

EEN PASSENDE WOORDKEUS

Het kiezen van woorden voor woordenschatlessen

Carry van de Guchte en Anne Vermeer, Babylon, Letteren, Universiteit van Tilburg

1 Inleiding

Woordenschatonderwijs staat de laatste jaren sterk in de belangstelling in het (basis)onderwijs. Steeds meer leerkrachten zien in dat een behoorlijke woordenschat in het Nederlands voor de leerlingen van essentieel belang is om het onderwijs goed te kunnen volgen, met name bij begrijpend lezen en wereldoriëntatie (Droop, 1999; Inspectie van het Onderwijs, 2001). Ook is duidelijk geworden dat voor veel leerlingen een grote woordenschat geen vanzelfsprekendheid is. De woordenschat van allochtone en taalzwakke autochtone kinderen blijft ver achter bij die van gemiddelde autochtone kinderen (Verhoeven & Vermeer, 1996; 2001). Op veel scholen, met name op scholen met veel allochtone kinderen, is dan ook een plaats voor woordenschatactiviteiten ingeruimd. Ook de nieuwere taal- en leesmethodes besteden expliciet aandacht aan woordenschatonderwijs, in die zin dat er een aparte leerlijn is voor woordenschat of voor strategieën om woorden te leren (Van de Guchte, 2002).

Een van de problemen waar leerkrachten en methodemakers tegenaan lopen, is het kiezen van geschikte woorden voor een woordenschatles. In dit artikel gaan we in op een aantal overwegingen rondom die keuze, aan de hand van de manier waarop woorden zijn gekozen voor de kleuter-cdrom *Woorden vangen* en de taalmethode *Zin in taal*. Allereerst bespreken we een aantal criteria waardoor een woord geschikt wordt voor een woordenschatles. Vervolgens laten we aan de hand van een voorbeeld zien dat er bij het kiezen van woorden voor bovengenoemde leermiddelen steeds gewerkt is vanuit twee kanten: vanuit een thema, en vanuit een woordenlijst. Tenslotte komt aan de orde dat de selectie van woorden ook afhankelijk is van het medium of de interactie waarmee de lesstof wordt overgedragen. Daarbij gaan we ook kort in op de verschillende rollen van leerkrachten en methodemakers.

2 Toeval en bewuste keuze

Is het nodig om veel aandacht te besteden aan het kiezen van woorden voor een woordenschatles? Kunnen we niet gewoon aan het toeval overlaten welke woorden behandeld worden? Er komen toch vanzelf allerlei woorden voor, in (voor)leesboeken bijvoorbeeld of in wat een leerling vertelt? Ja, dat is wel zo, en soms zijn dat geschikte woorden om aan te grijpen voor woordenschatonderwijs, maar soms ook niet. Neem bijvoorbeeld het woord *zorgeloos* in het begin van het boek *De eend en de vos* van Max Velthuis: *“Een eend ging eens een eindje wandelen. Het was lekker weer en ze voelde zich blij en*

zorgeloos”. Een leerkracht die dit boek aan kleuters voorleest zou eventueel kunnen uitleggen wat het woord *zorgeloos* betekent om duidelijk te maken hoe de eend zich voelt. Maar zonder dat boek zou een kleuterleerkracht dat woord niet gauw kiezen om het kleuters aan te leren, zeker niet om het hen actief te leren gebruiken, want het is wel een beetje wijsneuzig woord voor een kleuter.

Het uitleggen van de betekenis van een moeilijk woord, in een leestekst bijvoorbeeld, is dus soms alleen nuttig om ervoor te zorgen dat de leerlingen de tekst begrijpen. Ze hoeven dat woord niet per se te onthouden, al zal dat incidenteel wel gebeuren. Maar met een woordenschatles bedoelen we hier een les waarin gericht gewerkt wordt aan de opbouw van de woordenschat van de leerlingen. In zo'n les wordt de betekenis van woorden aan de leerlingen duidelijk gemaakt en er wordt met de woorden geoefend om ervoor te zorgen dat ze die onthouden. Het zou niet efficiënt zijn om kostbare tijd in zo'n les te besteden aan woorden waarvan alle kinderen in de klas de betekenis al kennen, aan te makkelijke woorden voor die kinderen dus. Een les biedt ook meer uitdaging voor de leerlingen als woorden die behandeld worden door hen als moeilijk of onbekend ervaren worden.

Ook is het jammer als er veel tijd besteed wordt aan woorden die de kinderen verder nergens meer tegenkomen of die ze nooit zelf kunnen gebruiken. Woorden zullen weer snel uit hun geheugen verdwijnen als ze niet functioneel zijn voor de leerlingen. Dat kan het geval zijn met woorden die uit een volwassenencontext komen waarmee kinderen van de desbetreffende leeftijd nog niet te maken hebben, zoals het woord *zorgeloos* uit het voorbeeld hierboven, of juist een woord uit een context van jongere kinderen. Het woord *poppenboek* bijvoorbeeld is een uiterst functioneel woord voor kleuters, maar voor kinderen die op twaalfjarige leeftijd beginnen met Nederlands leren is het niet een woord dat ze snel nodig zullen hebben. Als criteria waardoor een woord geschikt wordt voor een woordenschatles kunnen we dus formuleren:

- (1) De betekenis van het woord is onbekend voor tenminste een deel van de leerlingen in de groep;
- (2) Het woord is functioneel voor de leerlingen, dat wil zeggen het komt vaker voor in taal waarmee de leerlingen te maken hebben, dus in leesboeken en schoolboeken voor de desbetreffende leeftijdsgroep en mondelinge taal van leeftijdgenoten en volwassenen wanneer die spreken met kinderen van die leeftijd.

