

# **De invloed van buurwoorden in de context van tweetalige visuele woordherkenning**

**Ilse Angenent**

**3908836**

**16-07-2015**

**Eindwerkstuk**

**Shalom Zuckerman**

## Inhoudsopgave

Abstract	p. 3
1. Inleiding	p. 4
2. Theoretisch kader	p. 7
3. Onderzoeksvraag en hypothesen	p. 13
4. Methode	p. 14
5. Resultaten	p. 17
6. Discussie	p. 29
Bijlage 1	p. 33
Bijlage 2	p. 34
Bijlage 3	p. 35
Literatuur	p. 37

## **Abstract**

*The effect of the neighbourhood in visual word recognition is sometimes facilitatory and sometimes inhibitory. Carreiras, Perea and Grainger (1997) and Perea and Pollatsek (1998) proposed that words with higher frequent neighbours will be slower recognized when unique word identification is required in a task and that high-density words will be faster recognized when participants base their responses on summed lexical activity. I tested this hypothesis in a bilingual context with a new method, the word puzzle. I tested 24 Dutch “monolinguals” and 25 Dutch-English bilinguals with high proficiency in English. I searched for correlations between English neighbors and word recognition. Unfortunately I couldn’t find an effect, due to problems with assumptions of statistical methods, but I did find indications that the hypothesis is right in at least a monolingual context. Further research could again test the hypothesis in a bilingual context with some adjustments of the research method.*

## 1. Inleiding

### 1.1 De grote vragen binnen onderzoek naar woordherkenning

Voor veel mensen in de Westerse wereld is kunnen lezen een vanzelfsprekendheid. Soms lijkt het begrijpen van geschreven woorden zelfs automatisch te gaan. Denk hierbij aan het Stroop-effect: als ik u vraag de kleur van het woord “geel” te benoemen, dan doet u daar langer over dan als ik u vraag de kleur van een groen vlak te benoemen, omdat de betekenis van het woord “geel” voor interferentie zorgt. Om toch de kleur van het woord te kunnen benoemen, moet u de betekenis van “geel” onderdrukken en daar is concentratie voor nodig. (Stroop, 1935, p. 659). Het ophalen van de betekenis van woorden is dus een sterk geautomatiseerd proces, maar wat gebeurt er precies in de hersenen voordat we een woord begrijpen? En hoe ziet het woordenboek in ons hoofd, het mentale lexicon, eruit? Dat zijn twee grote vragen binnen het onderzoek naar visuele woordherkenning. Ik zal in deze inleiding globaal bespreken welke antwoorden op deze vragen gegeven worden en maak daarbij onderscheid tussen onderzoek met eentalige en tweetalige proefpersonen. Vervolgens beschrijf ik mijn eigen onderzoeksvraag, hoe ik deze ga beantwoorden en hoe mijn verdere scriptie is opgebouwd.

### 1.2 Woordherkenning door eentaligen

Het mentale lexicon is de plek waar alle representaties van woorden liggen opgeslagen en wordt voorgesteld als een verzameling van netwerken, waaronder het semantische, fonologische en orthografische netwerk. Dit betekent dat woorden op verschillende niveaus met elkaar verbonden zouden zijn, zoals via hun semantiek (betekenis), hun fonologie (klanken) en hun orthografie (spelling). Het herkennen van een woord gebeurt in dit model doordat er informatie over dat woord uit het mentale lexicon opgehaald wordt. Tijdens dit proces worden er meerdere mogelijke kandidaten actief die met elkaar in competitie gaan, voordat een woord daadwerkelijk herkend wordt (Dijkstra, 2005, p. 180). De kandidaten binnen het orthografische netwerk verschillen met één letter van de input, hebben een letter meer of minder of hebben twee letters omgedraaid, dus bij het lezen van het woord “boot” worden ook “poot”, “bloot” en “bot” geactiveerd en “boto” als dat een woord was geweest (Davis & Taft, 2005). Nadat er kandidaten binnen het orthografische netwerk geactiveerd zijn, worden er vervolgens ook kandidaten binnen het semantische en fonologische netwerk geactiveerd, zoals “schip” en “rood” (Yates, Locker & Simpson, 2003; Yates, Locker & Simpson, 2004). Het aantal geactiveerde buurwoorden wordt vervolgens steeds verder gereduceerd tot er één woord overblijft. Op dat moment kan de benodigde informatie opgehaald worden en is er sprake van woordherkenning (Dijkstra, 2005, p. 180).

Het meest overtuigende bewijs voor het netwerkmodel komt van priming-onderzoek. Dat is onderzoek waarbij de invloed van een voorafgaand woord of object op de identificatie van een woord wordt gemeten. Het priming-effect werd voor het eerst gevonden door Meyer en Schvaneveldt (1971). Zij ontdekten dat paren van geassocieerde woorden tijdens hun *lexical decision task* sneller herkend werden dan paren van ongeassocieerde woorden. De reactietijd was bijvoorbeeld bij het woordenpaar “bread” en “butter” sneller dan bij het woordenpaar “bread” en “doctor”. Dit fenomeen wordt tegenwoordig *semantic priming* genoemd en wordt als volgt verklaard: de *prime* (het eerste woord) activeert gerelateerde woorden binnen het semantische netwerk, waarna het verwerken van een gerelateerd *target* woord sneller gaat omdat het **al geactiveerd is** (McNamara, 2005, p. 15). Vandaar dat het priming-effect als steun voor het netwerkmodel kan worden beschouwd.

### 1.3 Woordherkenning door tweetaligen

Tot in de jaren '90 was het onderzoek naar woordherkenning vooral gericht op eentalige proefpersonen (Schwartz & Van Hell, 2012). Dit betekent dat ruim honderd jaar lang de focus lag op woordherkenning door eentaligen en dat er pas sinds kort belangstelling is voor het verschil tussen woordherkenning door eentaligen en tweetaligen. Omdat het grootste deel van de mensen op de wereld meertalig is, is onderzoek met alleen eentalige proefpersonen echter niet voldoende om woordherkenning volledig te begrijpen. Een vraag die niet beantwoord kan worden met eentalige proefpersonen, is hoe talen zich in het mentale lexicon ten opzichte van elkaar verhouden: zijn woorden per taal opgeslagen en de verschillende lexicons dus van elkaar gescheiden, of zijn alle woorden geïntegreerd in één groot lexicon? Een andere vraag die niet beantwoord kan worden met eentalige proefpersonen is hoeveel talen er actief zijn tijdens het woordherkenningsproces: alleen de taal waar het woord bij hoort, of alle talen?

