

# **MET DE DEUR OF MET DE PET GOOIEN**

Grammaticaliteitsoordelen  
over syntactische en interpretatieve operaties  
op idiomen en collocaties

Studentnummer	Fieneke Jochemsen 3925625
Studie Deelgebied	Universiteit Utrecht Nederlandse taal en cultuur (voltijd) Taalkunde
Begeleider Tweede corrector	Bachelorscriptie April 2015 prof. dr. Norbert Corver dr. Eefje Boef

## Abstract

Het brede veld van idiomaticiteit is een veelvuldig onderzocht terrein binnen de linguïstiek. Duidelijk is dat idiomen zoals het Engelse *to kick the bucket* ('sterven') op syntactisch gebied veel minder flexibel zijn dan andere expressies. De verklaring hiervoor is niet volledig duidelijk, maar men is het erover eens dat de mate waarin de delen van een idioom bijdragen aan de gehele betekenis ervan, onder andere bepalend is voor de flexibiliteit van de expressie.

Bevindingen uit eerder onderzoek suggereren dat op collocaties meer grammaticale operaties, zoals verplaatsing en morfologische manipulatie, toegepast kunnen worden dan op idiomen. Het huidig onderzoek toetst dit verschil door een participantengroep grammaticaliteitsoordelen te laten vellen over zinnen met idiomen en collocaties waar een syntactische operatie, topicalisatie, of een interpretatieve operatie, pronominalisatie, op is uitgevoerd.

De gemanipuleerde collocaties werden door de participanten als significant aanvaardbaarder beoordeeld dan de gemanipuleerde idiomen. Idiomen zijn dus inderdaad syntactisch minder flexibel dan collocaties, al hangt dit vooral af van de semantische analyseerbaarheid van de expressie. Het huidige onderzoek speculeert over de oorzaak van dit verschil tussen collocaties en idiomen, dat te maken zou kunnen hebben met de manier waarop deze soorten expressies in het lexicon zijn opgeslagen: idiomen zouden als geheel lexicale insertie ondergaan en collocaties niet.

# Inhoudsopgave

<b>1. Introductie</b>	<b>4</b>
<b>2. Theoretisch kader</b>	<b>6</b>
2.1 Noties en definities	6
2.2 De plaats van idiomen in het lexicon	7
2.2.1 De Chomskyaanse traditie: X <sup>0</sup> -insectie	7
2.2.2 Theory of Fixed Expressions	9
2.3 Syntactische beperkingen en betekenis	10
2.3.1 Syntactische beperkingen	10
2.3.2 Betekenisloze constituenten	10
2.3.3 Of toch betekenisvolle constituenten?	12
2.4 Idiomen en collocaties	13
2.4.1 Co-occurrence	13
2.4.2 Encodeerbaarheid en decodeerbaarheid	14
2.5 Verwerking en beoordeling van idiomatische uitdrukkingen	15
<b>3. Vraagstelling en hypothese</b>	<b>17</b>
3.1 Vraagstelling	17
3.2 Hypothese	18
<b>4. Methode</b>	<b>20</b>
4.1 Participanten	20
4.2 Materiaal	20
4.2.1 Top-conditie	20
4.2.2 Ana-conditie	22
4.2.3 Fill-conditie	24
4.2.4 Eisen aan het materiaal	24
4.3 Procedure	25
<b>5. Resultaten</b>	<b>26</b>
5.1 Exclusies	26
5.2 Conditiegemiddeldes	26
5.3 Het vergelijken van condities	28
5.4 Conclusie	29
<b>6. Discussie</b>	<b>30</b>
6.1 Algemeen	30
6.2 Condities	30
6.2.1 Topicalisatie	30
6.2.2 Pronominalisatie	32
6.2.3 Fillers	33
6.3 Het onderscheid tussen idiomen en collocaties	34
Literatuur	36
Bijlage I. Materiaal	40
Bijlage II. Instructietekst questionnaire	43

# 1. Introductie

Het idiomatische domein van een natuurlijke taal is veel groter dan je in eerste instantie zou verwachten. Spreekwoorden, gezegden, formules, clichés, collocaties en uitdrukkingen maken er deel van uit, om er maar een paar op te noemen. Het lijkt erop dat elk van deze soort expressies zich op een glijdende schaal van idiomaticiteit bevindt waarin de mate van compositionaliteit een grote rol speelt.

Het principe van compositionaliteit zegt dat de betekenis van een grammaticale complexe expressie, zoals een zin, een compositionele functie is van de betekenis van zijn grammaticale constituenten. Dit betekent dat de betekenis van zo'n complexe expressie volledig bepaald is en voorspeld kan worden door de betekenis van de constituenten en dat elke constituent een betekenis heeft die bijdraagt aan de expressie (Cruse 2000: 67).

Het is in ieder geval duidelijk dat expressies van de puurste vorm van idiomaticiteit, te weten idiomen, niet aan dit principe voldoen. In (1a) dragen de afzonderlijke constituenten van het idioom *de plaat poetsen* niet bij aan de betekenis van de gehele VP (verb phrase) en is die betekenis geenszins voorspelbaar door de betekenis van de constituenten op zich. Door dit gebrek aan compositionaliteit, ook wel 'ondoorzichtigheid', kunnen niet alle syntactische operaties worden uitgevoerd op idiomen, terwijl die wel bij andere expressies mogelijk zijn (Everaert e.a. 1995:7). Vergelijk de gewone, letterlijke expressie in (2) en het idioom in (1).

- (1)<sup>1</sup>
- a. Jan poetst de plaat
  - b. \*De plaat wordt gepoetst (door Jan)
  - c. \*De plaat poetst Jan

- (2)
- a. Jan aait de kat
  - b. De kat wordt geaaid door Jan
  - c. De kat aait Jan

Deze hier getoonde ondoorzichtigheid, het feit dat er niet zonder restricties syntactische operaties op het idiomatische zinsdeel uitgevoerd kunnen worden, is echter niet bij elk idioom in deze mate aanwezig. Voor sommige sprekers van het Nederlands zijn de gemanipuleerde zinnen in (3) bijvoorbeeld wél mogelijk, terwijl die voor anderen uiterst ongrammaticaal zijn, meent Schenk (1995:256-257).

- (3)
- a. Jan draaide mij een loer. Schenk (1995:257)
  - b. \*/√Jan draaide mij een verschrikkelijke loer.
  - c. \*/√De loer die hij mij draaide...
  - d. \*/√Wat een loer draaide hij mij!

---

<sup>1</sup> De zinnen (1b-c) kunnen geen idiomatische betekenis hebben en zijn daarom gemarkeerd met het gangbare symbool voor ongrammaticale zinnen, een asterisk. De √ duidt een grammaticale zin aan.

Schenk (1995) meent dat dit komt door (individuele) taalvariatie: bepaalde 'idiomen' zouden in de interne grammatica's van sommige sprekers van een taal geen idioom maar een collocatie zijn, dat wil zeggen een (min of meer) vaste woordgroep die semantisch wél doorzichtig is (zoals *sterke koffie* maar geen *krachtige koffie* en *de tafel dekken* maar niet *de tafel opmaken*). Collocaties zouden veel meer syntactische vrijheid hebben dan idiomen, wat de (per individu verschillende) grammaticaliteit in (3) lijkt te verklaren. Schenk (1995) voorspelt dan ook dat expressies zoals *X een loer draaien* wel consistent goed of fout zullen worden beoordeeld, d.w.z. voor iedere spreker die (3b) afkeurt, zal (3c-d) ook fout zijn.

Omdat collocaties dus in hogere mate compositioneel zijn, zouden er dus meer syntactische operaties op uitgevoerd moeten kunnen worden. Of een expressie een idioom of collocatie is, zou dus moeten kunnen voorspellen of grammaticale operaties op die expressie mogelijk zijn.

In deze scriptie zal ik nader onderzoeken wat de verschillen zijn tussen collocaties en idiomen wat betreft de mogelijkheden tot grammaticale operaties. In hoofdstuk 2 zet ik het theoretisch kader waarin ik te werk ga uiteen. In hoofdstuk 3 definieer ik mijn vraagstelling en hypotheses die ik aan de hand van de methode in hoofdstuk 4 probeer te beantwoorden. In hoofdstuk 5 bespreek ik de resultaten van dit onderzoek. In hoofdstuk 6, ten slotte, draag ik concluderend een theoretische verklaring voor de gevonden resultaten aan en doe ik suggesties voor vervolgonderzoek.

## 2. Theoretisch kader

### 2.1 Noties en definities

In het voorgaande gebruikte ik de term idioom alsof de definitie algemeen bekend is, maar niets is minder waar. Een bekende en gangbare definitie van idioom is 'geconventionaliseerde complexe expressie' (Everaert e.a. 1995). Een prototypisch idioom zoals *een boterbriefje halen* past daarbinnen. De expressie is complex omdat deze uit meer dan één woord bestaat en geconventionaliseerd omdat de betekenis volgens conventie tot stand is gekomen: de betekenis is niet compositioneel. Niet-compositioneel betekent dat de betekenis van het geheel *een boterbriefje halen* ('trouwen') geen grammaticale optelsom is van de betekenis van de constituenten *een boterbriefje* en *halen*; de betekenis is dus niet afleidbaar uit de losse delen.

Omdat de betekenis van idiomen niet-compositioneel is<sup>2</sup>, kunnen ze niet op dezelfde manier geleerd en verwerkt worden als een grammaticaal compositionele expressie, zoals *een appel geven*, waarvan de betekenis een optelsom is van *een appel* en *geven*, maar dat ze expliciet geleerd moeten worden (Everaert e.a. 1995:3).

De term idioom wordt meestal, zo ook in dit onderzoek, gebruikt als subklasse van 'idiomatische uitdrukkingen' voor expressies die semantisch ondoorzichtig of niet-compositioneel zijn. In deze definitie vallen veel idiomatische uitdrukkingen buiten de boot die zich afspelen op zinsniveau (*Hoe laat is het?*), op woordgroepniveau (*X laat staan Y*) en op woordniveau (*zwarte bes*), die ook allemaal niet alleen maar een optelsom van de constituenten zijn of in ieder geval niet zo verwerkt worden (Geeraerts 1995:58). De idiomen die meestal worden bedoeld met de term 'idioom' spelen zich af op VP-niveau; denk aan uitdrukkingen zoals *de bloemetjes buiten zetten*, *de draak steken met X*, *een uiltje knappen* en *de knoop doorhakken*.

In het algemeen is men het erover eens dat idiomen expressies zijn die bepaalde lexicale beperkingen laten zien (zo is bij *de geest geven* een idiomatische lezing mogelijk, maar *de geest aanbieden*, *leveren*, *schenken* of *verschaffen* niet) die niet uitgelegd kunnen worden door reguliere grammatica's die door middel van regels geconstrueerd zijn (Everaert e.a. 1995, Wray 2008). Onze grammaticale kennis en wereldkennis geven taalgebruikers geenszins een verklaring voor het feit dat we *de strijdbijl begraven* zeggen en niet *het wapen begraven* of *de kruisboog begraven*.

Maar welke expressies nu precies wel en geen idioom zijn en hoe het brede domein van de idiomaticiteit te ordenen en te definiëren is, daarover bestaat nog geenszins consensus. Taalkundigen verschillen vooral van mening over of de expressies die samen het idioom vormen nog steeds betekenis hebben en of er een relatie is tussen de betekenis van die losse expressies en het idioom als geheel (Everaert e.a. 1995:4).

---

<sup>2</sup> Om preciezer te zijn: de betekenis van de meeste idiomen is niet-compositioneel. Hierover meer in paragraaf 2.3.2.

## 2.2 De plaats van idiomen in het lexicon

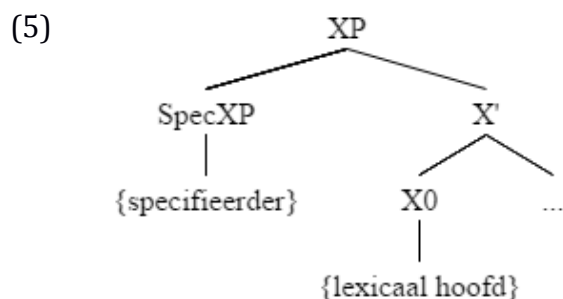
### 2.2.1 De Chomskyaanse traditie: X<sup>0</sup>-insertie

Gerelateerd aan de vraag of de constituenten van een idioom op zich (nog) betekenis dragen, is de kwestie 'waar' idiomen zich bevinden in het talige domein. Volgens de Chomskyaanse traditie bevindt zich in de kern van het talige vermogen van een mens slechts de Universele Grammatica, een taalsysteem dat elk mens bezit, waarin regels liggen vastgelegd die voor elke taal geldig zijn (Chomsky 1981). Idiomen behoren, omdat ze niet bepaald zijn door grammaticale regels, in deze theorie tot de periferie van de taal, het lexicon (Chomsky 1972:39). Het lexicon wordt traditioneel gezien als een lijst van morfemen en woorden die door grammaticale regels gecombineerd kunnen worden tot grote taaleenheden zoals zinnen (Chomsky 1957). Omdat die speciale semantiek van een idiomatische woordencombinatie zich op de een of andere manier toch in het lexicon moet bevinden, is een voor de hand liggende mogelijkheid dat die als gehele VP opgeslagen zijn. Een dergelijk lexicon ziet eruit zoals in (4).

- (4) N → hond, tafel, geest, ...  
V → eten, fietsen, geven, ...  
VP → de pijp uitgaan, de strijdbijl begraven, de geest geven, ...

Het nadeel van een dergelijke voorstelling is dat de idiomatische VP's in het lexicon geen interne syntaxis kennen. Er zouden dan in z'n geheel geen grammaticale operaties op toegepast kunnen worden, zelfs niet de operaties waarvan we weten dat idiomen die wél kunnen ondergaan. Zo zou er bij het idioom *de geest geven* geen inflectie plaatsvinden op *geven* (*hij geeft de geest, zij gaven de geest*) omdat het niet gemarkeerd is als V binnen de VP *de geest geven* (Jackendoff 1997:158).