Wil een woord geschikt zijn voor een woordenschatles dan moet het dus een passend woord zijn voor de leerlingen. Maar het moet ook een passend woord zijn in de les. Het woord moet op een zodanige wijze in de les voorkomen dat het door de leerlingen begrepen en onthouden kan worden. Dat betekent dat er beter niet zomaar allerlei losse woorden behandeld kunnen worden. Een woord wordt makkelijker begrepen als het wordt aangeboden in een context van een thema of verhaal en het wordt beter onthouden als het kan worden vastgelegd in een netwerk van woorden in het geheugen van de leerlingen (Verhallen & Verhallen, 1994). Nog een criterium is dus:

- (3) Het woord past in de context van een thema, een verhaal of een activiteit.

Een zorgvuldige keuze van woorden is ook nodig om ervoor te zorgen dat er voldoende woorden aangeboden en behandeld worden, niet te veel en niet te weinig. Als men het aan het toeval overlaat welke onbekende woorden er opduiken, heeft men

geen greep op het aantal. Ook heeft men dan geen greep op de variatie of diversiteit van de woorden. In een willekeurige tekst kunnen te veel of in het geheel geen moeilijke woorden zitten. En in wat leerlingen of leerkrachten vertellen komen doorgaans niet veel woorden voor die onbekend zijn. Het is voor leerlingen met een geringe woordenschat wel uiterst belangrijk dat ze een rijk aanbod krijgen van voor hen nieuwe woorden, zodat ze hun woordenschat kunnen uitbreiden. Pas dan kunnen zij wat makkelijker ook andere nog onbekende woorden leren (Mondria, 1996) en de leesteksten uit de wereldoriëntatievakken begrijpen. Een vierde criterium bij woordselectie is dus:

- (4) Er wordt een passend aantal woorden aangeboden.

Maar wat is ‘een passend aantal’? Wat betreft tekstbegrip houdt ‘een passend aantal’ woorden in dat er in een tekst niet teveel woorden mogen zitten die onbekend zijn. Het hangt ook van de soort tekst af, en van de kennis over het onderwerp van de tekst, maar voor een redelijk tekstbegrip mogen er niet meer dan 15% van de woorden in een tekst onbekend zijn voor de lezer (zie Vermeer, 1998 en de daar genoemde literatuur). Als het er meer zijn, begrijpt een lezer de tekst niet meer goed. Bovendien kan die dan ook de onbekende woorden niet meer raden, want het kunnen afleiden van de betekenis van een onbekend woord is afhankelijk van de kennis van de woorden er omheen.

Op de vraag wat ‘een passend aantal’ woorden is, kunnen we eigenlijk pas antwoord geven, als we weten hoeveel woorden iemand kent en hoeveel hij er zou moeten kennen. De geschatte omvang van de receptieve Nederlandse woordenschat van eentalige Nederlandse en tweetalige Turkse en Marokkaanse kinderen in Tabel 1 maakt duidelijk hoeveel woorden dat criterium ‘een passend aantal’ inhoudt.

Tabel 1 Geschatte omvang van de receptieve Nederlandse woordenschat van Nederlandse en allochtone kinderen in het basisonderwijs (totaal aantal woorden dat ze kennen)

Leeftijd	Eentalige Nederlanders	Tweetalige Turken/Marokkanen
vierjarigen	3200	1000
vijfjarigen	3800	1800
zesjarigen	4400	2500
zevenjarigen	5000	3100
achtjarigen	5600	3600
negenjarigen	7000	4000
tienjarigen	10000	5500
elfjarigen	13000	7800
twaalfjarigen	16000	9800

De schattingen zijn gebaseerd op de woordenschattaken van de TAK voor de onder- en bovenbouw (Verhoeven & Vermeer, 2001; 1993), waarin onder ‘woord’ wordt verstaan: ‘woordenboekingang’ (lemma), dus het vetgedrukte trefwoord in een woordenboek. De genoemde aantallen woorden komen overeen met elders in de literatuur genoemde, bijvoorbeeld voor zesjarigen (Schaerlackens, 1981) en

twaaalfjarigen (Sanders, 1990). Vanzelfsprekend gaat het om grove indicaties, waarbij de variatie binnen de groepen zeer groot is. Aan de hand van Tabel 1 kunnen we wat meer zeggen over wanneer een selectie voldoende groot is voor een bepaalde groep. Turkse en Marokkaanse kinderen bijvoorbeeld kennen begin groep 1 gemiddeld ongeveer duizend woorden. Eind groep 2 kennen autochtone Nederlandse kinderen er meer dan vierduizend. In ruim twee jaar in de kleutergroepen zouden die Turkse en Marokkaanse kinderen dus ruim drieduizend woorden erbij moeten leren om hun achterstand op autochtone Nederlandse kinderen in te halen. ‘Een passend aantal’ woorden betekent voor deze kinderen dus zo’n 120 woorden per maand. Ook voor de andere groepen worden minimaal zoveel te leren doelwoorden genoemd. Uit onderzoek lijkt het leervermogen van kinderen niet gauw een grens te bereiken: hoe meer woorden er aangeboden worden, hoe meer er ook onthouden worden (Appel & Vermeer, 1997). De grote aantallen woorden in Tabel 1 laten zien hoe gigantisch de leerlast is, en hoe belangrijk het dus is om voor de schaarse onderwijstijd die er is, een uitgekiende en zorgvuldige selectie van woorden te maken.

3 Bewuste woordkeuze: een les uit *Zin in taal*

Pagina 30 van taalboek C1 van *Zin in taal* (zie illustratie 1) hoort bij een les uit de leerlijn woordenschat. Het is les 2 van de eenheid die gaat over het thema ‘wilde dieren’. Op die pagina staan de plaatjes die horen bij het verhaal ‘*Slangenvoer*’, dat geschreven is door Peter Vervloed. De tekst van het verhaal staat op een apart kopieerblad. Het verhaal begint zo.

De kinderen van de klas zijn op bezoek in het reptielenhuis. Vanuit glazen bakken staren bagedissen en slangen hen onbeweeglijk aan. In een boek hangt een skelet van een uitgestorven reptiel.