Om een antwoord te vinden op deze vragen, zijn onderzoekers tweetalige proefpersonen gaan testen. Zij hebben bijvoorbeeld gemeten hoe lang het duurde om cognaten, homografen of homofonen te herkennen. Cognaten zijn woorden uit verschillende talen met eenzelfde oorsprong, maar meestal worden woorden bedoeld uit verschillende talen met zowel dezelfde vorm als betekenis. Een voorbeeld is het woord “film” dat zowel in het Engels als in het Nederlands voorkomt en in beide talen dezelfde betekenis heeft. Homografen en homofonen zijn woorden die in twee of meer talen dezelfde respectievelijk orthografische of fonologische vorm hebben, maar een andere betekenis, zoals het Nederlandse werkwoord “bellen” dat in het Duits “blaffen” betekent. Dit is toevallig een voorbeeld van zowel een homograaf als een homofoon.

Als er bij het verwerken van cognaten, homografen of homofonen meerdere representaties van een woord actief worden, zal dit invloed hebben op de snelheid waarmee

het woord herkend wordt. Een verschil tussen de snelheid waarmee deze speciale woorden en “gewone” woorden herkend worden, duidt dus op co-activatie van representaties in meerdere talen. Omdat het overgrote deel van het onderzoek heeft bevestigd dat het herkennen van cognaten sneller gaat dan het herkennen van niet-cognaten, en dat homografen en homofonen afhankelijk van het soort experiment langzamer of sneller herkend worden, lijkt het erop dat er inderdaad meerdere representaties van deze woorden actief zijn. Vandaar dat deze uitkomsten beschouwd worden als bewijs voor het altijd actief zijn van alle talen tijdens de woordherkenning. (Schwartz & Van Hell, 2012, pp. 31-132).

Tegen het onderzoek met homografen, homofonen en de meeste cognaten is echter het argument in te brengen dat een woord dat in twee talen voorkomt vanzelfsprekend beide talen activeert en dat het daarom niet bewijst dat beide talen altijd actief zijn. Overtuigender bewijs zal daarom gevonden moeten worden zonder expliciet beide talen te activeren. (Midgley, Holcomb, Van Heuven & Grainger, 2008, p. 124).

#### **1.4 Huidig onderzoek**

Ik heb mijn eigen onderzoek daarom niet gericht op cognaten, homografen of homofonen, maar ik ben op zoek gegaan naar cross-linguïstische invloed van buurwoorden. Het voordeel hiervan is dat er maar één taal geactiveerd wordt tijdens de test, waarbij er wordt gemeten of de herkenning van woorden in de ene taal afhankelijk van het aantal en de frequentie van buurwoorden in een andere taal sneller of langzamer gaat. Met behulp van een nieuwe methode, de woordzoeker, heb ik onderzocht of Engelse woorden die buurwoorden zijn van een minder frequent Nederlands *target* woord, de herkenning van dat woord beïnvloeden in Nederlands-Engels tweetaligen, dus of bijvoorbeeld het frequentere Engelse woord “hug” invloed heeft op de herkenning van het minder frequente Nederlandse “hut”. Mocht het woord “hut” langzamer en dus minder vaak herkend worden dan een vergelijkbaar woord dat geen frequentere Engelse buuren heeft, dan is dit bewijs voor invloed van het Engels op de woordherkenning van Nederlandse woorden bij Nederlandse-Engels tweetalige proefpersonen. Deze uitkomst zou suggereren dat tijdens woordherkenning altijd beide talen actief zijn en dat het dus niet mogelijk is een taal volledig uit te schakelen. Ook zou het suggereren dat woorden van verschillende talen zich bevinden in één groot lexicon, omdat het aantoont dat er verbindingen zijn tussen woorden van verschillende talen. Het belangrijkste doel van mijn onderzoek is echter niet het vinden van een effect van Engelse buurwoorden van Nederlandse woorden op de herkenning van die Nederlandse woorden, maar het vinden van een antwoord op de vraag welke buurwoorden er wanneer een effect hebben.

In het volgende hoofdstuk ga ik dieper in op de theoretische achtergrond van woordherkenning. Ik zal bespreken wat er al bekend is over variabelen die van invloed zijn op woordherkenning. Ik maak daarbij onderscheid tussen eigenschappen van stimuli

(waaronder ook eigenschappen van buurwoorden van die stimuli en van de taal van de stimuli), eigenschappen van de taak en eigenschappen van proefpersonen. Vanuit deze theoretische achtergrond kom ik uit om mijn specifieke onderzoeksvraag en hypothesen, die ik zal toelichten in het derde hoofdstuk. Vervolgens beschrijf ik de opzet van mijn onderzoek in het vierde hoofdstuk, rapporteer ik de resultaten van mijn onderzoek in hoofdstuk 5 en analyseer ik deze in hoofdstuk 6, wat tegelijkertijd ook de conclusie van mijn scriptie is.

## **2. Theoretisch kader**

Er zijn verschillende modellen gemaakt van eentalige en tweetalige woordherkenning, die opgesteld zijn op basis van resultaten van experimenteel onderzoek. Met die modellen wordt het gedrag van proefpersonen gesimuleerd met als doel de onderliggende processen van woordherkenning beter te begrijpen. Een voorbeeld van zo'n model is BIA+ (*Bilingual Interactive Activation Plus*) van Dijkstra en Van Heuven (2002). Dit is een connectionistisch model van de verwerking van taal door tweetaligen, dat uitgaat van één groot lexicon waarin het niet mogelijk is één van de talen uit te schakelen, met als gevolg dat er tijdens het proces van woordherkenning in eerste instantie woorden uit beide talen geactiveerd worden (Dijkstra & Van Heuven, 2002, p. 194). Volgens het model wordt de mate van activatie van woorden mede bepaald door het rustniveau van activatie, dat afhankelijk is van bijvoorbeeld de frequentie van het woord en de vaardigheid in de taal, en wordt het aantal kandidaten dat geactiveerd wordt, bepaald door onder andere het aantal buurwoorden van het woord en de frequentie van het woord en van zijn buurwoorden (p. 182). Heel specifiek gaan Dijkstra en Van Heuven echter niet in op de factoren die invloed hebben op de herkenning van woorden. Zij geven geen overzicht van de factoren en geven ook geen informatie over welke factor wanneer en in welke mate invloed heeft. In paragraaf 2.1 geef ik daarom een overzicht van alle bekende factoren, in volgorde van de grootte van het effect dat zij hebben. Vervolgens ga ik in paragraaf 2.2 verder in op het effect van buurwoorden, omdat het soort effect sterk afhankelijk blijkt te zijn van de taak. Ten slotte geef ik een samenvatting van het hoofdstuk.