Ervan uitgaande dat alleen lexicale hoofden, ook wel X<sup>0</sup>'s genoemd (zie (5)), geschikt zijn voor lexicale insertie (d.w.z. het opnemen in de syntactische boom), is een VP in het lexicon sowieso niet gewenst. Chomsky (1981:146) komt dan ook met een ander voorstel: een idioom als *de geest geven* wordt in het lexicon gemarkeerd als een V mét interne structuur (nl. een VP, bestaande uit een NP *de geest* en een V *geven*), zoals weergegeven in (6).



- (6) [V [VP [NP [Det de] [N geest]]] [V geven]]]

Dat het idioom hier in z'n geheel een V-knoop is, betekent dat het in de syntaxis ook alleen maar onder een V ingebracht kan worden. Dit kan vreemde syntactische consequenties opleveren, bijvoorbeeld bij idiomen die onderbroken worden door een

argument. Neem het Engelse *to sell X down the river* ('iemand verraden'), dat er dan uitziet zoals in (7).

(7) [v [v<sub>P</sub> [v sell][PP [P down] [NP [Det the] [N river]]]]]

In het Engels kent men een regel voor Particle Movement (Koster 1975), die zegt dat een voorzetsel in werkwoord-voorzetsel-combinaties zoals *to look up* en *to bring in* naar rechts verplaatst kan worden óver de eerste NP heen, zoals in (8). Is dit eerste pronomen een NP, dan is de verplaatsing zelfs verplicht, zie (9).

(8) a. John looked up the answer.  
b. John looked the answer up.

(9) a. \*The police brought in him.  
b. The police brought him in.

Koster (1975:115)

Door Emonds (1970) is al eerder beredeneerd dat dit soort idiomatische werkwoord-voorzetsel-combinaties (V+P) zijn afgeleid van constructies met een volle PP (V+PP). Omdat een PP in het Engels altijd na het object verschijnt (zie (10)), is het aannemelijk dat *to look up* de onderliggende structuur *to look X up* heeft en niet *to look up X* (Emonds 1970).

(10) a. \*She brought in the house her son/him.  
b. She brought her son/him in the house.

Terug naar (7): wanneer insertie van het thema-argument (het object), bijvoorbeeld *his mother*, plaatsvindt moet de PP *down the river* gepostponeerd (achteropgeplaatst) worden, omdat anders ongrammaticaliteit ontstaat, zoals in (11a) te zien is.

(11) a. \*He sold down the river his mother.  
b. He sold his mother down the river.

Dit postponeren zou dan via een nieuwe verplaatsingsregel moeten gebeuren, terwijl deze lexicaal V met interne structuur zo toch wel heel veel op een normale VP begint te lijken (Jackendoff 1997:159).

Als alleen X<sup>0</sup>-insertie vanuit het lexicon mogelijk is, levert dat dus wat problemen op voor idiomen. Een oplossing zoals 'platte' VP's in het lexicon of V's met interne syntaxis is niet bevredigend, zeker niet als we in acht nemen dat vrijwel alle idiomen een 'gewone' syntactische structuur hebben – de idiomen die we tot nu toe besproken hebben, lijken dezelfde structuur als gewone VP's te hebben (zie ook hoofdstuk 1: *een appel geven* tegenover *de geest geven*), laat staan zinsidiomen (ook wel 'spreekwoorden') zoals *het regent pijpenstelen*, *boontje komt om zijn loontje* en *zoals de waard is, vertrouwt hij zijn gasten* die niets anders kunnen zijn dan idiomen op S-niveau (Jackendoff 1997:160).



## 2.2.2 Theory of Fixed Expressions

Onder andere Jackendoff (1997) verwerpt het idee dat de inhoud van het lexicon beperkt is tot woorden. Met behulp van een enorm corpus aan *fixed expressions* ('vaste uitdrukkingen'), ook wel *formulaic language*<sup>3</sup> genoemd, uit een Amerikaanse televisiequiz<sup>4</sup> laat hij zien dat dat aantal uitdrukkingen zo groot is, dat die onmogelijk allemaal in het lexicon opgeslagen kunnen worden (Jackendoff 1997:156). Maar als deze expressies zich niet in het lexicon bevinden, waar dan wel? Een belangrijke theoretische reden dat idiomatische uitdrukkingen zich wél het lexicon zouden moeten bevinden, is dat idiomen zoals alle andere lexicale elementen een fonologische, syntactische en semantische structuur hebben en er geen andere faculteit binnen de menselijke geest is, bijvoorbeeld een meer algemeen deel van het geheugen, waarin deze eigenschappen opgeslagen kunnen worden (Jackendoff 1997:158).

Een lexicale eenheid kan in deze theorie dus groter zijn dan één grammaticaal woord ( $X^0$ ), dus ook een PP, VP of zelfs S. Jackendoff (1997) scheidt hiermee de algemeen gangbare Theory of Words ( $X^0$ -insertie) van zijn Theory of Fixed Expressions (ook grotere talige expressies kunnen lexicale insertie ondergaan) en trekt hiermee de harde grens tussen syntaxis en lexicon in twijfel. Hiermee wordt bedoeld dat idiomatische expressies op een hoger niveau dan het woordniveau hun intrede doen in de syntactische boom. De lexicale insertie vindt gebeurt dan op een later moment dan bij de creatie bij een  $X^0$ . Dit zou een verklaring kunnen zijn voor het feit dat niet alle grammaticale operaties toepasbaar zijn op idiomen, omdat die bijvoorbeeld plaats zouden moeten vinden voordat de lexicale insertie van het idioom plaatsvindt.

Gazdar, Klein, Pullum & Sag (1985) menen daarentegen dat idiomen niet als geheel in het lexicon zijn opgeslagen, maar dat elke constituent op zich is opgeslagen met een 'aantekening' dat het in combinatie met een ander woord een andere betekenis krijgt. Ook Everaert (1991) houdt deze theorie aan. Zo zouden *geest* en *geven* beiden met hun letterlijke betekenis in het lexicon zijn opgenomen, waarbij wordt aangemerkt dat wanneer deze lexicale elementen gecombineerd worden, er een andere, idiomatische betekenis (nl. 'sterven') ontstaat. In dat geval hebben de losse constituenten van een idioom dus nog wel betekenis, maar in een vrije woordcombinatie een andere betekenis dan in combinatie met het andere idioomdeel.

Een argument tegen een dergelijke benadering is dat een ondenkbaar grote hoeveelheid informatie in het lexicon opgeslagen moet worden, omdat bij elk woord dat niet alleen op zichzelf, maar ook in een idioom voorkomt (zoals *geest* in *de geest geven*) apart aangemerkt moet worden dat dat woord in combinatie met een ander woord (*geven*) een andere betekenis krijgt en dat die twee woorden samen een VP vormen. Een complexer idioom zoals *een kat in de zak kopen* kost nog veel meer opslagruimte: bij elk

---

<sup>3</sup>Een gangbare definitie van formulaic is: "een al dan niet onderbroken reeks van woorden of andere betekenisvolle elementen die geprefabriceerd is of dat lijkt te zijn d.w.z. als geheel opgeslagen en opgehaald uit het geheugen bij het gebruik in plaats van gegeneerd of geanalyseerd door grammatica" Wray & Perkins (2000).

<sup>4</sup>Het gaat hier om het *Wheel of Fortune*-corpus, bestaande uit zo'n 600 uitdrukkingen, verzameld gedurende slechts een paar maanden. Omdat de show al zo'n tien jaar op televisie is (ten tijde van Jackendoffs publicatie) en zes dagen per week wordt uitgezonden, moet het totaal aantal quiz-antwoorden 10.000 tot 15.000 uitdrukkingen beslaan (Jackendoff 1995:154-156).

afzonderlijk woord moeten alle andere woorden van het idioom genoemd worden en dat al die woorden alleen met z'n allen de betekenis 'een miskoop doen' vormen. Daarnaast moeten er nog syntactische regels voor het vormen van de VP worden opgenomen in het lexicon, zodat voorkomen wordt dat men *de kat in een zak kopen* kan vormen, of erger: *een zak in de kat kopen* (Jackendoff 1997).

Al met al lijkt het niet wenselijk om idiomen als functies van losse woorden te beschouwen die in combinatie een andere betekenis vormen dan in isolatie, om zo de lexicale insertie met louter  $X^0$  te behouden.

## 2.3 Syntactische beperkingen en betekenis

### 2.3.1 Compositionele en non-compositionele idiomen

Het is bekend dat veel idiomen moeite hebben met syntactische operaties. Zowel topicalisatie (vooropplaatsing in topic-positie) in (12b), clefting ('Het was \_ dat...') in (12c) en relativisatie ('De/het \_ die/dat...') in (12d) zijn niet mogelijk bij het idioom *to kick the bucket*. Ook gevallen van links-dislocatie in (13b) en (13c) zijn niet mogelijk.

- (12) a. Pete kicked the bucket.  
b. \*The bucket Pete kicked.  
c. \*It was the bucket Pete kicked.  
d. \*The bucket Pete kicked was astonishing. Schenk (1995:254)

- (13) a. Het huis, dat kan Jan kopen  
b. \*Het heen en weer, dat kan Jan krijgen.  
c. \*Zijn kerfstok, daar heeft Jan heeft veel op. Vat (1997:71)

Gazdar, Klein, Pullum & Sag (1985) en Wasow, Sag & Nunberg (1984) menen dat het bestaan van twee typen idiomen dit verschil verklaart: compositionele en niet-compositionele idiomen. Omdat de eerste soort (waaronder *to spill the beans* valt) niet semantisch te analyseren is, kunnen er syntactische operaties op uitgevoerd worden, zoals passivisatie (14a). De tweede soort (waaronder *to kick the bucket* valt) kan dit niet (14b). Ze beredeneren dat de constituenten van compositionele idiomen wél betekenis dragen, in tegenstelling tot die van idiomen van de tweede soort.

- (14) a. The beans were spilled by John.  
b. \*The bucket was kicked by John.

### 2.3.2 Betekenisloze constituenten

Schenk (1995) gaat echter tegen de conclusies van Gazdar e.a. (1985) en Wasow e.a. (1984) in. Als het inderdaad gaat om het onderscheid tussen compositionele en niet-compositionele idiomen, beargumenteert hij, dan zou men op basis daarvan verwachten dat compositionele idiomen élk soort operatie zouden moeten kunnen ondergaan. En dat is niet het geval, zoals te zien is in (15).

- (15) a. The beans were spilled by John.  
 b. \*The beans John spilled. Schenk (1995:59)  
 c. \*John instructs the beans to be spilled. Schenk (1995:261)

Schenk (1995) beargumenteert dat dit een gevolg is van de niet-compositionaliteit van álle idiomen, die als geheel één abstracte betekenis uitdrukken. Hij stelt dus dat de constituenten van een idioom geen betekenis dragen en daarom niet in posities kunnen voorkomen die voor betekenis dragende elementen bedoeld zijn, zoals de topic-positie<sup>5</sup> in (15b) en de control-positie in (15c)<sup>6</sup>.

Dat sommige syntactische operaties wél mogelijk zijn bij bepaalde idiomen, zie (5), moet dus een andere oorzaak hebben dan Gazdar e.a. (1985) en Wasow e.a. (1984) voorstelden.

- (16) a. John laid his failure at Mary's door.  
 b. At Mary's door John laid his failure.  
 c. Pete pulled Mary's leg.  
 d. Mary's leg Pete pulled. Schenk (1995:260)

Een logische verklaring voor de grammaticaliteit van deze zinnen is dat de gemanipuleerde delen van de expressie hier wél betekenis dragen, namelijk door de thètarol (van *Mary*) die ze bij zich hebben als gevolg van Pied Piping (Schenk 1995:260).

Pied Piping is een veel voorkomend fenomeen in talen (waaronder het Nederlands en Engels), waarbij de constituenten van complexe NP's elkaar volgen als één ervan verplaatst wordt (Van Riemsdijk & Williams 1986). Zo bestaat de complexe NP *Mary's leg* uit de twee NP's *Mary* en *leg*. Wanneer *Mary* getopicaliseerd wordt, moet *leg* ook meeverplaatsen, zie (17a). Vergelijk ook met Pied Piping bij de verplichte wh-verplaatsing in (17b).

- (17) a. [**Mary's leg**]<sub>i</sub> Pete pulled [<sub>i</sub>].  
 b. [**Welk boek**]<sub>i</sub> heeft hij gelezen [<sub>i</sub>]?

Dat (16b) en (16d) mogelijk zijn, komt dus omdat een deel van getopicaliseerde PP (*At Mary's door*) of NP (*Mary's leg*) betekenis draagt en daarmee de constituent de gelegenheid geeft zich naar posities te verplaatsen die voor betekenisdragende elementen gereserveerd zijn.

### 2.3.3 Of toch betekenisvolle constituenten?

Feit is dat veel idiomen bepaalde operaties wél kunnen ondergaan, ook als ze geen thètaroldragend element bij zich hebben, zoals in (16) het geval is. McClure (2011) houdt vast aan het eerder voorgestelde verschil tussen compositionele en niet-

---

<sup>5</sup> Zo kan het betekenisloze existentieel 'there' ook niet getopicaliseerd worden: *He believes there to be unicorns* / \**There he believes to be unicorns* (Schenk 1995:259).

<sup>6</sup> In control-constructies moeten zowel het controllerende element als het gecontroleerde element betekenis dragen. Zo kan het betekenisloze 'it' niet als controllerend/gecontroleerd element fungeren: \**John tries it to rain* / \**It tries to rain* (Schenk 1995:261; Van Riemsdijk & Williams 1986:129).

compositionalen idiomen. De eerste soort wordt gekenmerkt door een herkenbare betekenis in de afzonderlijke delen (18a-b). De tweede soort heeft dat niet (18c-d).

- (18) a. John buried the hatchet.  
b. The hatchet was buried by John.  
c. John kicked the bucket.  
d. \*The bucket was kicked by John.

In *to bury the hatchet* (vergelijkbaar met *de strijdbijl begraven*) kan vanwege het hoge metaforische karakter nog enige betekenis toegekend worden aan de verschillende constituenten *to bury* (zoiets als ‘wegstoppen’) *the hatchet* (zoiets als ‘onenigheid’). Voor Engelstaligen is *to bury the hatchet* dus in hogere mate compositioneel dan *to kick the bucket*, waarin de constituenten (schijnbaar<sup>7</sup>) op geen enkele manier bijdragen aan de betekenis van het geheel. McClure (2011) merkt ook op dat dit verklaart waarom sommige idioomdelen wel gemodificeerd kunnen worden (19a-d), terwijl andere daar niet vatbaar voor zijn (19e).