Een lange magere gids leidt de klas langs de bakken en vertelt iets bij elke bak. Steven en Farid staan achteraan.

In het verdere verloop van het verhaal begint Farid tegen Steven een beetje op te scheppen over wilde dieren in Marokko. Hij vertelt onder andere dat zijn vader oog in oog heeft gestaan met een panter. Steven gelooft er niet veel van, maar even later vertelt de gids ook dat er in Marokko nog panthers zijn gesignaleerd, terwijl men dacht dat die uitgestorven waren. Als de gids vraagt wie er die grote slang vast wil houden, roept Steven dat Farid dat wel wil. Farid durft zich bijna niet te bewegen met de slang om zijn nek. Zijn ogen gaan triomfantelijk glanzen als hij de slang door mag geven, maar waar is Steven dan opeens gebleven?

‘Die moest nodig naar de wc,’ zegt Ayoena. ‘Geef mij die slang maar. Als we moeten wachten tot Steven terugkomt, is dat beest allang uitgestorven.’

Op pagina 30 staat ook een rij woorden die in het verhaal voorkomen en die in de les behandeld worden, de doelwoorden van deze les: *het reptiel, de bagedis, onbeweeglijk, het skelet, uitsterven, de gids, het mormel, de sprinkhaan, geruisloos, met huid en haar, zich afspelen, oog in oog staan met, signaleren, zich bevinden, triomfantelijk*. Een aantal ervan zijn te vinden in bovenstaande samenvatting van het verhaal.



Illustratie 1. *Zin in Taal*, Taalboek C1, p. 30, les 2 uit het thema 'Wilde dieren':
Slangenvoer

De opzet van de les is als volgt. In de introductiefase bekijkt de leerkracht de pagina met de leerlingen en laat hij hen reageren op de titel en de illustraties. In de instructiefase leest de leerkracht het verhaal voor, terwijl de leerlingen het volgen aan de hand van de plaatjes. Vervolgens wordt de betekenis van de woorden in het rijtje op de pagina behandeld door terug te grijpen op de context van het verhaal en uitleg te geven over de betekenis van de woorden. Daarna maken de leerlingen de opdrachten op pagina 31, waarin de doelwoorden nog eens terug komen. Bij opdracht 1 bijvoorbeeld staat een tekstje waarin de doelwoorden blauw zijn afgedrukt. De leerlingen moeten in een kiesrij voor elk blauw woord iets zoeken wat hetzelfde betekent. In de reflectiefase worden de antwoorden op de opdrachten doorgenomen. Dit type les komt in *Zin in taal* in alle delen (voor jaargroep 4 t/m 8) geregeld voor. Het verhaal is dan ook het tweede uit een serie van 10, waarin steeds Steven, Farid en Ayoena voorkomen. Lessen van dit type zijn er ook rondom sprookjes en volksverhalen.

Zin in taal is dus een methode met lessen woordenschat, waarin het leren van woordbetekenissen het doel van de les is. Daarnaast omvat de leerlijn woordenschat ook lessen waarin strategieën voor het leren van woorden geleerd worden, leren leren dus (zie Van de Guchte, 2002). We hebben deze les hier weergegeven als voorbeeld van intentioneel woordenschatonderwijs, waarin doelgericht en systematisch gewerkt wordt aan de opbouw van de woordenschat van de leerlingen. Ook voor de leerlingen is het expliciet dat woorden leren het doel van de les is, want de doelwoorden staan in het taalboek (voor de leerlingen) op een rij (zie illustratie 1).

4 *Kiezen van twee kanten: thema's en woordenlijst*

Hoe zijn de woorden voor de woordenschatlessen van *Zin in taal* gekozen? Hoe is bijvoorbeeld het lijstje met doelwoorden uit de hierboven beschreven les '*Slangenvoer*'

tot stand gekomen? In de eerste plaats is altijd uitgegaan van een thema. De thema's waren vooraf voor de hele methode vastgesteld. In de tweede plaats is gebruik gemaakt van de woordenlijst 'Woorden in het basisonderwijs' (Schrooten & Vermeer, 1994). Woorden die in het kader van het thema geschikt leken, werden opgezocht in deze woordenlijst. Gekeken werd of de woorden in een voor de desbetreffende leeftijdsgroep relevant gedeelte van de lijst voorkwamen. Er werd bij het kiezen van woorden dus steeds gewerkt vanaf twee kanten: vanuit het thema en vanuit de woordenlijst. We gaan hieronder op beide in.

Woorden kiezen rond een thema

Het is een goede gewoonte in het taalonderwijs om te werken in de context van thema's. In de meeste methodes vindt men die gewoonte terug. Werken in thema's is van belang voor de beleving van de kinderen. Ze zien daardoor samenhang tussen verhalen en activiteiten en ze ervaren dat een bepaald onderwerp vanuit verschillende invalshoeken wordt uitgediept (zie ook Verhoeven & Aarnoutse, 2000). Voor de woordenschatlessen van *Zin in taal* zijn dan ook steeds verhalen geschreven of bestaande verhalen gezocht binnen de vooraf bepaalde thema's. In de lessen van het type zoals 'Slangenvoer' komen de doelwoorden ook voor in de context van een verhaal. Behalve de belevingscontext van het thema is er dus ook de talige context van een verhaal. Vooral voor de wat abstractere woorden of voor uitdrukkingen is talige context van belang om de betekenis goed uit te laten komen. Ook voor het onthouden van woorden is een talige en thematische context belangrijk. Woorden worden makkelijker in het geheugen opgeslagen als ze kunnen worden aangehaakt bij woorden die de leerlingen al kennen. Als leerlingen een tijdje aan een thema werken, liggen er als het ware al woorden uit dat themaveld in hun hoofd klaar om nieuwe woorden aan te verbinden. Ook zorgt een thema voor veel kans op herhaling van dezelfde woorden in verschillende lessen. Het woord *uitsterven* uit 'Slangenvoer' bijvoorbeeld komt in hetzelfde taalboek ook in een volgende les over het thema 'wilde dieren' voor (les 4), evenals het woord *reptiel* (les 13). En in les 6 in het Taalmaatje, in de differentiatiestof voor kinderen met een geringe woordenschat, worden de woorden *de hagedis*, *het reptiel* en *de sprinkhaan* nog eens herhaald.