### **2.1 Factoren van invloed op woordherkenning**

#### **2.1.1 Frequentie**

De factor met het grootste en stabielste effect op de herkenning van een woord is de frequentie van dat woord. Dit betekent dat de ervaring die iemand met een woord heeft, sterk bepaalt hoe snel dat woord herkend wordt. Dit lijkt heel vanzelfsprekend, maar nog onbekend is op welke manier de ervaring met een woord invloed heeft op de herkenning van dat woord: in modellen wordt dit soms verklaard doordat woorden die vaker voorkomen een

hoger rustniveau hebben, soms doordat de drempel voordat een woord geactiveerd wordt lager is bij frequentere woorden, soms doordat frequentere woorden hoger in de lijst van te doorzoeken woorden staan en soms doordat de verbindingen tussen de units van woorden met een hogere frequentie sterker met elkaar verbonden zijn (Dufau, Lété, Touzet, Glotin, Ziegler & Grainger, 2010, pp. 670-671). Ook kan de ervaring met woorden op meerdere manieren gemeten worden. Zo kan de frequentie waarmee woorden voorkomen in teksten (*printed frequency*) gemeten worden, of het totale aantal keren dat iemand in zijn leven in aanraking is gekomen met een woord (*cumulative frequency*). Deze laatste meting blijkt een beter schatting te geven van de ervaring die iemand heeft met een woord (Rastle, 2007, pp. 75-76).

### **2.1.3 Neighbourhood**

Een andere belangrijke factor die van invloed is op woordherkenning, is de *neighbourhood*: de buurwoorden van het te verwerken woord. Dit kan op verschillende manieren gemeten worden. Het aantal buuren kan geteld worden, er kan een som of gemiddelde berekend worden van de frequentie van de buurwoorden, en het aantal buuren met een hogere frequentie kan geteld worden. Het effect van deze variabele blijkt afhankelijk van de gekozen taak faciliterend, inhibiterend of niet significant te zijn. Dufau, L'éte, Touzet, Glotin, Ziegler en Grainger (2010) verklaren dit als volgt:

At the sublexical level, words or nonwords with many neighbours are processed more quickly because they typically have more frequent sublexical units and orthography-phonology correspondences (e.g., Ziegler & Perry, 1998) or they benefit from stronger lexical feedback (Andrews, 1997). At the lexical level, however, words with many neighbours might be processed more slowly because they suffer from lexical competition and lateral inhibition (Grainger, 1990; Grainger, O'Regan, Jacobs, & Segui, 1989).

Dit betekent dat woorden met veel buuren of buuren met een hogere frequentie in sommige gevallen sneller verwerkt worden en in sommige gevallen juist trager. Ik zal hier in paragraaf 2.2 verder over uitweiden.

### **2.1.2. Age of acquisition en concreteness**

De belangrijkste subjectief bepaalde factoren zijn volgens Brysbaert, Stevens, Deyne, Voorspoels en Storms (2014) de leeftijd waarop een woord geleerd is (*age of acquisition*) en de concreetheid van het woord (*concreteness*). De leeftijd waarop een woord geleerd is lijkt los van *printed* en *cumulative frequency* invloed te hebben op de woordherkenning (Cortese & Schock, 2013). Dit betekent dat de leeftijd waarop een woord geleerd is niet alleen mede bepaalt wat de cumulatieve frequentie is, maar dat ook de volgorde van het leren van woorden op zich al van invloed is op woordherkenning, doordat het bepaalt hoe het mentale



lexicon is georganiseerd. Concreetheid is een belangrijke factor omdat het invloed heeft op geheugen en op de manier van leren van een woord. Concretere woorden zijn makkelijker te onthouden, omdat ze naast een verbale code, ook een perceptuele code activeren (bijvoorbeeld een plaatje van een olifant bij het zien van dat woord) en worden vaker via een directe ervaring geleerd in plaats van via tekst en discourse (Brybaert, Stevens, De Deyne, Voorspoels & Storms, 2014, p. 81).

#### **2.1.4 Woordlengte**

Woordlengte kan onder andere gemeten worden door het aantal letters, het aantal fonemen of het aantal syllaben te tellen. New, Ferrand, Pallier en Brybaert (2006) hebben aangetoond dat zowel het aantal letters als het aantal syllaben onafhankelijk van elkaar invloed hebben op de snelheid waarmee woorden worden herkend. Lange woorden (van 8-13 letters lang) en korte woorden (van 3-5 letters lang) werden in hun experiment langzamer herkend dan woorden van gemiddelde lengte (5-8 letters lang). De langzamere herkenning van woorden van 3 tot 5 letters lang was een onverwachte uitkomst, maar is te verklaren doordat er meer woorden uit 5 tot 8 letters bestaan dan uit 3 tot 5 letters en er tijdens het lezen minder vaak wordt gefixeerd op hele korte woorden; die worden vaker overgeslagen. Het aantal syllaben van een woord bleek in lineair verband te staan met de snelheid waarmee het woord herkend werd: hoe meer syllaben, des te meer inhibitie dit opleverde. De onderzoekers hebben niet gemeten of het aantal fonemen invloed heeft op woordherkenning, omdat dit zeer sterk correleerde met het aantal letters en minder variantie verklaarde dan het aantal letters.

#### **2.1.5 Syllable Frequency**

Naast de frequentie van woorden zelf en de frequentie van hun buurwoorden, speelt in sommige talen ook de frequentie van de syllaben van het te identificeren woord een rol. Dit *syllable frequency effect* is gevonden in het Spaans, Frans en Duits: reactietijden in *lexical decision tasks* met Spaanse, Franse en Duitse woorden waren langer bij woorden met hoogfrequente syllaben dan bij woorden met laagfrequente syllaben (Barber, Vergara & Carreiras, 2004). Het effect is echter niet gevonden in het Engels (Macizo & Van Petten, 2007). Volgens Macizo en Van Petten ligt dit aan de slecht gemarkeerde syllabegrenzen in het Engels en de matige correspondentie tussen spelling en klank. Het Spaans, Frans en Duits zijn daarentegen talen met een sterke correspondentie tussen spelling en klank. Dit feit wijst erop dat ook de taal waarin getest wordt invloed heeft op woordherkenning.