- (19) a. John kept tabs on Mary.  
b. John kept close tabs on Mary.  
c. John gave Mary flak.  
d. John gave Mary well-deserved flak.  
e. \*John kicked the rusty bucket.

McClure (2011:2)

Omdat *flak* gezien kan worden als een synoniem voor ‘kritiek’, kan *flak* ook gemodificeerd worden zoals ‘kritiek’. Bij *the bucket* is dit niet het geval omdat het niet geassocieerd kan worden met een deelvertaling van *to kick the bucket* ‘sterven’<sup>8</sup>. Ook volgens McClure (2011) dragen de constituenten van sommige (nl. compositionele) idiomen dus wel betekenis.

Ook Jackendoff (1997) gaat uit van een interne semantiek door het metaforische (compositionele) karakter van sommige idiomen en verwijst daarbij ook naar de mogelijkheid tot passivisering van het idioom in (18b). *the hatchet* heeft door zijn metaforische karakter ‘gewoon’ een *thètarol* en kan dus allerlei syntactische operaties zoals passivisatie ondergaan. Dan blijft er nog een ander probleem met dit idioom: als het compositioneel is, waarom kan het dan niet alle syntactische operaties, zoals topicalisatie ondergaan (zie (20))?

- (20) \*The hatchet John buried.

---

<sup>7</sup> In feite is elk idioom een metafoer die is vastgeroest in een taal en die dus ooit een compositionele betekenis had. Zo is *iemand de loef afsteken* (iemand te snel/goed af zijn) niet meer herkenbaar als compositioneel idioom omdat men tegenwoordig niet meer bekend als een synoniem voor ‘wind’ (of ‘windzijde van een schip’). *Iemand de loef afsteken* komt dan ook uit de zeevaart, waarbij het ene schip het andere inhaalde door de wind uit iemand zeilen over te nemen, zodat de concurrent geen wind meer ving (Stoett 1923-1925:567-568).

<sup>8</sup> McClure (2011) noemt wel modificaties op niet-compositionele idiomen zoals *John kicked goddamned bucket* en *John kicked the metaphorical bucket* maar legt uit dat het hier gaat om expressieve respectievelijk metalinhuïstische modificaties, die niet het idioom(deel) modificeren maar de gehele expressie (i.e. zin).

De verklaring die Jackendoff (1997) geeft, heeft niet te maken met het al dan niet kunnen dragen van thètarollen, maar met de onmogelijkheid tot distributie van het getopicaliseerde element. *The hatchet* kan, ook al heeft het een thètarol, niet in semantisch contrast staan met een ander element omdat het niet inwisselbaar is – daar is het tenslotte een idiomatische expressie voor – en kan dus niet in een positie voor nadruk of contrast (zoals topicalisatie of focus) staan. Dit kan echter wel in het geval van een thètaroldragende getopicaliseerde constituent zoals in (5), waar *Mary's leg* in contrast kan staan met bijvoorbeeld *Tim's leg* of *Ellen's leg*.

Ook al zijn er op dit gebied nog wat open plekken, de strekking door de literatuur heen lijkt te zijn dat hoe minder compositioneel een idioom is, hoe minder syntactische operaties erop uitgevoerd kunnen worden.

## 2.4 Idiomen en collocaties

Als de mate van compositionaliteit bepalend is voor de mogelijkheid tot grammaticale operaties op een idioom, kan het nuttig zijn te kijken naar expressies waarvan we zeker weten dat ze zich op de schaal van idiomaticiteit bevinden maar in betekenis wel zeer doorzichtig zijn: collocaties. Belangrijk is hier wat precies het verschil is tussen een idioom en een collocatie. Daar bestaan verschillende ideeën over.

### 2.4.1 Co-occurrence

Een idioom in algemene zin verwijst naar een woordgroep waarvan de combinatie van woorden iets anders betekent dan de losse woorden op zich. Er is dus sprake van een abnormale semantiek en (in meerdere of mindere mate) van niet-compositionaliteit (Schenk: 255-256). Bij een collocatie daarentegen speelt niet-compositionaliteit geen rol, menen Cantos en Sánchez (2002). Een collocatie bestaat 'slechts' uit twee of meer woorden die een sterke neiging hebben samen voor te komen, wat ook wel een hoge co-occurrence genoemd wordt (Cantos & Sánchez 2002:201). Nu is dit een wel heel ruime, technische definitie die vooral in de corpuslinguïstiek gebruikt wordt. Een woordcombinatie als *ik ben* heeft statistisch gezien ook een sterke neiging om samen voor te komen, en toch noemt men die in het kader van idiomaticiteit over het algemeen geen collocatie.

Benson, Benson, and Ilson (1986) gebruiken voor collocaties de fraseologie werkzamere definitie van “gefixeerde, identificeerbare, niet-idiomatische frases en constructies”. Met niet-idiomatisch doelen zij op de enge notie ‘idioom’ en niet op de ruimere die wij hier gebruiken. Benson e.a. (1986) leggen uit dat zij met collocaties woordgroepen bedoelen die een optelsom zijn van de betekenissen van hun constituenten (compositioneel), maar dat die constituenten wel vaak samen gebruikt worden en eerder gebruikt worden dan vrije combinaties van woorden die precies hetzelfde zouden betekenen, ware het niet dat zij geconventionaliseerd zijn.

Benson e.a. (1986) maken onderscheid tussen grammaticale collocaties en lexicale collocaties. Grammaticale collocaties zijn een combinatie van V, A of N met een functiewoord, zoals een P. Het gaat dus om collocaties zoals *rekenen op*, *bang zijn voor* of *op school*. Lexicale collocaties bestaan uit meerdere lexicale elementen, waarin veel

variatie mogelijk is. Zo is een combinatie van A+N mogelijk, zoals in *een verpletterende nederlaag*, maar ook combinaties van onder andere N+N (*een school vissen*), Adv+V (*luidkeels roepen*) en V+N (VP's) zoals *vrede stichten*, die het meest lijken op de idiomen die hier tot nu toe behandeld zijn.

Een moedertaalspreker zal vrijwel nooit de voorkeur geven aan een vrije combinatie van woorden (zoals *krachtige thee*) als er een collocatie voorhanden is die in principe dezelfde betekenis draagt (*sterke thee*). Deze voorkeur is vaak zo sterk dat alternatieve formuleringen als ongrammaticaal gezien worden door de onverwachte combinatie (voor de bedoelde betekenis). Wat in het dagelijks taalgebruik vaak als 'verkeerde woordkeus' wordt bestempeld, is in feite een deel van een collocatie gecombineerd met een woord dat daar niet 'opgeroepen' hoort te worden, zie (21) en (22).

(21) a. de tafel dekken  
b. ?de tafel bekleden/opmaken

(22) a. zware roker  
b. ?zware eter

#### 2.4.2 Encodeerbaarheid en decodeerbaarheid

Cruse (2000) houdt er voor zowel idioom als collocatie een iets andere en technischere definitie op na en werkt met het verschil tussen encoding en decoding. Decoding is betekenis-toekennend proces vanuit het perspectief van de ontvanger van een complexe expressie. Decoding van de expressie is het omzetten van talige uitingen naar een abstracte betekenis. Dit herleiden van de betekenis gebeurt meestal door de betekenis van losse constituenten, een reeks woorden, 'bij elkaar op te tellen'. Dat gebeurt in ieder geval bij compositionele expressies. In het geval van een idioom kan de ontvanger niet decoderen als hij het idioom niet kent, omdat het niet compositioneel is.

Encoding daarentegen is juist het omgekeerde en draait om het proces dat de spreker doorloopt als hij een 'betekenisgeheel' in zijn hoofd heeft en dat wil uiten. Op dat moment encodeert hij die abstracte betekenis naar een reeks woorden. In het geval van een idioom bedenkt de spreker dus de betekenis 'sterven', waar hij woorden aan wil toekennen. Die abstracte boodschap encodeert hij dan naar de grammaticale constructie *de pijp uitgaan*. Natuurlijk kan de spreker ook kiezen voor andere verwoordingen zoals *doodgaan*, *de geest geven* of gewoonweg *sterven*.

Dit levert een interessante interpretatie op van het verschil tussen een idioom en een collocatie. Een idioom is in dat geval een expressie die je niet kunt decoderen én niet kunt encoderen als je deze niet kent. Een collocatie is een expressie die niet te encoderen is (denk aan foute woordkeuzes die tweedetaalleerders vaak maken in het Nederlands), maar wel te decoderen (Cruse 2000: 76).



## 2.5 Verwerking en beoordeling van idiomatische uitdrukkingen

In de afgelopen jaren wordt steeds meer psycholinguïstisch onderzoek gedaan, ook naar de verwerking van *formulaic language* (idiomatische uitdrukkingen), zoals behandeld in paragraaf 2.2.2. Deze onderzoekers zijn vooral geïnteresseerd in de voordelen van deze 'geprefabriceerde' uitdrukkingen voor het werkgeheugen, vandaar dat ze in de psycholinguïstische literatuur veelal 'prefabs' genoemd worden.

Deze expressies, die ik hier voor de consistentie idiomatische uitdrukkingen blijf noemen, werden in verschillende onderzoeken getest op leestijden van proefpersonen. Millar (2011) bijvoorbeeld manipuleerde collocaties door een alternatieve woordcombinatie aan te bieden die hij uit een geschreven corpus van tweedetaalleerders had gehaald. Hij liet moedertaalsprekers in een self-paced reading test (Jegerski 2014) zinnen met correcte collocaties en zinnen met foutieve collocaties lezen. De leestijden lieten zien dat de foutieve collocaties langzamer gelezen en verwerkt werden dan de correcte collocaties (Millar 2011:142). Deze resultaten ondersteunen de al wijdverspreide gedachte dat *formulaic language* voordelen bij taalverwerking bieden en bevestigt onder andere de definitie van Benson e.a. (1986) dat taalgebruikers liever collocaties verwerken dan originele combinaties van woorden.

Gibbs & Nayak (1989) onderzochten eerder al of VP-idiomen hun figuurlijke betekenis behouden wanneer ze syntactisch gemanipuleerd worden. Zij deden experimenten met onder andere passivisatie en pronominalisatie binnen idiomen en zagen dat er grote verschillen waren in de beoordeling tussen verschillende idiomen. De conclusie die Gibbs & Nayak (1989) trekken is dat de mate waarin idiomen als syntactisch flexibel worden beoordeeld in grote mate afhangt van de mate van *semantic decompositionality*. Er wordt hier gesproken over decompositionaleiteit in plaats van compositionaleiteit, omdat er niet per se wordt uitgegaan van de productie van idiomen – hoe de constituenten semantisch bijdragen aan het geheel – maar eerder van de manier waarop de constituenten kunnen helpen bij het semantisch analyseren van het geheel. Ook hier wordt dus geconcludeerd dat idiomen wel degelijk interne semantiek hebben, maar dat de mate daarvan (i.e. de mate waarin de spreker die interne semantiek kan ontrafelen) onder andere bepalend is voor de syntactische flexibiliteit van het idioom.

Zowel Gibbs & Nayak (1989) als Everaert e.a. (1995) stellen een glijdende schaal van idiomaciteit voor. Fraser (1970) ziet ook dat de syntactische flexibiliteit van idiomen erg verschilt per geval, maar stelt een zevenpuntsschaal van Frozenness Hierarchy voor van level 0 tot level 6, waarin level 0 idiomen bevat die absoluut geen syntactische operaties kunnen ondergaan (en dus *frozen* zijn) en waarin level 6 géén idiomen bevat, omdat die dan alle mogelijk grammaticale operaties moeten zouden kunnen ondergaan – en we weten dat zulke idiomen niet bestaan (Fraser 1970:37).

Opvallend is dat Fraser (1970) in de meest 'bevroren' categorie, level 0, idiomen zoals *dip into X's pocket* en *bite (off) X's tongue* plaatst, terwijl we weten dat andere idiomen met een nog in te vullen argument, zoals *to pull X's leg* wél topicalisatie kunnen ondergaan door de thètarol die het nog in te vullen argument draagt (Jackendoff 1997; Schenk 1995; zie (x)).

Het indelen van idiomen op basis van operationele criteria – Fraser (1970) gebruikt onder andere Particle Movement, passivisatie en nominalisatie als mogelijke operaties – lijkt een mooi handvat om idiomen in klassen in te delen, maar zoals voorgesteld door Fraser (1970) lijkt die in ieder geval (nog) niet werkzaam en lijkt een glijdende schaal zonder specifieke punten (Gibbs & Nayak 1989) een beter houvast om idiomatieit te behandelen.



## 3. Vraagstelling en hypothese

### 3.1 Vraagstelling

Uit de literatuur wordt duidelijk dat a) idiomen bepaalde syntactische beperkingen kennen en dat b) de hoeveelheid en soort syntactische beperkingen verschillen per idioom, waarvan de oorzaak waarschijnlijk te maken heeft met (de mate van) compositionaliteit. Ook is duidelijk dat idiomatische taalvormen zich op een glijdende schaal bevinden, waarop aan de ene kant idiomen te vinden zijn en aan de andere kant vrije, originele woordcombinaties.

Schenk (1995), Cruse (2000) en Beson e.a. (1986) suggereren dat collocaties zich tussen idiomen en vrije woordcombinaties in bevinden op deze schaal. Schenk (1995) stelt daarnaast dat taalgebruikers manipulaties op bepaalde uitdrukkingen binnen het idiomatische domein goedkeuren omdat het in hun grammatica collocaties zijn. Of een expressie een idioom of collocatie is – bijvoorbeeld volgens de definitie van Cruse (2000) – zou men dus moeten kunnen voorspellen in hoeverre er syntactische operaties op uitgevoerd kunnen worden dan wel interpretatieve relaties gelegd kunnen worden.

Er is al eerder onderzoek gedaan naar de verwerking van gemanipuleerde idiomatische uitdrukkingen (Millar 2011, Wray & Perkins 2000), maar nog niet eerder naar grammaticaliteitsoordelen. In dit onderzoek ga ik dan ook de aanvaardbaarheid van grammaticale operaties op idiomen én op collocaties testen.