Er komen verschillende tekstsoorten in de methode voor. De les die hier beschreven is bevat een realistisch verhaal over kinderen in hun dagelijks leven, in dit geval een uitstapje met school. Het verhaal is gebaseerd op de ervaringen van de schrijver (Peter Vervloed), die onderwijzer is op een basisschool. Er zijn ook lessen met verhalen waarin fantasie een belangrijke rol speelt: sprookjes en volksverhalen uit verschillende culturen en fragmenten uit klassieke verhalen. Dat zijn dus bestaande teksten waaruit woorden gekozen zijn voor de lessen. Bij het thema 'wilde dieren' is er een les aan de hand van 'Gloeskabi verandert de dieren', een verhaal van de indianen in Noord-Amerika. Bij het thema 'dierenverhalen' in hetzelfde taalboek is er een les rondom een fragment uit Paul Biegels bewerking van 'Reinaart de vos'. De lessen met realistische verhalen vallen steeds in het begin van een periode waarin aan een thema gewerkt wordt. De verhalen staan dichtbij de kinderen, in het hier en nu. De lessen met sprookjes en volksverhalen zitten steeds meer aan het eind van de themaperiode. De verhalen gaan over 'ver weg en vroeger'. Ook in de lessen Spreken/Luisteren van *Zin in taal* is deze opbouw te zien. Uit verschillende typen teksten komen ook verschillende typen doelwoorden voort. Uit de realistische verhalen zijn het veelal concrete of 'down-to-earth' woorden als *de hagedis*, *het skelet*, *het mormel*, enzovoorts. In de lessen met

sprookjes en volksverhalen zijn het vaker abstracte of wat ‘verheven’ woorden, bijvoorbeeld in de les over Reinaart ondermeer de volgende: *de pelgrim, sire, majesteit, overstemmen, machteloos, het gezag*.

Moet elk woord dat behandeld wordt dan voorkomen in een verhaal? Nee, dat hoeft ons inziens niet per se. Ook een thema kan genoeg context bieden om de leerlingen het gevoel te geven dat een les ergens over gaat. Naar aanleiding van een verhaal over dieren, kan een leerkracht best met de leerlingen inventariseren hoe lichaamsdelen van dieren heten, zonder dat elk lichaamsdeel in het verhaal voorkomt, of dieren uit het verhaal indelen in categorieën zoals reptielen en insecten en daarbij ook dieren betrekken die niet in het verhaal voorkomen.

Wat misschien opvalt is dat de doelwoorden uit het rijtje van ‘Slangenwoer’ niet allemaal dicht bij het thema ‘wilde dieren’ zitten. Sommige woorden wel, zoals *het reptiel, de hagedis, het skelet, de sprinkhaan*, maar er zijn ook vrij algemene woorden bij, zoals *onbeweeglijk, geruisloos, zich bevinden, triomfantelijk*. Als men alleen vanuit het thema al brainstormend woorden zou bedenken, zouden zulke algemene woorden niet gekozen zijn. Door tegelijk te werken met teksten rondom een thema komen ook die algemene woorden aan bod.

Voor een deel zijn de teksten in de woordenschatlessen van *Zin in taal* bestaande verhalen. Voor andere lessen zijn nieuwe verhalen geschreven. Een probleem dat zich voordoet bij het werken met bestaande teksten is dat er in een verhaal dat op zich prachtig is en goed past bij het thema, vaak niet genoeg relevante woorden zitten voor een woordenschatles. Het is soms erg zoeken naar mooie verhalen met een interessante woordenschat. De verhalen mogen doorgaans bovendien ook niet te lang zijn als ze in een les behandeld moeten worden. Als er nieuwe verhalen geschreven worden voor woordenschatlessen kan men de woordenschat daarvan een beetje sturen. De schrijvers die verhalen schreven voor de woordenschatlessen van *Zin in taal* kregen vooraf een keuzelijst met geschikte woorden, met het verzoek om 12 à 15 woorden daarvan te verwerken in een verhaal. De keuzelijst was lang (zo’n 60, 100 woorden of meer), zodat de schrijvers een ruime keus hadden en niet geforceerd bepaalde woorden moesten gebruiken. Voorop stond natuurlijk dat het een goed verhaal moest worden. Heel vaak doken er dan in zo’n verhaal ook nog heel interessante en geschikte woorden op die niet op de keuzelijst stonden. Vaak werd zo’n woord dan doelwoord in een les, zeker als het een mooie plaats had in het verhaal en de betekenis daarin goed uit de verf kwam. Het toeval is dus altijd wel een rol blijven spelen bij het kiezen van woorden.