#### **2.1.6 Overige variabelen**

Mogelijk speelt ook de frequentie van bigrammen een rol bij de woordherkenning. Dit zijn alle paren van elkaar opvolgende letters in een woord. Er zijn echter tegenstrijdige onderzoeksresultaten gevonden: de ene onderzoeker vond wel een effect van bigrammen, de ander niet (Duyk, Desmet, Verbeke & Brysbaert, 2004). Door de meeste onderzoekers wordt er echter wel gecontroleerd op frequentie van bigrammen en soms ook op frequentie van letters en van (bi)fonen. Via de Clearpond database, waarin eigenschappen van woorden uit verschillende talen opgezocht kunnen worden, is het mogelijk deze frequenties te berekenen. Van twee andere variabelen is het ook niet zeker of zij een rol spelen bij de woordherkenning. Brysbaert, Stevens, De Deyne, Voorspoels en Storms (2014) schrijven dat *imageability* en *familiarity* minder belangrijk zijn omdat de eerste variabele een hoge correlatie heeft met *concreteness* en teveel visueel is gericht, en omdat de tweede variabele van minimaal belang zal zijn als woordfrequentie en AoA al bekend zijn (p. 81).

### **2.1.7 Eigenschappen van proefpersonen**

Tot nu toe heb ik de invloed van de eigenschappen van stimuli besproken. Ook eigenschappen van de proefpersoon die de taak uitvoert kunnen echter van invloed zijn op de woordherkenning. Bij eentalige proefpersonen kan je dan denken aan verschillen in leesvaardigheid, leeftijd, grootte van de woordenschat en welke woorden ze recent nog gebruikt hebben. Ook is er tussen mensen enige variantie binnen de in eerdere subparagrafen genoemde variabelen: niet iedereen heeft op exact dezelfde leeftijd en in exact dezelfde volgorde eenzelfde groep woorden geleerd.

In het geval van twee- of meertalige proefpersonen wordt het aantal mogelijke verschillen tussen mensen nog veel groter en zijn er dus veel meer variabelen die de woordherkenning kunnen beïnvloeden: meertalige proefpersonen verschillen in hun vaardigheid in de talen, in de domeinen waarin zij de talen gebruiken en in de manier en leeftijd waarop zij de talen geleerd hebben (*Wei, 2000, p. 5*). Alleen het effect van vaardigheid in een taal is tot nu toe onderzocht. Uit onderzoek van Segalowitz, Segalowitz & Wood (1998) blijkt dat naarmate de vaardigheid in een taal toeneemt, de herkenning van woorden niet alleen versneld wordt, maar dat ook het proces van herkenning geautomatiseerd wordt, wat volgens hen een herstructurering van de onderliggende cognitieve componenten van woordherkenning inhoudt (p. 53). Van Hel en Dijkstra (2002) onderzochten of een zwakkere, tweede of derde taal invloed kan hebben op de herkenning van woorden uit de moedertaal. De zwakkere taal bleek inderdaad invloed te hebben op de woordherkenning in de moedertaal, maar die invloed was pas merkbaar vanaf het moment dat iemand een bepaald niveau heeft bereikt in een niet-dominante taal. Dit wijst op een verband tussen hoe vaardig iemand in een taal is en de invloed die die taal heeft op woordherkenning in andere talen. De vaardigheid in een taal lijkt dus zowel invloed te

hebben op de herkenning van woorden in de taal van die woorden als op de herkenning van woorden in andere talen. Dit suggereert ook dat het mentale lexicon is opgebouwd als één groot lexicon dat tijdens het proces van woordherkenning altijd alle talen actief zijn.

## **2.2 Het effect van buurwoorden: facilitatie of inhibitie?**

Het bijzondere van het effect van buurwoorden op de woordherkenning is dat dit afhankelijk van de taak soms facilitatie, soms inhibitie en soms geen effect oplevert. Omdat de uitkomsten per taak zo verschillend zijn, hebben De Groot, Borgwaldt, Bos en Van den Eijnden (2002) gesuggereerd dat de meeste experimenten ons slechts iets leren over woordherkenning in een specifieke taak maar niet over woordherkenning in het algemeen. Toch zijn de uitkomsten van taken niet zo verschillend van elkaar als ze in eerste instantie lijken. Carreiras, Perea en Grainger (1997) vergeleken vijf taken met elkaar:

*progressive demasking, lexical decision, lexical decision with blocking of neighborhood density, naming, en semantic categorization.* Het viel hen op dat in het geval van taken die unieke woordidentificatie vereisen om te kunnen reageren, woorden met frequentere buurwoorden langzamer herkend werden dan andere woorden. Als unieke woordidentificatie niet noodzakelijk was om te kunnen reageren, dan werden woorden met veel buurwoorden sneller herkend dan andere woorden.

Carreiras, Perea en Grainger verklaarden dit als volgt: er zijn andere processen gaande tijdens de *lexical decision task*, waarbij je alleen hoeft te weten of een gepresenteerde letterreeks in je mentale lexicon voorkomt, dan bij de *progressive demasking task*, waarbij je (in dit geval) ook informatie moet ophalen over de fonologie en articulatie, omdat de gepresenteerde letterreeks hardop voorgelezen moet worden. Tijdens de *lexical decision task* kunnen proefpersonen reageren op basis van de totale lexicale activiteit. Dit is een snelle strategie om te bepalen of een woord voorkomt in het mentale lexicon. Hoe meer orthografische buurwoorden er zijn, des te hoger de totale activatie en des te sneller een proefpersoon overtuigd is dat de gepresenteerde letterreeks een woord is. Het aantal buurwoorden heeft daarom een faciliterend effect als proefpersonen deze snelle strategie gebruiken (Carreiras, Perea & Grainger, 1997). Zowel laag als hoog frequente buurwoorden dragen bij aan de totale activatie en dus heeft het totale aantal buurwoorden effect op de woordherkenning (Perea en Pollatsek, 1998). In het geval van een taak waarbij er één woord uit het mentale lexicon geselecteerd moet worden voordat er gereageerd kan worden, zorgen buurwoorden echter juist voor inhibitie. Het is immers zo dat hoe meer woorden er in eerste instantie actief worden, des te langer het duurt voordat alle kandidaten behalve één onderdrukt zijn (Carreiras, Perea & Grainger, 1997). Vooral frequentere buurwoorden verstoren het selecteren van het doelwoord, omdat die woorden in eerste instantie actiever zijn en de herkenning van het target woord in de weg zitten (Perea en Pollatsek, 1998).

Grainger en Jacobs (1996) hebben aangetoond dat het faciliterende effect van buurwoorden tijdens de *lexical decision task* om te keren is in een inhibiterend effect als er in de instructie van de taak staat dat het geven van juiste antwoorden belangrijker is dan de snelheid waarmee vragen beantwoord worden. Proefpersonen lijken de snelle strategie in dat geval niet te gebruiken: zij zijn dan geneigd pas te reageren als ze een woord echt herkend hebben.