#### **Hoofdvraag**

Zijn grammaticale operaties op collocaties aanvaardbaarder dan grammaticale operaties op idiomen?

‘Aanvaardbaar’ wil hier zeggen dat een zin grammaticaal blijft na uitvoering van de manipulatie. Als taalgebruikers de ene taaluiting ‘aanvaardbaarder’ vinden dan een andere, betekent dat dat zij die ene uiting als correcter of grammaticaler (ook wel: welgevormder) beoordelen (binnen hun moedertaal) dan een andere.

Om deze hoofdvraag te beantwoorden richt ik mij op twee soorten manipulaties: syntactische en interpretatieve manipulatie. Van de eerste soort, is al gebleken dat die moeilijk toepasbaar is op idiomen, zoals bij topicalisatie (Schenk 1995:254,259). Over de mogelijkheid tot het leggen van interpretatieve relaties op idiomen, zoals pronominalisatie, is nog niet veel bekend. Topicalisatie en pronominalisatie zijn dan ook de twee soorten operaties die ik ga gebruiken om mijn onderzoeksvraag te staven.

De onderstaande deelvragen zullen helpen de hoofdvraag van dit onderzoek te beantwoorden.

#### **Deelvraag I**

Is het topicaliseren van een deel van een collocatie aanvaardbaarder dan het topicaliseren van een deel van een idioom?

### **Deelvraag II**

Is een pronominale verbinding met een onderdeel (NP) van een collocatie aanvaardbaarder dan een pronominale verbinding met een onderdeel van een idioom?

### **Deelvraag III**

Zijn topicalisatie en pronominale verbinding in zinnen zonder een idioom of collocatie aanvaardbaarder dan in zinnen met een idioom of collocatie?

Deelvraag I en II vloeien direct voort uit de hoofdvraag en de twee gekozen grammaticale operaties, topicalisatie en pronominalisatie. Deelvraag III stel ik omdat ik niet alleen idiomen en collocaties met elkaar wil vergelijken, maar ook deze beide met constructies zonder idiomen of collocaties. Het lijkt evident dat gemanipuleerde idiomen en gewone zinnen sterk van elkaar verschillen in grammaticaliteit, maar of collocaties beperkingen op syntactische flexibiliteit kennen, is nog onbekend. Ook over de semantische doorzichtbaarheid van collocaties is de literatuur niet sluitend over.

Zo zegt Schenk (1995) niets over de mogelijkheid tot topicalisatie of control-constructies; hij noemt alleen de mogelijkheden in hoofdstuk 1 (de zinnen in (3)). Cruse (2000) noemt collocaties weliswaar decodeerbaar, maar niet encodeerbaar, wat de syntactische flexibiliteit wellicht alsnog beïnvloedt. Door de idiomen en collocaties ook met gewone zinnen te vergelijken, hoop ik meer uitsluitsel te kunnen geven over de compositionaliteit van collocaties.

## **3.2 Hypothese**

Cruses (2000) definitie van collocaties stelt dat deze wél decodeerbaar en dus compositioneel zijn, terwijl (de meeste) idiomen dat niet zijn. Gazdar e.a. (1985) en Wasow e.a. (1984) menen dat of een idioom wel of niet compositioneel is, bepaalt of een idioom syntactische operaties kan ondergaan of niet. Alhoewel Jackendoff (1997) niet stelt dat alléén compositionaliteit doorslaggevend is voor de mate van syntactische flexibiliteit, zegt ook hij dat idiomen door een metaforisch, compositioneel karakter in staat zijn om syntactische operaties te ondergaan.

Schenk (1995) stelt dat onverklaarbare syntactische manipulaties op idiomen zoals modificatie van een NP (*hij draaide mij een verschrikkelijke loer*) correct kunnen zijn voor mensen die het idioom *X een loer draaien* beschouwen als een collocatie. Op basis van Schenks (1995) aanname en het feit dat compositionaliteit de syntactische flexibiliteit van een expressie bevordert, verwacht ik dat grammaticale operaties aanvaardbaarder zijn op collocaties dan op idiomen.

### **Hoofdhypothese**

Grammaticale operaties op collocaties zijn aanvaardbaarder dan grammaticale operaties op idiomen.

Uit deze hypothese vloeien de volgende deelhypotheses voort.

### **Deelhypothese I**

Het topicaliseren van een deel van een collocatie is aanvaardbaarder dan het topicaliseren van een deel van een idioom.

Jackendoff (1997) en Schenk (1995) hebben laten zien dat het topicaliseren van idioomdelen in de meeste gevallen niet mogelijk is. Op basis van de veronderstelde flexibiliteit van collocaties verwacht ik dan ook dat topicalisatie beter toepasbaar is op constituenten van collocaties dan van idiomen.

### **Deelhypothese II**

Een pronominale verbinding met een onderdeel (NP) van een collocatie is aanvaardbaarder dan een pronominale verbinding met een onderdeel van een idioom.

Gibbs & Nayak (1989) concludeerden al eerder dat idiomen die semantisch gemakkelijk te analyseren zijn, pronominalisatie redelijk goed ondergaan. Dit komt omdat de antecedenten in dat geval semantisch herkenbare eenheden zijn. Bij idiomen die niet-compositioneel zijn, zijn gepronominaliseerde elementen minder aanvaardbaar. Daarom verwacht ik dat collocaties, die in elk geval (de)compositioneel zijn, aanvaardbaarder zijn met een pronominale verbinding dan idiomen.

### **Deelhypothese III**

Topicalisatie en pronominale verbinding in zinnen zonder een idioom of collocatie zijn aanvaardbaarder dan in zinnen met een idioom of collocatie.

Omdat collocaties toch enigszins vaste woordgroepen zijn (Benson e.a. 1986) verwacht ik dat het toch enige verwarring kan opleveren voor de taalgebruiker wanneer er grammaticale operaties op uitgevoerd worden. Ook Cruse (2000) stelt dat collocaties dan wel decodeerbaar zijn, maar niet zonder meer encodeerbaar. Mijn verwachting is daarom dat in 'gewone' zinnen, zonder idiomen of collocaties, grammaticale operaties met nog minder problemen worden aanvaard dan in zinnen met collocaties.

## 4. Methode

### 4.1 Participanten

De informantengroep, die de grammaticaliteit van een aantal testzinnen beoordeelde, bestond uit 39 informanten tussen de 18 en 25 jaar. Door deze leeftijdsgrenzen kan ik er zekerder van zijn dat de informanten met globaal dezelfde kennis over de gebruikte idiomen en collocaties zijn opgegroeid. Onder de participanten waren 11 mannen en 28 vrouwen, en zij waren allen student op het hbo of wo.

De deelnemers werden via e-mail benaderd om de questionnaire in te vullen. Er stond geen vergoeding tegenover deelname aan het onderzoek.

### 4.2 Materiaal

Om de (deel)vragen uit hoofdstuk 2 te kunnen beantwoorden moesten er voor de testzinnen expressies gekozen worden, waarvan men volgens de definitie van Cruse (2000) zeker weet dat het collocaties zijn, omdat ze alleen decodeerbaar (ofwel receptief compositioneel) zijn, en expressies waarvan men zeker weet dat het idiomen zijn, omdat ze decodeerbaar noch encodeerbaar (ofwel zowel receptief als productief niet-compositioneel) zijn. De idiomen zijn uit verscheidene spreekwoordenboeken gehaald (Stoett 1923-1925; Lommerse & Vernooij 2014; Harrebomée 1990) en op basis van gelijkenis gekozen als paren met de collocaties.

Omdat het moeilijk is om te bepalen hoe *frozen* of compositioneel een uitdrukking is, zijn er collocaties gekozen waarvan in ieder geval één deel van de uitdrukking zeer letterlijk te nemen is (zoals *de tafel* in *de tafel dekken*). De aanname is dat dat het makkelijker maakt om operaties op de collocaties uit te voeren. Op intuïtie zijn een aantal idiomen uitgekozen die niet-compositioneel leken te zijn – en waarvan de participanten, die in dezelfde generatie zitten als ik, daar waarschijnlijk ook zo over denken.

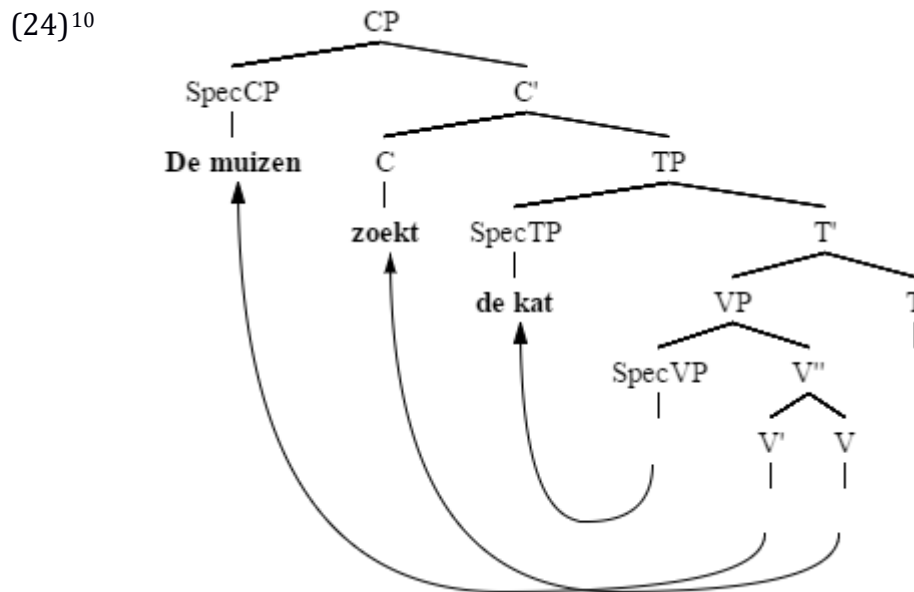
In de questionnaire zijn in totaal 42 testzinnen opgenomen, verdeeld over drie verschillende condities, waarbinnen een driedeling is gemaakt in zinnen met topicalisatie (Top-conditie, bestaande uit 18 zinnen), zinnen met anaforen (Anacconditie, 18 zinnen) en neutrale fillers (Fill-conditie, 6 zinnen). De eerste twee condities bestonden uit drie subcondities, die een idioom (Idi), een collocatie (Col) ofwel geen van beide bevatten (Fill). In totaal zijn er dus zeven condities.

#### 4.2.1 Top-conditie

Topicalisatie is een operatie van verplaatsing waarbij een element, net als bij wh-verplaatsing naar de SpecCP wordt verplaatst (Chomsky 1986) zodat dat element als topic van de zin (de kern van de zin, datgeen waar de zin over gaat) geïnterpreteerd

wordt.<sup>9</sup> In (23) en (24) is te zien hoe topicalisatie in zijn werk gaat. Het getopicaliseerde element wordt initieel gegenereerd in canonieke positie (in het geval van *de muizen* in V'-positie) en vervolgens gekopieerd in topicpositie, waarna het origineel weggelaten wordt in de Phonological Form (de uitgesproken talige uiting). In (24) zie je nog meer verplaatsingen, maar die hebben allemaal een ander doel, waar ik nu niet over zal uitweiden.

- (23) a. De kat zoekt de muizen.  
 b. De muizen zoekt de kat.



Het getopicaliseerde element wordt behalve door verplaatsing overigens vaak ook prosodisch gemarkeerd door de zinsklemtone.

In elk van de 18 zinnen in de Top(icalisatie)-conditie was het complement van de V, of dat nu een DP of PP was, getopicaliseerd. In een derde van zinnen was het getopicaliseerde element een deel van een idioom (conditie Idi-Top), in een derde een deel van een collocatie (conditie Col-Top) en in derde een deel van een gewone VP (Fill-Top).

<sup>9</sup> Er zijn echter ook taalkundigen die hebben voorgesteld dat de linkerperiferie van de zin (de C-projectie) uitgebreid moest worden met aparte projecties voor de informatiestructurele noties topic en focus. Bij zo'n aanname verplaatst het getopicaliseerd element naar de specificeringspositie van de topic-projectie, SpecTopP (Rizzi 1997). Voor dit onderzoek maakt het geen verschil welke theorie aangenomen wordt, maar vanwege de complexiteit van een cartografische benadering zoals die van Rizzi ga ik hier uit van de Chomskyaanse benadering.

<sup>10</sup> De oplettende lezer zal opmerken dat bij de standaardvolgorde in de Nederlandse hoofdzin, subject-(hulp)werkwoord-object, *de kat* in (11a) zich wel in de SpecCP moet bevinden en dus in de dezelfde positie als het topic staat. In een zin met canonieke woordvolgorde wordt het subject inderdaad als topic gezien (Langacker 2001). Overigens zijn er veel voorstellen geweest, waaronder die van Chomsky (1977), om de syntactische boom uit te breiden met een extra CP-laag waar het topic (en wh-elementen) naar zou verplaatsen, maar dat doet voor dit onderzoek niet ter zake dus ik bespreek dit hier verder niet.

a. **Idi-Top**

6 zinnen met een idioom,  
waarin de direct-objects-DP (3x) of de PP (3x) van het idioom getopicaliseerd is.  
Met de pet gooiden ze naar de herkansing van dat lastige statistiekvak.

In conditie Idi-Top bevonden zich zinnen met idiomen, zoals *met de pet gooien naar X*, dat maar liefst twee PP's kent. De PP die hier het dichtst bij het lexicale hoofd V *gooien* bevindt is *met de pet* en die constituent werd dan ook getopicaliseerd. *Met de pet* bevond zich dan in SpecCP en *gooiden* in de C, waarop de rest van de zin volgde.

b. **Col-Top**

6 zinnen met een collocatie,  
waarin de direct-objects-DP (3x) of de PP (3x) van de collocatie getopicaliseerd is.  
Met de deur gooide hij zo hard dat alle fotolijstjes van de vensterbank vielen.