Woorden kiezen vanuit een woordenlijst

Voor *Zin in taal* is gebruik gemaakt van de woordenlijst *Woorden in het basisonderwijs* (Schrooten & Vermeer, 1994). In het vreemde-talenonderwijs hebben woordenlijsten al vanaf begin vorige eeuw een belangrijke rol gespeeld bij woordenschatselectie (zie Bogaards, 1986a; 1986b voor een historisch overzicht). Immers, voor het leren van een vreemde taal is in ons voortgezet onderwijs slechts twee tot vier uur per week beschikbaar en lang niet alle woorden van een taal zijn even nuttig. Maar wat zijn nu eigenlijk de nuttigste woorden in een taal? ‘Klassieke’ criteria hierbij zijn ‘leerbaarheid’ (ook wel: ‘moeilijkheidsgraad’, ‘leerlast’), valentie (ook wel: ‘vervangbaarheid’, of je met een woord veel kunt zeggen), en oordelen van ‘native speakers’ zoals leerkrachten (denk aan de *Streeflijst Woordenschat zesjarigen*, Schaerlaekens, Kohnstamm & Lejaegere, 1999). Maar naast dit soort intuïtieve criteria speelde (en speelt) ook het begrip ‘frequentie’ een belangrijke rol. Zo verzamelde De la Court (of liever gezegd: zijn

kweekschoolleerlingen) uit Batavia in 1937 één miljoen woorden op kaartjes (met de hand!), om tot een lijst te komen ten behoeve van het onderwijs Nederlands in het toenmalige Nederlands-Indië. Overigens werden in die lijst de 200 meest frequente woorden na een tijdje niet meer geteld ('ontelbaar'). Met de komst van de computer is het veel gemakkelijker geworden grote verzamelingen (*corpora*) aan te leggen en frequentielijsten op te stellen, en praktisch alle vreemde-taalmethoden verwijzen nu voor de verantwoording van hun woordselectie naar bekende frequentielijsten.

Dat het 'nut' van een woord samenhangt met de frequentie ervan, wordt ondersteund door onderzoeksresultaten die laten zien dat er een relatie bestaat tussen de frequentie van een woord in het dagelijks taalaanbod en de verwerving ervan, dat wil zeggen, hoe frequenter een woord voorkomt in het aanbod, hoe eerder het geleerd wordt (zie o.a. Kibby, 1977; Anderson & Freebody, 1981, 101; Nation, 1990, 77; Brown, 1993, 277; Hazenberg, 1994, 129; Vermeer, 1997, 181). En, in relatie tot lezen: hoe frequenter een woord voorkomt in teksten, hoe sneller het ook herkend wordt (Rudell, 1993). Bovendien blijkt uit onderzoek dat (afhankelijk van de wijze van tellen, van de doelgroep en van het soort tekst, zie Hazenberg, 1994) de 4.000 tot 8.000 meest frequente woorden (lemma's) al bijna 85% van een willekeurige tekst dekken, welk percentage minimaal nodig is om tot redelijk tekstbegrip te komen.

Toch moet met het criterium frequentie opgepast worden. Ten eerste is een frequentielijst die gebaseerd is op krantenteksten niet geschikt voor de selectie van mondeling vocabulaire (de *domeinen* zijn verschillend), net zo min als een lijst die gebaseerd is op taalgebruik van volwassenen dat is voor kinderen (de *doelgroepen* zijn verschillend). Ten tweede zijn kleine verzamelingen woorden al gauw onbetrouwbaar, omdat het toeval daar veel te groot is. Volgens Alekseev (1984, 62) ligt de betrouwbaarheidsgrens (ongeacht de grootte van het corpus) bij een frequentie van 35, dat wil zeggen dat slechts die woorden voldoende betrouwbaar geacht worden om op te nemen in een lijst, als die minimaal 35 keer voorkomen in het corpus (in de woordenlijst van Schrooten & Vermeer hebben zo'n 3000 woorden een frequentie van 35 of meer). Ten derde zijn de 200 tot 400 meest frequente woorden niet automatisch de nuttigste. Zeer frequente woorden (*maken, doen, gaan, de, in*) hebben een grote inzetbaarheid, maar een geringe informatieve waarde. Ten slotte zijn sommige woorden niet heel erg frequent, maar komen wel in alle soorten teksten voor (ze hebben een grote *spreiding*), en zijn daarom nuttig. Om die reden wordt vaak niet alleen naar de frequentie, maar ook naar de spreiding van woorden gekeken.

Voor het kiezen van woorden voor woordenschatlessen in het basisonderwijs moet een lijst dus aan een aantal voorwaarden voldoen. De lijst moet gebaseerd zijn op de betreffende doelgroep en taaldomeinen, het corpus moet voldoende groot zijn, en de woorden moeten verzameld zijn uit veel verschillende onderwerpen en situaties. Bovendien moet er niet louter naar de frequentie van woorden gekeken worden, maar ook naar de spreiding over die domeinen. Met dit doel en deze criteria voor ogen is de woordenlijst van Schrooten & Vermeer (1994) tot stand gekomen. Uit het schriftelijke en mondelinge taalaanbod in alle groepen in het basisonderwijs, zijn (verdeeld over 92 verschillende thema's) uit prentenboeken, leesboeken, taal-, reken- en wereldoriëntatiemethoden, en uit de instructie door leerkrachten in de klas, bijna twee miljoen woorden (tokens, woordvormen) verzameld, wat in totaal ruim 26.000 verschillende lemma's opleverde. Van elk lemma is het *geometrisch gemiddelde* berekend. Het geometrisch gemiddelde is een samengestelde maat waarin, naast de pure frequentie van woorden in de lijst, ook rekening gehouden wordt met de spreiding van een woord over de

verschillende domeinen. In principe komt het erop neer dat een woord dat bijvoorbeeld in 40 domeinen steeds maar één keer voorkomt een veel hoger geometrisch gemiddelde heeft dan een woord dat 40 keer voorkomt in maar één domein, terwijl de frequentie voor beide hetzelfde is (zie voor een uitvoerige beschrijving van deze maat Van Hout & Vermeer, 1992).

Ten behoeve van het kiezen van woorden voor lesmateriaal is de totale lijst onderverdeeld in negen deellijsten. De eerste lijst (voclist1) omvat de duizend lemma's met het hoogste geometrische gemiddelde, de tweede lijst de daarop volgende duizend, enzovoorts, met dien verstande dat de zesde en zevende lijst ruim 1.500 lemma's omvatten, de achtste lijst ruim 4.500, en de negende lijst (voclist9) bijna 14.000 lemma's. Overigens zijn in voclist 8 en 9 de frequenties zo laag (veelal minder dan 6) dat het niet meer zinvol is een volgorde in moeilijkheidsgraad aan te geven. Het aantal tokens, lemma's en het geometrisch gemiddelde van de negen lijsten is weergegeven in Tabel 2, en in de laatste kolom staat voor welke groep in het basisonderwijs uit welke lijsten geput kan worden.