Deze uitkomst bevestigt nogmaals dat buurwoorden een inhibiterend effect hebben tijdens unieke woordidentificatie en een faciliterend effect als unieke woordidentificatie niet nodig is, maar er op basis van totale lexicale activiteit gereageerd wordt. Dit betekent bovendien dat er kan worden afgeleid of proefpersonen tijdens een taak gebruik maakten van het snelle gokmechanisme of niet (Carreiras, Perea & Grainger, 1997).

Het gebruik van het snelle gokmechanisme wordt door veel onderzoekers als ongewenst gezien, omdat de proefpersonen de gepresenteerde woorden dan niet echt herkennen. Perea en Pollatsek (1998) hebben daarom in de instructie van hun *lexical decision task* aangegeven dat zij het belangrijker vonden dat proefpersonen juiste antwoorden gaven dan dat zij snelle antwoorden gaven, om de kans te vergroten dat proefpersonen hun reactie te baseren op unieke woordidentificatie. Het bijzondere aan hun onderzoek is dat zij een vergelijking hebben gemaakt tussen *lexical decision* en *silent reading*, door in beide taken dezelfde stimuli te gebruiken. In de tweede taak kregen proefpersonen zinnen te lezen waarin de stimuli voorkwamen en werden hun oogbewegingen bijgehouden. Uit de *lexical decision task* bleek dat een woord langzamer herkend werd en vaker als non-woord werd gelabeld als dat woord frequentere buurwoorden had. Uit de *silent reading task* bleek dat er ook vaker teruggekeken werd naar woorden met frequentere buurwoorden. Beide resultaten wijzen in de richting van een inhibiterend effect van frequentere buurwoorden. Omdat de *silent reading task* veel lijkt op hoe mensen normaal woorden herkennen, bevestigt dit onderzoek dat inhibitie van buurwoorden het "echte" effect is.

### **2.3 Samenvatting**

Zoals uit bovenstaande paragrafen blijkt, zijn er veel factoren van invloed op de herkenning van woorden. De belangrijkste daarvan zijn de frequentie, buurwoorden, *age of acquisition*, *concreteness* en lengte van woorden. Daarnaast speelt in sommige talen de frequentie van syllaben een rol. Andere variabelen, waarvan het niet zeker is of ze een rol spelen bij woordherkenning, zijn de frequentie van bigrammen, grammen, bifonen en fonen, *imageability* en *familiarity*. Een bijzondere factor is de invloed van buurwoorden: afhankelijk van de taak, is het effect van buurwoorden op de herkenning van een target woord inhibiterend of faciliterend. Volgens Carreiras, Perea en Grainger (1997) heeft dit te maken met of er wel of

geen unieke woordidentificatie vereist is om de taak uit te kunnen voeren. Inhibitie lijkt echter het “echte” effect te zijn, omdat dit effect ook gevonden wordt in leesexperimenten. De inhibiterende dan wel faciliterende invloed van burens kan op verschillende manieren gemeten worden, maar omdat volgens Perea en Pollatsek (1998) alleen frequentere burens storend zijn voor de herkenning van een woord is het te verwachten dat de frequentie van buurwoorden of het aantal frequentere burens een betere maat is van de inhibitie op een woord dan (de frequentie van) het totale aantal buurwoorden. In het geval van facilitatie is wel de frequentie van het totale aantal burens van invloed, omdat alle woorden bijdragen aan de totale activatie en op basis daarvan een beslissing genomen wordt.

### 3. Onderzoeksvraag en hypothesen

In het model BIA+ (Dijkstra & Van heuven, 2002) wordt er geen onderscheid gemaakt tussen taken die wel of geen unieke woordidentificatie vereisen. Ook wordt er geen verband gelegd tussen het effect dat gevonden wordt - inhibitie of facilitatie - en de burens die het effect veroorzaken - alleen hoger frequente burens of alle burens; nergens wordt genoemd dat mogelijk (de frequentie van) het totale aantal burens in het ene geval een beter maat is en in het andere geval (de frequentie van) het aantal burens met een hogere frequentie dan de target: Dijkstra en Van Heuven rekenen altijd met het totale aantal burens.

In mijn onderzoek heb ik de hypothesen van Perea en Pollatsek (1998) en Carreiras, Perea en Grainger (1997) proberen te testen. Als de uitkomsten van mijn onderzoek suggeren dat de hypothesen waar zijn, dan zou dat betekenen dat het model van Dijkstra en Van Heuven (2002) verder verfijnd kan worden. Hun hypothesen zijn gebaseerd op onderzoek naar woordherkenning door eentaligen, wat het noodzakelijk maakt om te controleren of de hypothesen generaliseerbaar zijn naar woordherkenning door tweetaligen, voordat aangenomen kan worden dat het model BIA+ aangepast moet worden. Ik heb daarom op basis van de hypothesen een onderzoek opgezet binnen een tweetalige, namelijk Nederlands-Engelse, context en ik heb onderzocht of hoger frequente Engelse burens een betere voorspeller van de herkenning van Nederlandse woorden zijn als er sprake is van unieke woordidentificatie en dus inhibitie, en of alle Engelse burens een betere voorspeller als er geen sprake is van unieke woordidentificatie en dus facilitatie.

Omdat ik een nieuwe methode heb gebruikt, was het van te voren niet bekend of woorden met veel of frequente burens langzamer of juist sneller herkend zouden worden en omdat ik maar één methode gebruikt heb (en daarbinnen ook iedereen op dezelfde manier getest is), stond al wel vast dat ik maar één kant zou kunnen onderzoeken: in het geval van inhibitie van burens of inderdaad de hoger frequente Engelse burens een betere voorspeller waren van woordherkenning en in het geval van facilitatie of alle burens een betere

voorspeller waren. Omdat de woordzoeker het opschrijven van gevonden woorden vereist wat unieke woordidentificatie stimuleert, en er bovendien geen hele sterke tijdsdruk was, verwacht ik inhibitie, geen facilitatie. Dat zou betekenen dat hoger frequente buurwoorden een betere voorspelling zijn van de snelheid van herkennen. Mocht er facilitatie optreden, dan gebruikten de proefpersonen tijdens het maken van de woordzoeker het gokmechanisme en dan zouden alle buurwoorden een betere voorspeller moeten zijn. Verder verwacht ik dat alle eerder genoemde factoren, zoals woordfrequentie en *age of acquisition* een effect hebben op de woordherkenning en dat deze effecten terug te vinden zijn in de resultaten.