De Col-Top-conditie bestond uit zinnen met collocaties, zoals *met de deur gooien*. Ook hier werd de PP *met de deur* naar de SpecCP verplaatst om zo gemarkeerd te worden als topic. *Gooide* nam wederom de C-positie in.

c. **Fill-Top**

6 zinnen zonder idioom of collocatie,  
waarin de direct-objects-DP (4x), de PP (1x) of de AP (1x) getopicaliseerd is.  
Het hele brood had Gabriël met liefde willen opeten, maar alle starende mensen hielden hem tegen.

Ook uit 'gewone' VP's werd getopicaliseerd, zoals in de voorbeeldzin hierboven gebeurd is met de VP *het hele brood (hebben willen) opeten*. De DP *het hele brood* heeft topicwaarde in de SpecCP-positie, waarna de V *had*, het subject en de rest van de zin volgde.

#### 4.2.2 Ana-conditie

Bij operatie pronominalisatie wordt een NP of DP vervangen door een pronomen, zie (25). Een pronomen heeft altijd een referent buiten zijn minimale domein (de CP)<sup>11</sup>, wat ook wel 'vrije binding' wordt genoemd. In (25b) bijvoorbeeld kan het persoonlijk pronomen *haar* in geen geval dezelfde referent hebben als *Marie* d.w.z. *haar* en *Marie* kunnen niet naar dezelfde persoon in de werkelijkheid verwijzen.

- (25) a. Marie ziet Mieke niet staan.  
b. Marie ziet haar niet staan.  
c. Jan legt de sleutels op tafel.  
d. Jan legt ze op tafel.

In de testzinnen in deze conditie wordt de NP of DP vervangen door een aanwijzend pronomen *die* of *dat*.<sup>12</sup> Een inherente eigenschap van pronomina is dat ze niet

---

<sup>11</sup> Technischer verwoord: volgens de Bindingstheorie (Chomsky 1981) mag een pronomen  $\beta$  niet gebonden worden door een antecedent  $\alpha$  binnen zijn bindende categorie  $\gamma$ .

<sup>12</sup> De zinnen vormen hierop de uitzondering. In de zin 'Ik zag de eerste tien minuten geen steek, maar de rest van mijn familie zag er zelfs een half uur geen' is *steek* in het tweede deel van de zin vervangen door

geïnterpreteerd kunnen worden zonder context: in de zin *hij sloeg Jan* kun je niet weten wie *hij* is zonder talige context of kennis over de werkelijkheid buiten die zin (bijvoorbeeld de spreker die, terwijl hij de zin uitspreekt, iemand van het mannelijk geslacht aanwijst; dan is die persoon waarschijnlijk de referent van *hij*).

Vanwege deze eigenschap moesten de testzinnen in deze conditie dus referenten bevatten, i.e. het oorspronkelijke element dat vervangen is door een anafoor (verwijswoord). Deze conditie heet daarom de Ana(foor)-conditie. De pronominale anaforen in de testzinnen volgden altijd op hun antecedent (de constituent waarnaar verwezen werd).

In elk van de 18 zinnen in de Ana-conditie was het complement van de V, of dat nu een DP of PP was, vervangen door een anafoor. In een derde van zinnen verwees de anafoor naar een deel van een idioom (conditie Idi-Ana), in een derde naar een deel van een collocatie (conditie Col-Ana) en in derde naar een deel van een gewone VP (Fill-Ana).

a. **Idi-Ana**

6 zinnen met een idioom,

waarin naar de direct-objects-DP (6x) verwezen wordt met een anafoor (in combinatie met de V van het VP-idioom).

*Ik zou direct de biezen pakken als ik hoorde dat mijn echtgenoot een ander had, maar hij zei dat hij die niet meteen zou pakken.*

In de Idi-Ana-conditie werd in een eerste deelzin een idioom, zoals *de biezen pakken* geïntroduceerd, waarna in een tweede deelzin het idioom werd herhaald. De tweede keer werd echter de DP ervan, in dit geval *de biezen*, vervangen door een pronominale anafoor, hier *die*. De V van het idioom, *pakken*, werd wel herhaald, omdat de anafoor naar een idioomdeél moest verwijzen, niet naar de gehele VP.

b. **Col-Ana**

6 zinnen met een collocatie,

waarin naar de direct-objects-DP (6x) verwezen wordt met een anafoor (in combinatie met de V van de VP-collocatie).

*Zij wil altijd meteen de bus pakken als het regent, maar ik pak die nooit zomaar, omdat ik veel liever fiets.*

De Col-Ana-zinnen bevatten een collocatie, zoals *de bus pakken* in het voorbeeld hierboven, waarvan de DP *de bus* in het tweede deel van de zin werd vervangen door een anafoor (*die*). Ook hier werd de V *pakken* een tweede keer genoemd.

---

(kwantitatief) *er*, wat in dit geval ook als een anafoor geldt. In 'Ik zag niet direct een uitweg en mijn reisgenoot zag er eerlijk gezegd ook geen' is *niet een uitweg* vervangen door *geen*, omdat *niet een* vervangen kan worden door *geen* en een extra anafoor *die* de tweede deelzin ongrammaticaal zou maken.



c. **Fill-Ana**

6 zinnen zonder idioom of collocatie,  
waarin naar de direct-objects-DP (6x) verwezen wordt met een anafoor (in  
combinatie met de V van de VP).

*Jonge stedelingen halen graag brood op de Voorstraat, terwijl mensen van het  
platteland dat liever halen bij de dichtstbijzijnde bakker.*

Als derde de subconditie met zinnen zonder idioom of collocatie, Fill-Ana. Hierin werd  
de DP uit een 'gewone' VP zoals *brood* in *brood halen* in de tweede deelzin vervangen  
door een pronominale anafoor, in bovenstaand voorbeeld *dat*.

#### 4.2.3 Fill-conditie

De Fill-conditie, met als enige subconditie Fill-Neutraal, diende ter controle. Deze  
conditie bestond uit zinnen die geen idiomen en collocaties bevatten, maar ook geen  
topicalisatie of anaforen. Dit moesten zinnen zijn die iedere moedertaalspreker van het  
Nederlands goed zou moeten keuren.

a. **Fill-Neutraal**

Zinnen zonder idioom of collocatie,

waarin geen sprake is van topicalisatie of anaforen (in de VP)

*Ik weet dat iedereen moet kunnen zwemmen, maar ik heb als jong kind nu eenmaal  
nooit mijn A-diploma gehaald.*

#### 4.2.4 Eisen aan het materiaal

De idiomen en collocaties in de Top-conditie en de Ana-conditie waren zoveel mogelijk  
minimale paren, te zien in de voorbeelden bij Idi-Top en Col-Top en de voorbeelden bij  
Idi-Ana en Col-Ana. Er was hier telkens sprake van een gelijke V met een verschillend  
direct object (DO), die de zin ofwel een collocatie ofwel een idioom maakte. Door de  
beperkingen van het aantal Nederlandse idiomen en collocaties konden niet alle  
testzinnen aan deze voorwaarde voldoen. In die gevallen is er wel voor een paar met  
dezelfde syntactische structuur gekozen. In het algemeen zijn er alleen idiomen en  
collocaties in de vorm DP-V en PP-V gebruikt, om de twee delen van een paar (collocatie  
en idioom) goed met elkaar te kunnen vergelijken.

In de Ana-conditie noem ik expliciet dat de anafoor (na de volledige uitspelling van het  
idioom of de collocatie in het eerste deel van de zin) in het tweede deel van de zin  
voorkomt met de V van het idioom, de collocatie ofwel de 'gewone' V uit het eerste deel  
van de zin. Ik ben me ervan bewust dat deze herhaling enigszins onnatuurlijk kan  
overkomen – in taal in een natuurlijke omgeving wordt zo min mogelijk herhaald als dat  
niet nodig is – maar deze herhaling is van belang omdat ik er zo zeker van kan zijn dat  
de anafoor geïnterpreteerd wordt als verwijzing naar de voorafgaande DO-DP van de  
idioom-VP of collocatie-VP, en niet naar de gehele VP of een geheel andere entiteit. We  
willen tenslotte testen of het mogelijk is een déél van een idioom of collocatie te  
vervangen door een anafoor. Bij de fillers is deze structuur ten behoeve van de  
gewenste gelijkenis ook aangehouden.

Alle (minimale) paren en testzinnen zijn te vinden in bijlage 1.



### 4.3 Procedure

De questionnaire bestond uit een Grammaticality Judgement Task (GJT) die digitaal werd afgenomen. De participanten kregen voor de daadwerkelijke vragenlijst met zinnen een korte instructie: er werd hun verteld dat zij moeten beoordelen hoe grammaticaal ze de aan hen voorgeschotelde zinnen vinden op een schaal van 1 t/m 7, waarin 1 staat voor absoluut niet grammaticaal en 7 voor absoluut grammaticaal.

In de instructietekst stond tevens vermeld dat er soms een idioom voorkomt in een zin. In dat geval is de betekenis van het idioom toegevoegd. Deze opzet was nodig om er zeker van te zijn dat de participanten de zin niet letterlijk opvatten, wat het oordeel over de zin kan beïnvloeden. Zo is *Een bok schiet zijn vader* allerm minst ongrammaticaal wanneer *een bok* letterlijk wordt opgevat en niet als onderdeel van een idioom. De enige manier om letterlijke lezing te voorkomen is door open kaart te spelen en de participanten in te lichten over het gebruik van idiomen.

De gehele instructietekst voor de deelnemers is te vinden in bijlage 2.

## 5. Resultaten

Van alle testzinnen (6 per conditie) zijn de gemiddelde scores en standaarddeviaties<sup>13</sup> berekend, waarbij de Top-condities (plus Fill-Neutraal) gescheiden zijn van de Ana-condities (plus Fill-Neutraal).

### 5.1 Exclusies

Na rondvraag onder de informanten bleek dat veel van hen moeite hadden met een van Fill-Ana-zinnen, die zij syntactisch anders geïnterpreteerd hadden dan in de opzet bedoeld was.<sup>14</sup> Daardoor had deze zin niet alleen een zeer lage score in vergelijking met de andere Fill-Ana-zinnen (Fill-Ana 6: 3,97; gemiddelde geïnccludeerde Fill-Ana's: 5,67) maar ook de verschillen tussen de beoordelingen waren erg groot. De zeer lage én zeer hoge beoordelingen zijn terug te zien in de hoge standaarddeviatie van Fill-Ana 6 (Sd = 1,926). Om het vertekenende effect van deze zin op de totale (conditie)score te neutraliseren is besloten om deze zin niet op te nemen in de resultaten.

Ook Fill-Ana 5 is uitgesloten van analyse vanwege een spelfout die de beoordeling van de grammaticaliteit van de zin zou kunnen hebben beïnvloed (score: 3,90 en Sd = 1,861).

### 5.2 Conditiegemiddeldes

In onderstaande tabellen zijn de resultaten opgenomen van de Topcondities (tabel 1) en de Ana-condities (tabel 2), in beide gevallen naast de resultaten voor de Fill-Neutraal-conditie.

In de meest linkse kolom van beide tabellen zijn de nummers van de testzinnen te zien en in verticale lijn de scores op de testzinnen – de minimale score is 1 en de maximale score is 7 – en tussen haakjes de standaarddeviatie. Merk op dat elke zin in Idi-Ana en Idi-Top een minimaal paar vormt met Col-Ana en Col-Top, maar dat de zinnen Fill-Ana, Fill-Top en Fill-Neutraal in willekeurige volgorde staan. Als we per paar kijken, heeft het dus alleen zin om de Idi-condities met de Ana-condities te vergelijken.

---

<sup>13</sup> De standaarddeviatie (Sd) geeft aan in hoeverre getallen afwijken van het gemiddelde aan. Hoe hoger de standaarddeviatie, hoe meer de getallen (in dit geval: de scores) verspreid zijn. Je berekent de standaarddeviatie door het gemiddelde van een reeks getallen te berekenen en van elk getal de afstand tot het gemiddelde te nemen. Neem het kwadraat van die afstanden en bereken het gemiddelde van die kwadraten. De wortel van de uitkomst daarvan is de standaarddeviatie.

<sup>14</sup> Zij hadden in deze zin (*Mijn demente opa schrijft zijn naam op al zijn servies, terwijl de andere bejaarden op zijn gang die niet op hun spullen schrijven.*) het pronomen *die* dat verwees naar *zijn naam* gelezen als een relatiefpronomen bij *de andere bejaarden op zijn gang*. Deze foutieve lezing vonden sommige participanten plausibel omdat ze al eerder in de questionnaire zinnen waren tegengekomen die op z'n minst niet helemaal grammaticaal genoemd kunnen worden, zo vermoed ik.

	<b>Idi-Top</b>	<b>Col-Top</b>	<b>Fill-Top</b>	<b>Fill-Neutraal<sup>15</sup></b>
1	4,00 (1,622)	5,18 (1,699)	6,26 (1,141)	5,41 (1,634)
2	3,69 (1,894)	4,21 (1,894)	4,08 (1,869)	6,69 (0,614)
3	3,41 (1,568)	4,72 (1,905)	5,33 (1,797)	6,51 (0,683)
4	2,90 (1,635)	5,72 (1,432)	4,87 (1,838)	6,23 (1,307)
5	2,97 (1,630)	3,31 (1,809)	5,10 (1,683)	5,41 (1,728)
6	3,31 (1,575)	5,10 (1,569)	5,28 (1,605)	6,56 (0,968)
<b>Totaal</b>	<b>3,38 (1,316)</b>	<b>4,71 (1,111)</b>	<b>5,15 (1,140)</b>	<b>6,14 (0,744)</b>

Tabel 1. De gemiddelde scores op een schaal van 1 t/m 7 voor alle soorten zinnen met topicalisatie en de neutrale filler-zinnen, met de standaarddeviaties tussen haakjes.

