> # staat voor een afbreking aan het eind van de regel of onterecht geschreven spatie.

<sup>11</sup> De resultaten worden niet anders als er wordt uitgegaan van een token telling.

<sup>12</sup> De token telling laat een vergelijkbare lijn zien, maar daar ligt de lijn over het algemeen iets hoger.

<sup>13</sup> 4.2a: meervoud of werkwoord op *-en*, vormen op *-e*, verkleinwoord op *-etje*.

<sup>14</sup> 4.2b: geen meervoud, wel gesloten lettergreep gevolgd door lettergreep met stomme *e*.

<sup>15</sup> Het corpus is op dit moment ook nog in ontwikkeling en zal met de tijd uitgebreid worden.

## Bibliografie

- Booij, G., Hamans, C., Van Minnen, C., Verhoeven, G. & Balk-Smit Duyzentkunst, F. 1979. *Spelling*. Groningen: Wolters-Noordhoff.
- Cranshoff, B. & Zuidema, J. 2010. *Van Dale Basisspellinggids*. Utrecht/Nederland: Van Dale Uitgevers.
- De Schryver, J. & Neijt, A. 2005. *Handboek Spelling*. Mechelen: Wolters Plantyn.
- Ghesquière, P. 1997. 'Spellingproblemen in het secundair onderwijs. Een exploratief onderzoek in Vlaanderen.' Ghesquière, P. & Ruijssenaars, A. *Leerproblemen in het middelbaar onderwijs*. Leuven: Acco, 69-82.
- Ghesquière, P. & Ruijssenaars, A. 1997. *Leerproblemen in het middelbaar onderwijs*. Leuven: Acco.
- Kloots, H., De Schutter, G., Gillis, S. & Swerts, M. 2003. 'Verdoffende vocalen en klinkers die verdwijnen: een casestudy.' *Nederlandse Taalkunde*, 8 (3): 231-254.
- Neijt, A. 2002. 'The Interfaces of Writing and Grammar.' Neef, M., Neijt, A. & Sproat, R. *The Relation of Writing to Spoken Language*. Tübingen: Niemeyer, 11-34.
- Neijt, A., Peters, M. & Zuidema, J. 2012. 'The 12321 model of Dutch spelling acquisition.' *Linguistics in the Netherlands*, 29: 111-112.
- Ormrod, J.E. & Jenkins, L. 1989. 'Study strategies for learning spelling: correlations with achievement and developmental changes.' *Perceptual and Motor Skills*, 68: 643-650.

- Ruijsenaars, A., Claes, E. & Caluwe, M. 1992. 'Lees- en spellingproblemen in de overgangsfase van basisonderwijs naar secundair onderwijs.' *Tijdschrift voor Orthopedagogiek*, 31: 354-370.
- Schiffelers, I., Bosman, A. & Van Hell, J. 2002. 'Uitspreken-wat-er-staat: een effectieve spellingtraining voor woorden met inconsistente foneem-grafeem relaties.' *Tijdschrift voor orthopedagogiek*, 41: 320-331.
- Vanderlocht, M. & Ghesquière, P. 1997. 'Schrijf ik er nu één of twee? Onderzoek naar de beheersing van de regels over medeklinkerverdubbeling en klinkerverenkeling bij eerstejaarsleerlingen in het Vlaams secundair onderwijs.' Gesquière, P. & Ruijsenaars, A. *Leerproblemen in het middelbaar onderwijs*. Leuven: Acco, 83-95.
- Verhoeven, G. 1979. 'Verbeelding en werkelijkheid; Spelfouten in de opstellen uit verschillende typen van onderwijs.' *Tijdschrift voor de taalbeheersing*, 1-2: 146-163.
- Verhoeven, L., Schreuder, R. & Baayen, R. 2006. 'Learnability of graphotactic rules in visual word identification.' *Learning and Instruction* 16 (6): 538-548.
- Zuidema, J. & Neijt, A. 2012. *Verkennd onderzoek naar de wenselijkheid en de haalbaarheid van een verrijking van de Woordenlijst Nederlandse Taal ten behoeve van spellingonderwijs*. Nijmegen: Nederlandse Taalunie. Weblink: (<http://taalunieversum.org/sites/tuv/files/downloads/rapport%20VWS%2015022013.pdf>)