Tabel 2 Aantal tokens, lemma's (afgerond), en geometrisch gemiddelde, per deelwoordenlijst

woorden lijst	aantal tokens per voc.list	aantal lemmas per voc.list	aantal lemmas cumulatief	geometr gemidd	keuze voor jaargroep									
					1	2	3	4	5	6	7	8		
voclist1	1669651	1000	1000	> 115	1									
voclist2	117031	1000	2000	115 - 68	1	2								
voclist3	51222	1000	3000	67 - 58	1	2	3							
voclist4	29718	1000	4000	57 - 53	1	2	3	4						
voclist5	19196	1000	5000	52 - 51			3	4	5					
voclist6	19002	1500	6500	50 - 49				4	5	6				
voclist7	13348	1600	8100	49 - 48					5	6	7			
voclist8	20340	4600	12700	47						6	7	8		
voclist9	18470	13890	26590	46							7	8		
totaal	1957978	26590												

Uit het overzicht van de omvang van de woordenschat van kinderen (Tabel 1) bleek dat kinderen van 4 tot en met 6 jaar tussen de 1000 (vierjarige allochtonen) en 4400 woorden (zesjarige autochtonen) kennen. Woorden uit de eerste vier voclijsten zijn dus voor een deel van de kinderen in deze leeftijdsgroep onbekend, maar wel functioneel om te leren. Op dezelfde wijze kan met behulp van Tabel 2 een passende woordkeus gemaakt worden voor bijvoorbeeld groep 3. De kinderen uit deze groep kennen (zie Tabel 1) tussen de 2500 (zesjarige allochtonen) en 5000 (zevenjarige autochtonen) woorden. Voor groep 3 kunnen dus de woorden uit voclist 3, 4 en 5 gekozen worden. Als er veel taalzwakke kinderen tot de doelgroep behoren, zou ook nog uit voclist2 geput kunnen worden.

Voor de vier niveaus die in het woordenschatprogramma voor kleuters *Woorden vangen* worden onderscheiden, zijn woorden gekozen uit de eerste vier voclijsten uit de woordenlijst. *Woorden vangen* is bedoeld voor kleuters in groep 1 en 2, vooral voor allochtone en taalzwakke autochtone kinderen. Het omvat twee

woordenschatpakketten, elk bestaande uit een knieboek met een voorleesverhaal, een klein prentenboekje met de illustraties van het verhaal, een daarbij behorende cd met het verhaal en een cd-rom met het verhaal en woordenschatspelletjes. Thema's zijn 'school' en 'spelen'. Het programma selecteert in de spelletjes automatisch de woorden uit het niveau waar het kind aan toe is. Als bijvoorbeeld blijkt dat het kind de woorden op niveau 2 (voclist2) kent, dan schakelt het programma automatisch over naar woorden uit niveau 3 (voclist3). Blijkt het kind sommige woorden van niveau 2 wel en andere niet te kennen, dan krijgt het de woorden uit niveau 2. Als het kind te weinig woorden kent uit niveau 2, dan schakelt het programma terug naar niveau 1 (voclist1).

Ook voor *Zin in taal* is voor de selectie gekeken naar hoeveel woorden de zwakste en de sterkste leerdersgroepen gemiddeld kennen, en is vervolgens nagegaan uit welke voclisten geput kon worden. Voor groep 4 (deel A) bijvoorbeeld is dat tussen de 3100 en 5600 woorden (zie Tabel 1), dus uit voclist 3 tot en met 6. Woorden uit deze voclisten kunnen als onbekend verondersteld worden voor tenminste een deel van de leerlingen in die leeftijdsgroep en tevens zijn het voor hen wel nuttige woorden die ze vaker tegen zullen komen. Woorden die uit de thema's naar voren kwamen, in bestaande verhalen voorkwamen, of voor nieuwe verhalen op de keuzelijst voor de auteur gezet werden, zijn dus in de woordenlijst opgezocht als check bij het bepalen of het geschikte doelwoorden waren. Overigens is de lijst met soepelheid gehanteerd. Regelmatig zijn er woorden die niet in de lijst of in de 'juiste' voclist voorkwamen toch tot doelwoorden gebombardeerd, met name woorden die essentieel waren in een verhaal of die heel mooi in de context pasten en waarvan men op grond van intuïtie kon zeggen dat het nuttige woorden voor de leerlingen waren. Bij het gebruiken van de woordenlijst moet men steeds bedenken dat de lijst een indicatie geeft voor de geschiktheid van woorden voor een woordenschatles, maar niet als wet van Meden en Perzen voorschrijft welke woorden behandeld moeten worden. Ook intuïtie moet bij de keuze een rol spelen.

5 Een keuze die past: woorden voor boeken, computers en interactie

Behalve met het thema van de les en met de taalvaardigheid van de leerlingen moet men bij de keuze van woorden ook nog rekening houden met het medium door middel waarvan de lesstof wordt overgedragen: boeken, computer, of interactie in de klas. Elk van die manieren van overdracht heeft zo zijn voor- en nadelen. Bij het semantiseren van woorden ligt de kracht van de computer bijvoorbeeld in geluid en (bewegend) beeld.