	<b>Idi-Ana</b>	<b>Col-Ana</b>	<b>Fill-Ana</b>	<b>Fill-Neutraal</b>
1	3,36 (1,630)	4,23 (1,769)	5,67 (1,595)	5,41 (1,634)
2	3,54 (1,790)	5,00 (1,717)	6,03 (1,112)	6,69 (0,614)
3	3,00 (1,556)	5,18 (1,502)	5,51 (1,430)	6,51 (0,683)
4	2,26 (1,251)	5,69 (1,542)	5,49 (1,537)	6,23 (1,307)
5	3,41 (2,009)	5,18 (1,684)	3,90 (1,861) (excl.)	5,41 (1,728)
6	2,92 (1,440)	4,28 (1,701)	3,97 (1,926) (excl.)	6,56 (0,968)
<b>Totaal</b>	<b>3,08 (1,153)</b>	<b>4,93 (1,041)</b>	<b>5,67 (1,00)</b>	<b>6,14 (0,744)</b>

Tabel 2. De gemiddelde scores op een schaal van 1 t/m 7 voor alle soorten zinnen met anaforen en de neutrale filler-zinnen, met de standaarddeviaties tussen haakjes.

Bij de Idi/Col-paren valt op dat (het verschil tussen) de scores per paar erg uiteenlopend zijn. Idi-Ana 4 (2,26) en Col-Ana 4 (5,69) laten bijvoorbeeld een erg groot verschil (van 3,43) zien, tegenover het veel kleinere verschil (van 0,34) tussen Idi-Top 5 (2,97) en Col-Top 5 (3,31). In alle gevallen, zowel bij de Top- als Ana-condities, scoorden de Col-zinnen hoger dan de bijbehorende Idi-zinnen.

Hoewel de scores van de (minimale) paren al interessant op zich zijn, zijn de gemiddeldes van de 6 zinnen per condities (te zien in de laatste regel van de tabellen) geschikter voor een overzichtelijke interpretatie. De algemene tendens was als volgt: de twee Idi-condities scoorden het laagst, gevolgd door beide Col-condities en daarna de Fill-condities, waarbij Fill-Neutraal het hoogst van alle condities scoorde.

Terugvertalend naar de testzinnen: deze resultaten laten zien dat de zinnen met collocaties in alle gevallen beter beoordeeld worden dan hun idiomatische equivalenten (vergelijk in tabel 1 en 2 de tweede kolom met de derde kolom). Het totale gemiddelde van de zinnen met collocaties is zowel bij topicalisatie als bij de anafoorzinnen dan ook hoger dan de zinnen met idiomen: bij topicalisatie gaat het om een gemiddelde van 3,38 (idioom) tegen een gemiddelde van 4,71 (collocatie) en bij de anafoorzinnen verschillen

<sup>15</sup> Er waren slechts 6 volledig neutrale testzinnen. De 6 testzinnen in de conditie Filler-Neutraal in tabel 1 en 2 zijn dan dezelfde, net als de bijbehorende scores.

de gemiddeldes nog meer met 3,08 (idioom) en 4,93 (collocatie).

Aangezien de fillerzinnen niet gekoppeld zijn aan specifieke zinnen met collocaties en idiomen, is het niet zinvol die per zin met elkaar te vergelijken. De gemiddeldes laten wel zien dat de fillerzinnen met topicalisatie (5,15) of anafoor (5,67) hoger beoordeeld werden dan de collocatiezinnen, maar lager dan de fillerzinnen zonder manipulaties (6,14).

### 5.3 Het vergelijken van de condities

De vraag is nu of de verschillen in tabel 1 en tabel 2 significant<sup>16</sup> zijn. Om daar antwoord op te kunnen geven heb ik een paired-samples t-test<sup>17</sup> uitgevoerd op de totale gemiddeldes van de condities (zie tabel 3, 4 en 5).

Conditie	Significant?	Waardes <sup>18</sup>
Idi-Ana & Col-Ana	Ja	$p < 0,001$ ; $t = 9,954$
Col-Ana & Fill-Ana	Ja	$p < 0,001$ ; $t = 6,525$
Idi-Ana & Fill-Ana	Ja	$p < 0,001$ ; $t = 13,117$

Tabel 3. Uitkomsten van drie paired-samples t-tests tussen de verschillende Ana-condities.

Conditie	Significant?	Waardes
Idi-Top & Col-Top	Ja	$p < 0,001$ ; $t = 8,399$
Col-Top & Fill-Top	Ja	$p < 0,001$ ; $t = 4,183$
Idi-Top & Fill-Top	Ja	$p < 0,001$ ; $t = 10,373$

Tabel 4. Uitkomsten van drie paired-samples t-tests tussen de verschillende Top-condities.

Conditie	Significant?	Waardes
Fill-Ana & Fill-Neutraal	Ja	$p = 0,002$ ; $t = 3,382$
Fill-Top & Fill-Neutraal	Ja	$p < 0,001$ ; $t = 5,386$

Tabel 5. Uitkomsten van twee paired-samples t-tests tussen de Fill-conditie en de Fill-Ana- respectievelijk Fill-Top-conditie.

Alle tests gaven significantie als uitkomst (voor alle tests geldt  $p < 0,05$ ). De grootste effecten werden aangetroffen tussen Idi-Ana en Fill-Ana ( $t = 13,117$ ) en tussen Idi-Top en Fill-Top ( $t = 10,373$ ). Ook Idi-Ana en Col-Ana ( $t = 9,954$ ) en Idi-Top en Col-Top ( $t = 8,399$ ) is een groot effect te zien.

Dit betekent dat de collocatiezinnen met een anafoor significant grammaticaler worden beoordeeld dan de idioomzinnen met anafoor. Hetzelfde geldt voor de zinnen met

---

<sup>16</sup> Ik heb de standaard p-waarde voor significantie in empirisch onderzoek aangehouden:  $p < 0,05$  betekent een significant effect.

<sup>17</sup> Een paired-samples t-test wordt in de statistiek gebruikt om een of twee groepen participanten met elkaar te vergelijken. In dit geval gaat het om een groep participanten die telkens worden vergeleken op twee variabelen, namelijk op Idi-Ana en Col-Ana, op Col-Ana en Fill-Ana, etc. Deze t-test toetst of de gemiddelde scores in relatie tot het aantal participanten en de standaarddeviatie significant van elkaar verschillen, en zo ja, in welke mate.

<sup>18</sup> De p-waarde vertelt of het gevonden verschil significant is en de t-waarde laat de grootte van het effect oftewel verschil zien; hoe groter de t-waarde, hoe groter het effect.

topicalisatie: collocaties doen het beter dan idiomen. De voorbeelden van testzinnen in (26) en (27) laten dit zien.

(26) a. Op het vinkentouw zat Tim al lange tijd, voor hij eindelijk een argument van zijn oom had gevonden dat hij onderuit kon halen.  
(idioom: score van 4,00)

b. In moeilijkheden zat Maria zeker, toen haar fraude, ondanks haar zorgvuldigheid, aan het licht was gekomen.  
(collocatie: score van 5,18)

(27) a. Ik zou direct de biezen pakken als ik hoorde dat mijn echtgenoot een ander had, maar hij zei dat hij die niet meteen zou pakken.  
(idioom: score van 3,00)

b. Zij wil altijd meteen de bus pakken als het regent, maar ik pak die nooit zomaar, omdat ik veel liever fiets.  
(collocatie: score van 5,18)

Opvallend is dat ook de Fill-Neutraal-conditie (6,14) significant hoger scoort dan de Fill-Ana-conditie (5,67;  $p = 0,002$ ) en Fill-Top-conditie (5,15;  $p < 0,001$ ). Al zijn deze effecten met  $t = 3,382$  (Fill-Neutraal en Fill-Ana) en  $t = 5,386$  (Fill-Neutraal en Fill-Top) kleiner dan de meeste andere gevonden effecten, toch is het effect onverwacht.

## 5.4 Conclusie

Samenvattend: er is sprake van een significant afnemende score tussen de condities in de volgende volgorde: Fill-Neutraal > Fill-Top > Col-Top > Idi-Top en Fill-Neutraal > Fill-Ana > Col-Ana > Idi-Ana; algemener: 'gewone' zinnen > gemanipuleerde 'gewone' zinnen > gemanipuleerde collocaties > gemanipuleerde idiomen. De zinnen met manipulaties op idiomen worden dus het slechtst beoordeeld. Een voorbeeld van de steeds hoger wordende scores in de Ana-condities is te zien in (28).

(28) a. Dat zijn vader een bok schiet bij het kerstdiner is zeker, maar Karel voorspelt dat hij die zelf niet schiet.  
(idioom: score van 3,54)

b. Of de deelnemer een garnaal pelt zoals het hoort is nog maar de vraag, maar de chefkok pelt die in ieder geval wel correct.  
(collocatie: score van 5,00)

c. Jonge stedelingen halen graag brood op de Voorstraat, terwijl mensen van het platteland dat liever halen bij de dichtstbijzijnde bakker.  
(filler: score van 6,03)

d. Ik weet dat iedereen moet kunnen zwemmen, maar ik heb als jong kind nu eenmaal nooit mijn A-diploma gehaald.  
(neutrale filler: score van 6,69)

## 6. Discussie

### 6.1 Algemeen

Uit dit onderzoek is kortgezegd gebleken dat zinnen met idiomen die deels getopicaliseerd of gepronominaliseerd zijn, significant slechter scoren in een Grammaticality Judgement Task dan zinnen met collocaties waarop dezelfde operaties waren uitgevoerd.

De hoofdhypothese wordt daarmee bevestigd: gemanipuleerde idiomen zijn inderdaad minder aanvaardbaar dan collocaties. Er kunnen dus minder makkelijk grammaticale operaties op uitgevoerd worden, zoals voorspeld aan de hand van Gazdar e.a. (1985), Wasow e.a. (1984) en (Jackendoff 1997). Zij stelden dat de mate van compositionaliteit van een idioom bepaalt in hoeverre er grammaticale operaties op uitgevoerd kunnen worden. Collocaties, die hoe dan ook compositioneel zijn, scoren dan ook beter op de geteste operaties.

De deelhypotheses I en II zijn (als inherent gevolg daarvan) bevestigd, omdat zowel topicalisatie als pronominalisatie significant aanvaardbaarder is bij collocaties dan bij idiomen. Ook deelhypothese III is bevestigd: zowel idiomen als collocaties ondergaan minder graag de twee geteste zinsmanipulaties dan 'gewone' zinnen waarin geen idioom of collocatie voorkomt.

### 6.2 De condities

#### 6.2.1 Topicalisatie

De resultaten bij de topicalisatieconditie bevestigen wat Schenk (1995) en Jackendoff (1997) al eerder lieten zien: idioomdelen kunnen over het algemeen niet getopicaliseerd worden, zie (29a). Collocaties daarentegen, doen het volgens verwachting beter, omdat ze semantisch analyseerbaar zijn (Gibbs & Nayak 1989), zie (29b)

- (29) a. Een speld zocht de buurman in een hooiberg gedurende de hele week, nadat hij meende de Staatsloterij gewonnen te hebben en zijn lot kwijt was. (Idi-Top 4: 2,90)  
b. Troost zocht de weduwe al jaren bij haar zonen, maar die woonden in Frankrijk en schonken haar na de dood van hun vader geen aandacht meer. (Col-Top 4: 5,72)  
c. Op het vinkentouw zat Tim al lange tijd, voor hij eindelijk een argument van zijn oom had gevonden dat hij onderuit kon halen. (Idi-Top 1: 4,00)

Opvallend is dat Schenk (1995) en Jackendoff (1997) (en anderen met hen) erg stellig zijn over de mogelijkheid tot topicalisatie; het kan wel of het kan niet. Mijn resultaten laten echter zien dat het niet zo zwart-wit ligt. De laagste score op een getopicaliseerd idioom is 2,90, te zien in (29a) en niet 1,00. De meningen over de mogelijkheid tot

topicalisatie zijn dus blijkbaar nogal verdeeld. Het best scorende getopicaliseerde idioom, weergegeven in (29c), haalde zelfs 4,00, wat precies het midden van de zevenpuntsschaal is.

Deze twee idiomen vergelijkende ziet men weinig verschil in compositionaliteit. Beide idiomen zijn semantisch ondoorzichtig. Het is zelfs mogelijk om (29a), de slechtst scorende, als doorzichtiger te beschouwen. Gibbs & Nayak (1989) gaven als definitie van compositionaliteit of doorzichtigheid niet slechts het kunnen toekennen van betekenis aan de losse constituenten van een idioom, maar eerder hoe *salient* (opvallend, treffend) een idioomdeel is in het gehele idioom. Als een idioomdeel semantisch te relateren is aan de figuurlijke betekenis van datzelfde deel, achten taalgebruikers een idioom eerder semantisch analyseerbaar oftewel compositioneel (Gibbs & Nayak 1989:128). Een voorbeeld: in *stoom afblazen* ('kalmeren', 'woede/frustratie uiten') bevindt *stoom* zich in hetzelfde semantische veld als *woede*, dat vaak gerelateerd wordt met innerlijke hitte en druk. Ook *stoom* is gerelateerd aan hitte en druk; taalgebruikers zullen dus weinig moeite hebben om een dergelijk idioom semantisch te ontleden omdat de constituent *stoom* makkelijk te relateren is aan de referent 'woede' (Gibbs & Nayak 1989:128).

Andere compositionele idiomen zijn minder makkelijk semantisch analyseerbaar. Neem *to spill the beans*: *spill* kan refereren aan 'onthullen' en *the beans* aan 'een geheim'. *the beans* raakt echter op geen enkele manier hetzelfde semantische veld als 'een geheim' - behalve dan een historisch-semantisch veld waaruit het idioom ontstaan is, maar dat voor de hedendaagse taalgebruiker niet meer te achterhalen is en dus buiten beschouwing gelaten moet worden. Het idioom *to spill the beans* is wel compositioneel (in traditionele zin) maar niet semantisch analyseerbaar in de definitie van Gibbs & Nayak (1989). Hun verwachting is dan ook dat dit idioom minder aanvaardbaar is bij grammaticale manipulatie dan een semantisch compositioneel idioom zoals *to let off steam*.

Gibbs & Nayak (1989) leggen hier het idiomatische veld op een andere manier bloot: niet per se compositionaliteit, maar, preciezer, semantische analyseerbaarheid speelt een belangrijke rol in de syntactische flexibiliteit. Daarnaast factoren zoals leeftijd van het idioom, de frequentie in de taalomgeving, en het bewustzijn van taalgebruikers van de herkomst van een idioom.