## 4. Methode

### Proefpersonen

De data die ik gebruik heb voor mijn onderzoek maken deel uit van een grotere dataset, verzameld onder 222 proefpersonen door Shalom Zuckerman. Ik heb daar de tweetaligen Nederlands-Engels uit geselecteerd die aangegeven hadden een hoog niveau Engels te hebben: een 8, 9 of 10 op een schaal van 0 tot 10 waarbij 10 moedertaal niveau betekende. Ik heb gekozen voor deze groep, omdat mensen die een 8 of hoger invullen dat niet zomaar zullen doen. Het overgrote deel van deze groep vulde een 9 in, wat betekent dat zij bijna perfect Engels spreken ( $N = 25$ ,  $M = 8,92$ ). Een voorwaarde was dat zij geen andere taal hadden ingevuld met een cijfer hoger dan een 5. Daarnaast heb ik ook een groep “eentaligen” geselecteerd, als controle-groep ( $N = 24$ ). Geen enkele Nederlandse student zal helemaal eentalig zijn, ook al hebben ze geen tweede taal ingevuld, maar deze proefpersonen zullen in ieder geval het Engels lang niet zo goed beheersen als de andere groep.

### Stimuli

De proefpersonen kregen elk een woordzoeker met dertig woorden daarin verstopt, die op vaste plaatsen van links naar rechts geplaatst waren in een gerasterd vierkant, waaraan vervolgens random letters werden toegevoegd om de lege vakjes op te vullen. De woorden in de puzzel zijn geselecteerd uit een lijst van 360 woorden, die op hun beurt semi-random gekozen zijn uit SUBTLEX\_NL, een database van Nederlandse woordfrequenties gebaseerd op zo'n 44 miljoen woorden van film- en televisieondertitelingen: er zijn 400 woorden uit deze database geselecteerd waarvan er vervolgens 80 zijn weggestreept omdat deze woorden niet voorkomen in het Nederlands of de frequentie niet representatief is voor het woordgebruik in het Nederlands (de database is Vlaams). Er komen in de lijst geen

vervoegde werkwoorden voor, maar wel verbuigingen zoals meervoudsvormen en verkleiningen van zelfstandige naamwoorden. De lijst met 360 woorden is 10 keer random gesorteerd en steeds in twaalf groepen van dertig woorden verdeeld waarmee twaalf puzzels gevuld zijn. Van elke puzzel zijn er twee kopieën. Er waren dus 120 verschillende puzzels en 240 puzzels in totaal, waarvan er 222 zijn ingevuld. Een voorbeeldpuzzel en de lijst met alle 360 woorden zijn te vinden in respectievelijk bijlage 1 en 2. Op de achterkant van de puzzel was een vragenlijst te vinden. Een voorbeeld is te vinden in bijlage 3.

Doordat de woorden random zijn gekozen, komen er ook cognaten (waaronder “school”), homografen (waaronder “kind”) en homofonen (waaronder “blaf”) voor in de lijst met gekozen woorden. Bovendien zijn er soms korte Engelse woorden ontstaan doordat de puzzel met random letters is opgevuld en soms ook zitten er Engelse woorden verstopt in gekozen Nederlandse woorden (“door” in “doorstoten” en “us” in “neus”). De meest ideale situatie zou een test zijn die volledig in het Nederlands is, omdat dan uitgesloten is dat het Engels actief geactiveerd wordt. De oorzaak van de gevonden cross-linguïstische invloed kan dan niet gezocht worden in de activatie van beide talen. Het is echter haast onmogelijk een volledig Nederlandse woordzoeker te maken. Er blijven weinig woorden en letters over als alle toevallig ontstane woorden eruit worden gefilterd en bovendien zie je deze woorden ook snel over het hoofd.

Waarschijnlijk is dit echter geen groot probleem. Verschillende keren is er onderzocht of *language cues* in de context van een woord invloed hebben op de herkenning van dat woord en die invloed werd in de meeste gevallen niet gevonden. Een onderzoek waarin er wel invloed werd gevonden van de taal van de context is dat van Elston-Güttler, Gunter & Kotz (2005). Zij lieten proefpersonen vooraf aan de test een korte Engelse of Duits film zien, waarna de proefpersonen steeds een zin te zien kregen en daarna een woord waarvan ze moesten beslissen of dat een bestaand woord was. De zinnen waren in het Engels en eindigden in een interlinguïstische homograaf. Vervolgens was de betekenis van de target gerelateerd aan de Duitse betekenis van de homograaf. Er werd een primingeffect gevonden van de homograaf, maar alleen bij zinnen in het eerste blok van het experiment én alleen als proefpersonen van tevoren de Duitse film gezien hadden. (Adelman, 2012, p. 135) Het lijkt er dus op dat de taal van contextwoorden pas invloed heeft op woordherkenning als proefpersonen in een *language mode* zitten: daar is meer voor nodig dan één enkel woord of één enkele zin in een taal te zien of horen, zoals bleek uit dit experiment en uit andere experimenten waarin geen effect werd gevonden. Ik verwacht daarom dat enkele Engelse woorden in een verder Nederlandse woordzoeker niet de oorzaak zijn van een mogelijk gevonden cross-linguïstisch effect: ze hebben geen effect op de herkenning van de andere woorden in de woordzoeker. Natuurlijk zullen cognaten, homografen en homofonen zelf wel apart van de “gewone” woorden geanalyseerd moeten worden, omdat deze woorden

voordeel (of eventueel nadeel) hebben van hun voorkomen in twee talen ten opzicht van andere woorden.

### **Procedure**

De puzzels zijn verzameld in klasverband, tijdens hoorcolleges. Alle proefpersonen zijn bachelorstudenten en een enkele keer docenten, van de Universiteit Utrecht. De proefpersonen kregen tussen de twee en drie minuten de tijd om de puzzel in te vullen. De uitleg over de puzzel werd mondeling gegeven. Er werd onder andere gemeld dat woorden van links naar rechts in de puzzel te vinden waren en dat ze twee à drie minuten de tijd kregen om het experiment uit te voeren. Ook werd verzocht de woorden in de puzzel niet door te strepen maar ze onderaan de puzzel op te schrijven.