Voor de cd-roms van *Woorden vangen* is bij het kiezen van woorden in de eerste plaats het thema bepalend geweest. Voor elk van de vier niveaus werden woorden rondom het thema verzameld. Maar bij het kiezen van woorden voor de cd-roms was ook een van de criteria dat de betekenis van het woord met de computer duidelijk gemaakt moest kunnen worden aan kleuters, dus aan kinderen die nog niet kunnen lezen. Er zijn dan ook vooral woorden gekozen waarvan de computer een (eenduidig) plaatje kan laten zien, eventueel ook een animatie als het om een bewegingswoord gaat. Ook moest er van een woord een omschrijving gegeven kunnen worden in eenvoudige, voor kleuters te begrijpen bewoordingen. Verder wordt de betekenis van de woorden in het programma verduidelijkt door een contextzin, een geluid indien dat van toepassing is (bijv. bij *knippen*) en een vertaling in drie talen (Engels, Marokkaans-Arabisch en

Turks). Een deel van de woorden komt voor in het verhaal van de cd-rom. Woorden die niet goed duidelijk te maken waren door middel van een plaatje, een animatie of een eenvoudige omschrijving zijn dus niet gekozen voor de spelletjes. Een woord als *klas* bijvoorbeeld is afgefallen, hoewel het een belangrijk woord is in het kader van het thema van de eerste cd-rom 'school'. De betekenis van dit woord is moeilijk door een plaatje weer te geven. Het is immers altijd maar een deel van de klas wat je op een plaatje ziet. Het woord *klas* kan veel beter aangeboden worden in de werkelijkheid, door om je heen te wijzen als je met een kind midden in de klas staat, dus in de interactie tussen leerkracht en leerling. Voor aanbieden op de computer is dit woord minder geschikt, althans voor kleuters die het vooral moeten hebben van beeld. Voor oudere kinderen die al kunnen lezen of die over meer taal beschikken zijn er natuurlijk ook op de computer meer mogelijkheden.

Zo heeft de computer dus zijn beperkingen wat betreft het semantiseren van woorden. Boeken bieden wat dat betreft andere mogelijkheden. De talige context van sprookjes en andere fantasieverhalen lenen zich juist voor de keuze van meer 'literaire' woorden, zoals we in paragraaf 4 al aangegeven hebben. De voorleesverhalen en zelf te lezen verhalen zijn de 'ankers' die het mogelijk maken een rijk en gevarieerd aanbod van woorden te garanderen (Verhoeven & Aarnoutse, 2000). Door interactie rondom die boeken is er nog meer mogelijk. De leerkracht kan verduidelijken, dingen laten zien, voordoen, de kinderen laten uitbeelden, het verhaal met hen reconstrueren en verbinden aan eigen ervaringen: daar kan geen computer tegenop.

Welke woorden men kiest en op grond van welke criteria hangt dus mede af van de onderwijssituatie en de manier van overdracht. Een kleuterleerkracht die al brainstormend woorden kiest voor de invulling van een thema kiest waarschijnlijk woorden die in de omgeving van de kinderen van belang zijn. Bij het thema 'school' bijvoorbeeld zal de keuze voor een deel afhankelijk zijn van wat er in en om die school te zien en te beleven is. Dat is prachtig, want dat zullen woorden zijn die die kinderen op die school nodig hebben. Voor een groot deel zullen het waarschijnlijk woorden zijn die op iedere school van toepassing zijn. Een leerkracht kan ook woorden kiezen door (voor)leesteksten voor te bereiden en daaruit woorden te kiezen die essentieel zijn voor het begrip van die tekst (Kienstra, 1999). Als de betekenis van die woorden behandeld wordt, hebben de leerlingen daar meteen profijt van voor het begrip van de tekst. Ook kan een leerkracht ingaan op woorden waarvan de leerlingen vragen wat ze betekenen.

Een leerkracht kan bij het kiezen van woorden voor een les in eigen klas vooral afgaan op intuïtie en ervaring. Ook is het dan mogelijk om bij een paar leerlingen even te checken of de gekozen woorden inderdaad onbekend zijn, al zal dat nooit een compleet beeld van alle kinderen in de klas opleveren. De keuze van woorden die een leerkracht maakt zal daarmee meestal wel wat onsystematisch en van het toeval afhankelijk zijn. Leerkrachten die veel leerlingen met een geringe woordenschat hebben, laten het daar hopelijk niet bij en gebruiken daarnaast ook een taalmethode met een systematisch opgebouwde woordenschatleerlijn.

Makers van taalmethodes, boeken en cd-roms staan voor een andere situatie. Zij moeten de keuze van woorden afstemmen op alle - of in ieder geval vele - scholen in het land. Er kan tijdens de ontwikkeling niet even snel bij een grote groep kinderen nagegaan worden of de gekozen woorden geschikt zijn. Bovendien moet er in een methode een verantwoorde opbouw zitten. Een methodemaker zal dus meer tijd in de selectie van woorden moeten steken.

De aanpak van het kiezen van woorden die in dit artikel beschreven is, is ontstaan

in de praktijk van leermiddelenontwikkeling voor NT2-onderwijs. De eerste auteur van dit artikel was betrokken bij het samenstellen van verschillende NT2-methodes. Voor het onderdeel woordenschat van *Van horen en zeggen* (Werkgroep Nederlands voor anderstaligen, 1983) werden door de leerkrachten in de auteursgroep al brainstormend woorden rondom thema's verzameld. Bij de ontwikkeling van *Mondeling Nederlands bij Veilig leren lezen* (Van de Guchte, 1988) ontstond bij het kiezen van woorden behoefte aan meer fundering dan alleen intuïtie en ervaring. Voor dit pakket werden woorden geselecteerd uit de leesboekjes van de toen bestaande versie van *Veilig leren lezen*. Van die woorden werd bij een beperkte groep kinderen op een paar Tilburgse scholen nagegaan of de betekenis bekend was, aan de hand van een proefversie van het materiaal.