Terug naar (29a): een mogelijke reden voor de lage score van deze zin is de structuur van het idioom. *Een speld in een hooiberg zoeken* is geen simpel VP-idioom, maar een VP-idioom met niet alleen een DO-DP, maar ook een PP. Het is lastig te bepalen of de PP *in een hooiberg* zich in de boomstructuur bij *een speld* moet scharen of als losse bepaling bij de *V zoeken*. In de testzin is alleen de constituent *een speld* getopicaliseerd. De participanten hadden de zin wellicht aanvaardbaarder gevonden wanneer *een speld in een hooiberg* in z'n geheel was getopicaliseerd, omdat zij dat als één constituent beschouwen.

Een mogelijke reden voor de hoge score van (29b) - en andere idiomen in beide condities - is de aard van de taak die de participanten moesten doen. De instructie die zij kregen was om de zin te beoordelen op grammaticaliteit, niet op inhoud. Door de betekenis van de idiomen toe te voegen hoopte ik dat men het al dan niet behouden van



de idiomatiche betekenis mee zou nemen in het grammaticaliteitsoordeel, maar ik betwijfel of dat gelukt is. Ik vermoed dat veel participanten vooral beoordeeld hebben op wat volgens hen grammatica is: woordsoorten, woordvolgorde, congruentie, etc., en niet de betekenis van een zin. Een zin als (29b) is in letterlijke zin, hoewel semantisch opmerkelijk, niet ongrammaticaal.

Dit wijst meteen al op een belangrijke onvolkomenheid van dit onderzoek: er is niet per se getest op het behoud van de idiomatiche betekenis, al was de vooronderstelling dat participanten dat wel zouden doen. Voor vervolgonderzoek raad ik dan ook aan om, zoals Gibbs & Nayak (1989) dat deden, te vragen naar het behoud van de betekenis van idiomen in testzinnen.

Bij de topicalisatie-zinnen zonder idioom of collocatie (Fill-Top), zoals in (30) is een mogelijke verklaring voor de lagere score dat we hier te maken hebben met een sterk discourse-afhankelijke informatiestructurele notie. Topicalisatie is namelijk markering van het zinstopic door middel van vooropplaatsing (Ward & Birner 2004). Deze operatie is zonder context vaak niet gewenst en de grammaticaliteit ervan is zonder talige omgeving dan ook lastig te beoordelen. Een zin met topicalisatie mét context, zoals in (31), wordt waarschijnlijk beter beoordeeld dan (30). Dit zou kunnen verklaren waarom de testzinnen, die zonder context aan de participanten gepresenteerd werden, in de conditie Fill-Top lager beoordeeld werden dan in de conditie Fill-Neutraal.

- (30) Het hele brood had Gabriël met liefde willen opeten, maar alle starende mensen hielden hem tegen. (Fill-Top 2: 4,08)
- (31) *Op de picknicktafel lagen een heel brood en wat cherrytomaatjes die overduidelijk al even over datum waren.* Het hele brood had Gabriël met liefde willen opeten, maar alle starende mensen hielden hem tegen.

### 6.2.2 Pronominalisatie

De resultaten van de zinnen met anaforische pronomina laten zien wat Gibbs & Nayak (1989) al concludeerden over (semantisch niet-analyseerbare) idiomen, zoals in (32): die zijn niet aanvaardbaar als ze deels gepronominaliseerd zijn. Dit komt omdat de antecedenten in geen geval semantisch herkenbare eenheden zijn. Collocaties, die semantisch wel analyseerbaar zijn, kunnen pronominalisatie inderdaad beter ondergaan, zie (33).

- (32) Ik zag de eerste tien minuten geen steek, maar de rest van mijn familie zag er zelfs een half uur geen. (Idi-Top 4: 2,26)
- (33) Ik zag niet direct een uitweg en mijn reisgenoot zag er eerlijk gezegd ook geen. (Col-Top 4: 5,69)

Zoals in hoofdstuk 4 al genoemd is, werden alle zinnen met pronominale anaforen aangeboden met herhaling van de V van het idioom, zoals in (34).

- (34) Ik zou direct de biezen pakken als ik hoorde dat mijn echtgenoot een ander had, maar hij zei dat hij die niet meteen zou pakken.



Deze herhaling kan onnatuurlijk overkomen, maar is van belang omdat de anafoor geïnterpreteerd wordt als verwijzing naar de voorafgaande DO-DP (zoals in (34) *de biezen*) van de idioom-VP of collocatie-VP, en niet naar de gehele VP (*de biezen pakken*) of een geheel andere entiteit. We willen tenslotte testen of het mogelijk is een déél van een idioom of collocatie te vervangen door een anafoor. Er is echter ook een manier om dit te testen zónder de V te herhalen, namelijk door een pronominale toevoeging ‘van [persoonlijk pronomenen]’. Een zin zoals in (35) zou dan ook een betere testzin zijn, maar niet alle idiomen zijn geschikt voor deze operatie, zie (36).

- (35) Ik zou direct de biezen pakken als ik hoorde dat mijn echtgenoot een ander had, maar hij niet meteen die van hem.
- (36) a. Dat zijn vader een bok schiet bij het kerstdiner is zeker, maar Karel voorspelt dat hij die zelf niet schiet.  
b. Zijn vader schiet zeker een bok bij het kerstdiner, maar Karel die van hem niet.

*De biezen* (‘de spullen’) kunnen opgevat worden als het eigendom - bij een lichaamsdeel is ‘onderdeel’ meer van toepassing - van het subject, zoals in (33) *ik of mijn echtgenoot*. De constituent kan daarvoor vervangen worden door ‘van [persoonlijk pronomenen]’, maar bij *een bok* (‘een fout, een blunder’), die geen (letterlijke) bezitter heeft, is dat niet mogelijk. Niet alle idiomen kunnen dus op deze manier aan participanten voorgelegd worden.

In vervolgonderzoek lijkt het me daarom om bovengenoemde reden beter om met andere grammaticale manipulaties dan pronominalisatie, zoals topicalisatie of andere soorten verplaatsing, te werken.

### 6.2.3 Fillers

Uit de Fill-Neutraal-conditie blijkt dat ook ‘gewone’ zinnen, dus zonder idioom of collocatie, niet optimaal beoordeeld werden wanneer er een van de twee grammaticale operaties op toegepast is, zie (37a) met de laagste score en (37b) met de hoogste score binnen de conditie.

- (37) a. Marie vertelde dat jij haar een boek probeerde te geven, maar dat ze weigerde omdat ze zich schuldig voelde. (Fill-Neutraal 1: 5,41)  
b. Ik weet dat iedereen moet kunnen zwemmen, maar ik heb als jong kind nu eenmaal nooit mijn A-diploma gehaald. (Fill-Neutraal 2: 6,69)

Aangezien zinnen met deze manipulaties mogelijk zijn binnen het Nederlands, is dit een onverwachte uitkomst. Er zijn verschillende verklaringen te bedenken voor deze resultaten. Bij de fillerzinnen met anaforen is een plausibele reden voor de lage score in de Grammaticality Judgement Task dat deze zinnen door de herhaling van de V als onnatuurlijk overkomen. Vanwege praktische beperkingen, genoemd in paragraaf 6.2.2 acht ik het echter niet mogelijk deze herhaling te vermijden. Wel lijkt het me verstandig om bij soortgelijk onderzoek ook zinnen mét idiomen (en/of collocaties) op te nemen die géén manipulaties bevatten. Het moet natuurlijk niet zo zijn dat participanten zinnen met idiomen an sich als slechter beoordelen. Dit valt niet te controleren in de huidige opzet, waarin alleen fillers zonder idiomatische uitdrukkingen gebruikt zijn.

### 6.3 Het onderscheid tussen idiomen en collocaties

Een ander punt van aandacht is het theoretische uitgangspunt (Cruse 2000) dat ik gebruikt heb om een onderscheid te maken tussen idiomen en collocaties en de selectie die ik daarop gebaseerd heb. In de meeste gevallen is het duidelijk dat er sprake is van een groot verschil tussen de paren van idioom en collocatie die ik gebruikt heb. Zo is *een zandwinkeltje opzetten* duidelijk een ondoorzichtige expressie die niet de optelsom is van de losse delen *een zandwinkeltje* en *opzetten* en is *de tent opzetten* in ieder geval een stuk doorzichtiger. Cruse (2000) gaat er vanuit dat dat komt omdat *de tent opzetten* hier misschien niet encodeerbaar is, maar in ieder geval wel decodeerbaar, terwijl *een zandwinkeltje opzetten* geen van beiden is.

Deze definitie levert echter problemen op bij idiomen met een (enigszins) doorzichtige betekenis. Het idioom *de strijdbijl begraven* lijkt namelijk wél decodeerbaar te zijn; al is dit idioom geen letterlijke functie van de betekenis van de delen *de strijdbijl* en *begraven*, toch kan een hoorder uitstekend in staat zijn om de betekenis te achterhalen als hij de betekenis van de losse delen kent. Dit is het resultaat van de herleidbaarheid van de metafoor achter het idioom (Gibbs & Nayak 1989, Jackendoff 1997), dat veel makkelijker te doorgronden is dan een idioom zoals *op het vinkentouw zitten*, dat natuurlijk een historisch-metaforische oorsprong heeft, maar die is voor de hedendaagse hoorder allang niet meer kenbaar.

De vraag is dus of het onderscheid in idiomen en collocaties wel zit in het verschil in encodeer- en decodeerbaarheid: werkt deze analyse ook voor doorzichtige /semantische analyseerbare idiomen zoals *de strijdbijl begraven*, *iemand van zijn voetstuk stoten* of *iets onder de loep nemen*? De mijns inziens doorzichtigste idiomen in dit onderzoek, *op poten zetten* (Idi-Top5), *geen steek zien* (Idi-Ana4) en *een steentje bijdragen* (Idi-Top3) scoren niet beter dan andere idiomen met gemiddeldes van respectievelijk 2,97, 2,26 en 3,41. Deze drie gevallen zijn echter niet genoeg om hier conclusies uit te trekken.

Het is echter erg lastig om in een onderzoek naar de glijdende schaal van (de)compositionaliteit waarop alle idiomen zich bevinden, omdat het per taalgebruiker verschilt waar hij een idioom plaatst op de schaal en dus in hoeverre die syntactisch flexibel is (Gibbs & Nayak 1989). Wil men als onderzoeker een groep van doorzichtige en een groep van niet-doorzichtige idiomatische uitdrukkingen maken, dan spelen ook nog veel andere factoren mee, die voor een taalgebruiker een idioom kunnen doen verschuiven op de schaal van semantische analyseerbaarheid.

Nu lijkt compositionaliteit dus een kleinere rol te spelen bij het onderscheid tussen idiomen en collocaties, omdat er ook (min of meer) compositionele idiomen zijn zoals hierboven genoemd. Is het enige verschil dan echt dat collocaties net iets compositioneler zijn dan de meest compositionele idiomen? Wat moeten we dan met collocaties zoals *de bus pakken*, die in zekere zin misschien wel ondoorzichtiger zijn dan *de biezen pakken*? *Pakken* heeft een meer figuurlijke betekenis in *de bus pakken* dan in *de biezen pakken*; en *de biezen* moet door sommige taalgebruikers juist heel makkelijk te relateren zijn aan *de spullen* die gepakt worden bij het vertrek.

Wellicht is het onderscheid te maken vanuit het perspectief van het lexicon, waar op de een of andere manier idiomen en ook collocaties in opgeslagen moeten worden, aldus Jackendoff (1997). Hij stelt voor de insertie vanuit het lexicon ook open te stellen voor elementen groter dan  $X^0$ , zoals VP-idiomen. Idiomen hebben op die manier hun figuurlijke betekenis in het lexicon en kunnen makkelijk in een zinsstructuur insertie ondergaan, met de nodige syntactische consequenties van het niet meer kunnen ondergaan van alle mogelijke grammaticale operaties (Jackendoff 1997:167-168).

De grote vraag is: behoren collocaties daar ook tot de eenheden groter dan  $X^0$  die in één keer lexicale insertie ondergaan? Door deze twee gelijk te stellen in het lexicon is het nog moeilijker om onderscheid te maken. Beter lijkt het me om te stellen dat idiomen - doorzichtig, niet-doorzichtig of daar tussenin - als geheel in het lexicon zijn opgeslagen en lexicale insertie ondergaan, wat mogelijk is in Jackendoffs (1997) voorstel. Collocaties, die toch niet echt een geheel zijn maar eerder twee woorden die vaak samen gebruikt worden, kunnen beter behandeld worden als losse woorden die op de een of andere manier bij de insertie worden 'gecheckt' op elkaars aanwezigheid.

Het idee van 'checken' is eerder geopperd voor woorden die selectionele restricties d.w.z. semantische beperkingen of eisen kennen (Kozlowski, McCoy & Vijay-Shanker 2002). Voorbeelden van woorden met selectionele restrictie zijn *drinken* en *sterven*, die alleen gecombineerd kunnen worden met iets vloeibaars respectievelijk iets levends. De zinnen in (38) zijn, hoewel grammaticaal correct, semantisch gezien nonsens.

- (38) a. De kat drinkt het huis.  
b. De stoeprand sterft.

Kozlowski e.a. (2002) stellen dat de eisen die woorden zoals *drinken* en *sterven* stellen, op de een of andere manier in het lexicon vast moet liggen en vanuit de syntaxis 'raadpleegbaar' moeten zijn. Eenzelfde idee voor de opslag van collocaties lijkt bruikbaar: in het lexicon moet staan aangemerkt dat *een bus* gecombineerd kan worden met *pakken*.

Een bezwaar hiertegen is dan echter meteen duidelijk: in een collocatie stelt het ene woord niet per se harde eisen aan het andere woord. *Een bus* kan ook voorkomen in combinatie met *nemen*, *besturen*, *kapen*, etc. Het gaat bij collocaties niet om eisen aan het bijbehorende woord, maar om een zeer sterke neiging van twee woorden om met elkaar voor te komen.

Het lijkt me ondanks de bovengenoemde bezwaren wel vruchtbaar om collocaties als iets structureel anders te beschouwen dan idiomen. Of collocaties en idiomen nu semantisch doorzichtig zijn of niet, het staat vast dat collocaties altijd compositioneel zijn en idiomen lang niet altijd. Idiomen drukken altijd één primitieve, onderliggende betekenis uit (of kunnen in ieder geval geparafraseerd worden met een dergelijke primitieve betekenis), terwijl de twee (of meer) delen van een collocatie altijd bijdragen aan de gehele betekenis. Daarom stel ik voor meer onderzoek te doen naar syntactische beperkingen op beide soorten idiomatische expressies om zo dichter bij een theorie over de opslag van collocaties en idiomen in het lexicon.