### **Metten**

De woordzoeker is een geheel nieuwe methode, waarmee snel veel data verzameld kan worden. Deze methode lijkt misschien nog het meeste op de *lexical decision task*, omdat er tijdens het zoeken in de woordzoeker lettercombinaties genegeerd moeten worden die geen woorden zijn. Een groot verschil tussen deze twee methoden is echter dat er bij de *lexical decision task* reactietijden gemeten kunnen worden, terwijl er bij de woordzoeker een andere manier van meten gevonden zal moeten worden, zoals het noteren van de volgorde waarin woorden gevonden worden of het bijhouden hoe vaak woorden gevonden worden. Dit betekent dat er niet exact gemeten kan worden hoeveel sneller een woord ten opzichte van een ander woord herkend wordt, maar wel hoeveel sneller een woord relatief gezien herkend wordt (bij meting van de volgorde waarin woorden gevonden zijn) of hoe vaak een woord ten opzichte van een ander woord herkend wordt (bij meting van hoe vaak woorden gevonden zijn). Ik heb op beide manieren gemeten. In de eerste telling (de detectiescore met volgorde) heb ik steeds het eerste woord dat door een proefpersoon was gevonden 10 punten gegeven, het tweede woord 9 punten, en zo verder tot 1 punt. Als er meer woorden gevonden werden kregen die 0 punten. In de tweede telling (de detectiescore) heb ik steeds 1 punt voor elk gevonden woord gegeven en dit per woord gedeeld door het totale aantal keer dat het woord voorkomt in de puzzels. Voor beide tellingen geldt dat ik alleen woorden heb meegeteld die expres in de woordzoeker gestopt zijn, dus niet de deelwoorden van die woorden of woorden die ontstaan zijn door de random toegevoegde letters.



## 5. Resultaten

### 5.1 Problemen

De individuele frequenties van de woorden die in de puzzels voorkomen zijn erg laag. De gemiddelde frequentie van woorden in de controlegroep was nu 2,00 (SD = 1,18) en in de tweetalige groep 1,98 (SD = 1,39). Als gevolg daarvan zegt een detectiescore zonder volgorde van 100 procent weinig: een woord is dan in veel gevallen 2 van de 2 keer gevonden. Ook is er weinig differentiatie in deze scores. De scores 0, 50 en 100 zijn het meest frequent, omdat gemiddeld een woord maar twee keer voorkomt. De verdeling van de detectiescore mét volgorde ziet er beter uit, omdat er meer differentiatie is, maar ook daar is er een enorme piek bij nul, omdat woorden vaker niet dan wel gevonden zijn. De scores zijn dus niet zo betrouwbaar, hoewel het grote aantal individuele woorden dit wel enigszins goedmaakt.

Een probleem dat hiermee samenhangt is dat de scores niet normaal verdeeld zijn. Ook logtransformaties zorgen er niet voor dat de scores normaal verdeeld worden. Voor veel statistische methoden is het echter van belang dat data normaal verdeeld zijn. Ik heb daarom zoveel mogelijk mijn keuze van statistische toetsen proberen aan te passen, maar alsnog zullen de resultaten met een korreltje zout genomen moeten worden.

#### 5.2.1 Woordfrequentie, *Age of Acquisition*, *Concreteness* en woordlengte

Bij eerder onderzoek naar de invloed van buurwoorden op de herkenning van een woord zijn er voornamelijk de methoden *progressive demasking* (het zo snel mogelijk identificeren van een woord dat steeds zichtbaarder wordt), *lexical decision* (classificatie van woorden en nonwoorden) en *naming* (hardop lezen) gebruikt. De woordzoeker is een geheel nieuwe methode, dus voordat ik een antwoord kon geven op mijn onderzoeksvraag, heb ik eerst de methode moeten testen. Dit heb ik gedaan door te onderzoeken of dezelfde variabelen die bij andere taken een invloed hebben op woordherkenning, ook een rol spelen bij woordherkenning tijdens het maken van de woordzoeker.

Ik heb zes variabelen getest: woordfrequentie, age of acquisition, concreteness en aantal letters, fonemen en syllaben. Een missende variabele is de frequentie van de syllaben van een woord. Ik heb deze variabele overgeslagen omdat er nog geen handige tool bestaat om syllabefrequenties te berekenen. De invloed van de *neighborhood* komt later aan bod.

De woordfrequenties komen uit de database SUBTLEX\_NL en de daarin ontbrekende frequenties uit de database Clearpond. Ik heb deze frequenties zelf getransformeerd met een logtransformatie, zodat ze normaal verdeeld werden. Woordfrequenties zijn namelijk niet lineair verdeeld, maar er komen heel veel woorden voor met lage frequenties en heel weinig met hoge frequenties. Ik heb eerst de frequentie per miljoen woorden gedeeld door een

miljoen, daar de log10 van genomen en vervolgens de uitkomst vermenigvuldigd met 6 (alleen log10 is noodzakelijk om de data normaal verdeeld te maken). De gemiddelde woordfrequentie in groep 1 is 123,32 keer per miljoen (SD = 1037,09) en na logtransformatie -31,96 (SD = 6,51). Van groep 2 is het gemiddelde 125,95 per miljoen (SD = 1069,03) en na logtransformatie -31,89 (SD = 6,47).

De *age of acquisition* (AoA) van de woorden in de puzzels heb ik van Brysbaert, Stevens, De Deyne, Voorspoels en Storms (2014). Zij hebben een document gemaakt met daarin de gemiddelde schattingen van 30.000 woorden de leeftijden waarop de woorden geleerd worden. De gemiddelde AoA in groep 1 is 7,17 (SD = 2,51) en in groep 2 7,23 (SD = 2,50).

De variabele *concreteness* heb ik ook van Brysbaert et alii (2014) en is verzameld onder 75 studenten in Leuven. Zij kozen op een schaal van 1 tot 5 hoe concreet 6.000 van de ruim 30.000 woorden zijn, dus elk woord is 15 keer beoordeeld. De gemiddelde *concreteness* van de woorden in groep 1 is 3,49 (SD = 1,02) en in groep 2 3,52 (SD = 1,01).

Het aantal letters en fonemen kan zowel via Clearpond als via SUBTLEX opgevraagd worden. Ik heb beide databases gecombineerd, omdat niet alle woorden in de databases voorkomen. Het aantal syllaben heb ik zelf bepaald, omdat het niet via de bovengenoemde databases op te vragen is. Het gemiddelde aantal letters, fonemen en syllaben in groep 1 was respectievelijk 6,82 (SD = 2,02), 6,04 (SD = 1,89) en 2,13 (SD = 0,81) en in groep 2 6,85 (SD = 2,03), 6,05 (SD = 1,90) en 2,12 (SD = 0,82).