Ten behoeve van het kiezen van woorden bij het ontwikkelen van NT2-leermiddelen is in opdracht van het Ministerie van O&W (via de Projectgroep NT2) de woordenlijst *Woorden in het basisonderwijs* samengesteld, waarbij de tweede auteur van dit artikel betrokken was. Beide lijnen zijn terug te vinden in de hier beschreven aanpak van woordselectie. Die aanpak is te karakteriseren als kiezen van twee kanten: vanuit thema's en vanuit een woordenlijst. De thema's en teksten zorgen voor een zinvolle betekeniscontext en herkenbaarheid voor de leerlingen; de woordenlijst bevordert de systematische opbouw van de woordenschat en efficiënte besteding van de onderwijstijd. Voor het kiezen van woorden voor woordenschatlessen vormen deze twee sporen een goede combinatie.

6. Bibliografie

- Alekseev, P.M. (1984). *Statistische lexikographie*. Bochum: Brockmeyer.
- Anderson, R.C., & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and Teaching: Research Reviews* (pp.77-117). Newark, DE: International Reading Association.
- Appel, R., & Vermeer, A. (1997). *Woordenschat en taalonderwijs aan allochtone leerlingen*. Tilburg: TUP.
- Bogaards, P. (1986a). Op zoek naar nuttige woorden. Een geschiedenis van de vocabulaire-selectie voor het vreemde-talenonderwijs. *Levende Talen* 411: 328-336.
- Bogaards, P. (1986b). Hoe nuttig zijn frequente woorden? *Levende Talen* 416: 626-630.
- Brown, Ch. (1993). Factors affecting the acquisition of vocabulary: frequency and saliency of words. In Th. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp.263-286). Norwood, NJ: Ablex.
- Droop, M. (1999). *Effects of linguistic and cultural diversity on the development of reading comprehension*. Nijmegen: proefschrift Katholieke Universiteit Nijmegen.
- Guchte, C. van de (1988). *Mondeling Nederlands bij Veilig leren lezen. Een aanvullend programma voor aanvankelijk leesonderwijs aan anderstalige en dialectspreekende kinderen*. Tilburg: Zwijsen.
- Guchte, C. van de (2002). Trucs om taal te leren; taalleerstrategieën in Zin in taal. *Moer* 2, 61-69.
- Hazenberg, S. (1994). *Een keur van woorden*. Amsterdam: proefschrift Vrije Universiteit.
- Hout, R. van & Vermeer, A. (1992). Frequenties van woorden en het geometrisch gemiddelde. *Gramma/Tijdschrift voor Taal- en Tekstwetenschap* 1, 2, 125-132.
- Inspectie van het Onderwijs (2001). *Onderwijsverslag over het jaar 2000*. Utrecht: Inspectie

- van het Onderwijs.
- Kibby, M.W. (1977). Note on relationship of word difficulty and word frequency. *Psychological Reports*, 41, 12-14.
- Kienstra, M. (1999). *Woordenschatactiviteiten rond prentenboeken en voorleesverhalen*. Nieuwsbrief 4. Nijmegen: EN/KUN.
- Mondria, J.-A. (1996). *Vocabulairevererving in het vreemde-talenonderwijs. De effecten van context en raden op de retentie*. Groningen: proefschrift Rijks Universiteit Groningen.
- Nation, I.S.P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle Publishers.
- Rudell, A. (1993). Frequency of word usage and perceived difficulty: Ratings of Kucera and Francis words. *Behavior Research Methods, Instruments and Computers*, 25, 455-463.
- Sanders, M. (1990). Met andere woorden. Een onderzoek naar de omvang van de receptieve woordenschat van anderstalige brugklasleerlingen. *Levende Talen*, 455, 424-427.
- Schaerlaekens, A.-M. (1981). *De taalontwikkeling van het kind. Een oriëntatie in het Nederlandstalig onderzoek*. Groningen: Wolters-Noordhoff.
- Schaerlaekens, A.-M., Kohnstamm, D. & Lejaegere, M. (1999). *Streeflijst Woordenschat voor Zesjarigen*. Derde herziene versie gebaseerd op nieuw onderzoek in Nederland en België. Lisse: Swets & Zeitlinger.
- Schrooten, W., & Vermeer, A. (1994). *Woorden in het basisonderwijs*. 15.000 woorden aangeboden aan leerlingen. Tilburg: TUP.
(<http://cwis.kub.nl/~fdl/general/people/avermeer/wrdlst/wrdlstin.htm>)
- Verhallen, M., & Verhallen, S. (1994). *Woorden leren woorden onderwijzen*. Hoevelaken: CPS.
- Verhoeven, L., & Aarnoutse, C. (2000). Interactief taalonderwijs: Balans tussen constructief en instructief leren. *Spiegel*, 17/18, 9-29.
- Verhoeven, L., & Vermeer, A. (1996). *Taalvaardigheid in de bovenbouw*. Tilburg: TUP.
- Verhoeven, L., & Vermeer, A. (2001). *Taaltoets Alle Kinderen*, Arnhem: Cito.
- Vermeer, A. (1997). Breedte en diepte van woordenschat in relatie tot toenemende taalvererving en frequentie van aanbod. *Gramma/Tijdschrift voor Taal- en Tekstwetenschap* 6, 3: 169-187.
- Vermeer, A. (1998). Tekstdekking in groep 5. Woordkennis als oorzaak van het Mattheüs-effect in het basisonderwijs. *Toegepaste TaalWetenschap in Artikelen*, 60, 9-18.
- Werkgroep Nederlands voor anderstaligen (1983). *Van horen en zeggen. Een methode Nederlands voor anderstalige kinderen in het basisonderwijs*. Groningen: Wolters-Noordhoff.
- Leermiddelen
- Woorden vangen. Een woordenschatprogramma voor kleuters*. Boot, I., Guchte, C. van de, Linden, W. van der, & Vermeer, A. (2003). Tilburg: Zwijzen.
- Zin in taal*. Kouwenberg, B., Guchte, C. van de, & Gein, J. van de, Tilburg: Zwijzen 1997-2001.