# Literatuur

Benson, M., Benson, E. & Ilson, R. (1986). *Lexicographic Description of English*. Amsterdam/Philadelphia: Benjamins.

Chomsky, N. (1972). *Language and Mind*. New York: Harcourt Brace Jovanovitch.

Chomsky, N. (1977). On Wh-movement. In Culicover, P.W., Wasw, T. & Akmajian, A. (Eds.), *Formal syntax* (pp. 71-132). San Francisco/London: Academic Press.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of Language. Its Nature, Origin, and Use*. New York: Praeger Publishers.

Cruse, D.A. (2000). *Meaning in Language. An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.

Emonds, J.E. (1970). *Root and structure-preserving transformations*. Proefschrift MIT.

Erman, B. & Warren, B. (2000). The idiom principle and the open choice principle. *Text* (20) 1:29-62.

Everaert, M. (1991). *The lexical representation of idioms and the morphology-syntax interface*. Ongepubliceerd manuscript. Universiteit Utrecht.

Everaert, M., Van der Linden, E., Schenk, A. & Schreuder, R. (1995). Introduction. In Everaert, M., Van der Linden, E., Schenk, A. & Schreuder, R. (Eds.), *Idioms: Structural and Psychological Perspectives* (pp. 1-13). New Jersey: Lawrence Erlbaum.

Gazdar, G., Klein, E., Pullum, G.K. & Sag, I.A. (1985). *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell.

Geeraerts, D. (1995). Specialization and Reinterpretation in Idioms. In Everaert, M., Van der Linden, E., Schenk, A. & Schreuder, R. (Eds.), *Idioms: Structural and Psychological Perspectives* (pp. 58-73). New Jersey: Lawrence Erlbaum.

Gibbs, R.W. & Nayak, N.P. (1989). Psycholinguistic Studies on the Syntactic Behavior of Idioms. *Cognitive Psychology* (21), 1:100-138.

Gregoire, N. (2009). *Untangling Multiword Expressions. A study on the representation and variation of Dutch multiword expressions*. Proefschrift UU.

Harrebomée, P.J. (1990). *Spreekwoordenboek der Nederlandsche taal* (3 delen). Hoevelaken: Verba.

Jackendoff, J. (1997). *The Architecture of the Language Faculty*. Cambridge, Massachusetts: MIT Press.

Jegerski, J. (2014). Self-paced reading. In Jegerski, J. & VanPatten, B. (Eds.), *Research Methods in Second Language Psycholinguistics* (pp. 20-49). New York: Routledge.

Koster, J. (1975). Dutch as an SOV Language. *Linguistic analysis* (1), 2:111-136.

Kozlowski, R., McCoy, K.F. & Vijay-Shanker, K. (2002). Selectional restrictions in natural language sentence generation. In *Proceedings of the 6<sup>th</sup> World Multiconference on Systemics, Cybernetics, and Informatics*.

Langlotz, A. (2006). Occasional Adnominal Idiom Modification – A Cognitive Linguistic Approach. *International Journal of English Studies* (6), 1:85-108.

Lommerse, R. & Vernooij, R. (2014). *Spreekwoorden nader verklaard*. Hillegom: 5 Publishers.

McClure, S. (2011). Modification in noncombining idioms. *Semantics & Pragmatics* (4), 7:1-7.

Millar, N. (2011). The processing of malformed formulaic language. *Applied Linguistics* (32), 2:129-148.

Riemsdijk, H. van & Williams, E. (1986). *Introduction to the Theory of Grammar*. Cambridge, Massachusetts: MIT Press.

Riemsdijk, H. van (1997). Left Dislocation. In E. Anagnostopoulou, H. van Riemsdijk & F. Zwarts (Eds.), *Materials on Left Dislocation* (pp. 13-30). Amsterdam/Philadelphia: John Benjamins.

Riemsdijk, H. van & Zwarts, F. (1997). Left Dislocation in Dutch and Status of Copying Rules. In E. Anagnostopoulou, H. van Riemsdijk & F. Zwarts (Eds.), *Materials on Left Dislocation* (pp. 13-30). Amsterdam/Philadelphia: John Benjamins.

Rizzi, L. (1997) The fine structure of the left periphery. In Haegeman, L. (Ed), *Elements of grammar: A handbook of generative syntax* (pp. 281–337). Dordrecht: Kluwer.

Schenk, A. (1995). The syntactic behavior of idioms. In Everaert, M., Van der Linden, E., Schenk, A. & Schreuder, R. (Eds.), *Idioms: Structural and Psychological Perspectives* (pp. 58-73). New Jersey: Lawrence Erlbaum.

Langacker, R.W. (2001). Topic, subject, and possessor. In Simonsen, H.G. & Endresen, R.T. (Eds), *A Cognitive Approach to the Verb. Morphological and Constructional Perspectives* (pp 11-48). Berlin: Mouton de Gruyter.

Stoett, F.A. (1923-1925). *Nederlandsche spreekwoorden, spreekwijzen, uitdrukkingen en gezegden*. Vierde druk. Zutphen: W.J. Thieme & Cie.

Svenonius, P. (2005). Extending the Extension Condition to Discontinuous Idioms. *Linguistic Variation Yearbook* (5), 1:227-263.

Vat, J. (1997). Left Dislocation, Connectedness and Reconstruction. In E. Anagnostopoulou, H. van Riemsdijk & F. Zwarts (Eds.), *Materials on Left Dislocation* (pp. 67-92). Amsterdam/Philadelphia: John Benjamins.

Ward, G. & Birner, B. (2004). Information Structure and Non-canonical Syntax. In Horn, L.R. & Ward, G. (Eds), *The Handbook of Pragmatics* (pp. 153-174). Malden, Mass: Blackwell Publishing Ltd.

Wasow, T., Sag, I.A. & Nunberg, G. (1984). Idioms: An interim report. In Hattori, S. & Inoue, K. (Eds), *Proceedings of the XIIIth International Congress of Linguists* (pp. 102-115). Tokyo: Nippon Toshi Center.

Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.

Wray, A. & Perkins, M.R. (2000). The functions of formulaic language: an integrated model. *Language & Communication* (20), 1:1-28.

# Bijlagen

# Bijlage 1: Materiaal

## TOP-CONDITIE

### Paren

Idiomen	Collocaties
Op het vinkentouw zitten	In moeilijkheden zitten
Met de pet gooien naar	Met de deur gooien
Een steentje bijdragen	De verantwoordelijkheid dragen
Een speld in een hooiberg zoeken	Troost zoeken (bij)
Op poten zetten	Onder druk zetten
Een loer draaien	Een dienst draaien

### Idi-Top

1. Op het vinkentouw zat Tim al lange tijd, voor hij eindelijk een argument van zijn oom had gevonden dat hij onderuit kon halen.
2. Met de pet gooiden ze naar de herkansing van dat lastige statistiekvak.
3. Een steentje draagt Verona bij, maar dat komt haar niet altijd even goed uit.
4. Een speld zocht de buurman in een hooiberg gedurende de hele week, nadat hij meende de Staatsloterij gewonnen te hebben en zijn lot kwijt was.
5. Op poten zette Harold zijn eigen zaak, nadat hij ontslagen was door zijn gierige baas.
6. Een loer draaide Peter zijn zus vaak genoeg in groot gezelschap, wat zij meestal niet erg waardeerde.

### Col-Top

1. In moeilijkheden zat Maria zeker, toen haar fraude, ondanks haar zorgvuldigheid, aan het licht was gekomen.
2. Met de deur gooide hij zo hard dat alle fotolijstjes van de vensterbank vielen.
3. De verantwoordelijkheid draagt Marc, of hij het nu wil of niet.
4. Troost zocht de weduwe al jaren bij haar zonen, maar die woonden in Frankrijk en schonken haar na de dood van hun vader geen aandacht meer.
5. Onder druk zette Marijke haar dochter Carlijn, toen zij al twee keer haar topografietoets niet gehaald had.
6. Een dienst draaide hij over het algemeen doordeweeks en bij uitzondering 's avonds of in het weekend.

### Fill-Top

1. Dansen in een balzaal heeft Marije altijd gewild, maar haar vriendje weerhield haar daar altijd van.
2. Het hele brood had Gabriël met liefde willen opeten, maar alle starende mensen



hielden hem tegen.

3. Opzij ga ik voor niemand, niet voor bekenden en niet voor onbekenden.

4. Spijkers heeft de timmerman alleen in het weekend bij zich, ook al werkt hij vijf tot zes dagen per week.

5. In het Griftpark maakte ik gisteren een ommetje, toen ik de hond van mijn buurman zomaar los zag lopen.

6. De erfenis van zijn oma verspeelt Johan op dinsdagen in casino's, terwijl zijn moeder dat nadrukkelijk verboden had.

## ANA-CONDITIE

### Paren

Idiomen	Collocaties
De zinnen spitsen	De oren spitsen
Een bok schieten	Een garnaal pellen
De biezen pakken	De bus pakken
Geen steek zien	Geen uitweg zien
Een zandwinkeltje opzetten	Een tent opzetten
De kroon spannen	De tafel dekken

### Idi-Ana

1. Tijdens het spelprogramma spitste Rob zijn zinnen op de hoofdprijs en zijn tegenstander die van hem op de poedelprijs.

2. Dat zijn vader een bok schiet bij het kerstdiner is zeker, maar Karel voorspelt dat hij die zelf niet schiet.

3. Ik zou direct de biezen pakken als ik hoorde dat mijn echtgenoot een ander had, maar hij zei dat hij die niet meteen zou pakken.

4. Ik zag de eerste tien minuten geen steek, maar de rest van mijn familie zag er zelfs een half uur geen.

5. De arme weduwe zette een zandwinkeltje op, nadat haar geliefde echtgenoot die een jaar eerder had opgezet.

6. De derde deelnemer van de talentenjacht spant de kroon met zijn jongleeract, zeker omdat de jongen vóór hem die absoluut niet had gespannen.

### Col-Ana

1. Toen de hoorcollegedocent begon te praten, spitste Pauline haar oren en Pieter die van hem.

2. Of de deelnemer een garnaal pelt zoals het hoort is nog maar de vraag, maar de chefkok pelt die in ieder geval wel correct.

3. Zij wil altijd meteen de bus pakken als het regent, maar ik pak die nooit zomaar, omdat ik veel liever fiets.

4. Ik zag niet direct een uitweg en mijn reisgenoot zag er eerlijk gezegd ook geen.

5. John en zijn broertjes mochten op vakantie nooit eigenhandig een tent opzetten van hun moeder, want die zette ze uit voorzorg liever zelf op.
6. Zijn vader dekt de tafel bij voorkeur niet en zijn broertje had die sowieso nog nooit gedekt.

### Fill-Ana

1. Bram doet graag de was voor zijn moeder, hoewel ze die liever zelf doet vanwege zijn gebrek aan huishoudelijke vaardigheden.
2. Jonge stedelingen halen graag brood op de Voorstraat, terwijl mensen van het platteland dat liever halen bij de dichtstbijzijnde bakker.
3. Als geograaf moet je de belangrijkste begrippen wel weten uit de sociologie en psychologie, terwijl biologen die absoluut niet hoeven te weten.
4. Als je moeder en ik het paradijs gezien hebben, dan heb jij dat zeker ook gezien.
5. Toen het rolmodel de loodzware doos probeerde op te tillen, waren de andere geneigd die ook op te tillen. (geëxcludeerd)
6. Mijn demente opa schrijft zijn naam op al zijn servies, terwijl de andere bejaarden op zijn gang die niet op hun spullen schrijven. (geëxcludeerd)

## FILL-CONDITIE

### Fill-Neutraal

1. Marie vertelde dat jij haar een boek probeerde te geven, maar dat ze weigerde omdat ze zich schuldig voelde.
2. Ik weet dat iedereen moet kunnen zwemmen, maar ik heb als jong kind nu eenmaal nooit mijn A-diploma gehaald.
3. De boer moet al jaren een nieuwe schuur bouwen, maar hij is te eigenwijs om daar een bedrijf voor in te schakelen.
4. De burgemeester had gehoord dat hij niet voor het benefietconcert was uitgenodigd vanwege zijn schandelijke optreden tijdens een lezing van een tijdje geleden.
5. Karel ging alleen naar de markt met die rare jongen, zodat hij hem daarna van zijn lunchgeld kon beroven.
6. Je kunt keramisch glas gewoon reinigen met afwasmiddel en een theedoek, maar om het extra te laten glanzen zou ik ook wat reinigingscrème gebruiken.

## Bijlage 2: Instructietekst questionnaire

### Een taalkundig onderzoek

Beste deelnemer,

Fijn dat je wilt meewerken aan dit onderzoek. Alle resultaten zullen anoniem worden verwerkt. Deze test zal ongeveer 10 minuten in beslag nemen.

Je krijgt zo een aantal zinnen te zien. Ik wil je vragen elke zin hardop voor te lezen en die vervolgens te beoordelen op een schaal van 1 t/m 7. 1 staat voor absoluut ongrammaticaal en 7 voor absoluut grammaticaal. Let op: het gaat dus niet om de inhoud van deze zinnen, maar puur om de grammaticale correctheid.

Een voorbeeld van een absoluut grammaticale zin (een 7 op de schaal):  
"Mijn broertje houdt niet van gebakken aardappels, maar mijn zusje is er dol op."

Een voorbeeld van een absoluut ongrammaticale zin (een 1 op de schaal):  
"Een verfspuit schilderde de man de gehele zolder ermee."

De testzinnen zullen in sommige gevallen idiomen (vaste uitdrukkingen) bevatten. Ik heb in die gevallen de betekenis van de idiomen toegevoegd om te voorkomen dat je de betekenis van de zin niet begrijpt.

Tot slot: dit is een intuïtieve test en er is geen sprake van goede en foute antwoorden.

Alvast hartelijk dank voor je medewerking!

Fieneke Jochemsen