Omdat de scores niet normaal verdeeld zijn, heb ik ervoor gekozen met behulp van Spearman's rangcorrelatiecoëfficiënt de woordzoekermethode te testen. In tabel 1 zijn de correlaties te vinden tussen de verschillende scores en de zes variabelen. Opvallend aan de correlaties is het verschil tussen de twee groepen: de correlaties van groep 1 zijn over het algemeen een stuk hoger, behalve bij de correlatie tussen de scores en *concreteness*. Daarnaast zijn de correlaties over het algemeen zwak: de hoogste correlatie is -0,295. Woordfrequentie en *concreteness* correleren positief met de scores, dus hoe frequenter een woord of hoe concreter een woord, des te vaker en des te sneller het gevonden werd. Dit klopt met de verwachtingen. *Age of acquisition* daarentegen correleert negatief met de scores, dus woorden die op jongere leeftijd geleerd zijn, werden sneller en vaker gevonden. Ook dit klopt met de verwachtingen. Negatieve correlaties zijn ook gevonden voor het aantal letters, fonemen en syllaben: kortere woorden worden vaker gevonden. Van deze drie variabelen is het aantal syllaben in groep 1 de sterkst met de score correlerende variabele en in groep 2 is dat het aantal fonemen. Omdat in eerder onderzoek woordlengte een parabolische relatie leek te hebben met de snelheid waarmee woorden herkend worden en Spearman's rangcorrelatiecoëfficiënt uitgaat van een monotonische relatie tussen de variabelen, is het mogelijk dat de berekende correlaties lager zijn dan in werkelijkheid. In

spreidingsdiagrammen (*scatter plot*) was echter in veel gevallen een duidelijk lineair verband te zien tussen de scores en de woordlengtevariabelen en anders wel een monotonisch verband. In groep 2 is in enkele gevallen de correlatie zo zwak dat niet te zien is of het verband wel of niet monotonisch is, maar in ieder geval is het in die gevallen niet overduidelijk amonotonisch. Het laatste opvallende is dat de detectiescore zonder volgorde in de meeste gevallen sterker correleert met de variabelen dan de detectiescore met volgorde.

Tabel 1

Spearman correlaties tussen detectiescores met en zonder volgorde en zes variabelen. Groep 1 is de controlegroep en groep 2 de tweetalige groep.

	Groep 1 (N = 325)		Groep 2 (N = 305)	
	Detectiescore zonder volgorde	Detectiescore met volgorde	Detectiescore zonder volgorde	Detectiescore met volgorde
Woordfreq	.273*	.224*	.137*	.179*
AoA	-.272*	-.256*	-.138*	-.144*
Concreteness	.165*	.179*	.209*	.178*
Aantal letters	-.204*	-.171*	-.095	-.028
Aantal fonemen	-.252*	-.208*	-.118*	-.054
Aantal syllaben	-.295*	-.261*	-.088	-.034

\*correlatie is significant ( $p < 0.05$ )

### 5.2.2 Nederlandse buurwoorden

Naast de in de vorige paragraaf besproken variabelen, heb ik onderzocht of er mogelijk invloed is van buurwoorden op de scores. Ik heb buurwoorden opgesplitst in acht variabelen. Ten eerste het aantal orthografische burenen (Nr Orth burenen) en de frequentie van die burenen (Freq Orth burenen), die per woord uit de database van Clearpond zijn op te halen. Ten tweede het aantal hoger frequente orthografische burenen (Nr HF Orth burenen) en de frequentie van die burenen (Freq HF O burenen), ook op de halen uit de database van Clearpond. Ten derde het aantal fonologische burenen (Nr Fon burenen) en de frequentie van die burenen (Freq Fon burenen) en als laatste het aantal hoger frequente fonologische burenen (Nr HF Fon burenen) en de frequentie ervan (Freq HF F burenen). In tabel 2 zijn per groep de berekende correlaties te vinden tussen de acht variabelen en de scores.

Tabel 2

Spearman correlaties tussen scores en acht variabelen. Groep 1 is de controle groep en groep 2 de tweetalige groep.

	Groep 1 (N = 325)		Groep 2 (N = 305)	
	Detectiescore zonder volgorde	Detectiescore met volgorde	Detectiescore zonder volgorde	Detectiescore met volgorde
Nr Orth buren	.201*	.176*	.108	.077
Freq Orth buren	.234*	.194*	.068	.057
Nr HF Orth buren	.080	.059	.045	-.007
Freq HF O buren	.125*	.115*	.022	-.007
Nr Fon buren	.222*	.187*	.107	.080
Freq Fon buren	.210*	.165*	.077	.068
Nr HF Fon buren	.108	.096	.052	-.017
Freq HF F buren	.125*	.125*	.039	-.006

\*correlatie is significant ( $p < 0.05$ )

Opvallend in deze tabel is het verschil tussen de correlaties van de twee groepen. De correlaties zijn bij groep 1 een stuk hoger dan bij groep 2 en bovendien allemaal positief, terwijl in groep 2 hogere frequente buren een negatieve correlatie hebben met de detectiescore met volgorde. De correlaties zijn over het algemeen vrij zwak, maar in groep 2 in de meeste gevallen wel significant. Verder valt op dat correlaties met de detectiescore met volgorde altijd lager zijn dan de correlaties met de detectiescore zonder volgorde (hoewel in één geval gelijk) en dat correlaties tussen hogere frequente buren en de scores een stuk lager zijn dan de correlaties tussen alle buren en de scores.

### 5.2.3. Engelse buurwoorden

Niet alleen Nederlandse buurwoorden hebben mogelijk invloed op woordherkenning, maar ook Engelse buurwoorden, omdat zowel de controlegroep als de tweetalige groep naast Nederlandse woorden ook Engelse woorden kent. In onderstaande tabel zijn de correlaties te vinden tussen de scores en Engelse buurwoorden, verdeeld in acht variabelen zoals ook bij de Nederlandse buurwoorden.

Tabel 3

Overzicht van Spearmans rho tussen twee scores en acht variabelen. Alle bureen zijn Engelse bureen.

	Groep 1 (N = 325)		Groep 2 (N = 305)	
	Detectiescore zonder volgorde	Detectiescore met volgorde	Detectiescore zonder volgorde	Detectiescore met volgorde
Nr Orth bureen	.203*	.169*	.085	.032
Freq Orth bureen	.221*	.173*	.093	.031
Nr HF Orth bureen	.094	.037	.048	-.049
Freq HF O bureen	.110*	.041	.070	-.040
Nr Fon bureen	.230*	.225*	.032	.010
Freq Fon bureen	.234*	.229*	.043	.017
Nr HF Fon bureen	.195*	.190*	.035	-.004
Freq HF F bureen	.198*	.196*	.043	.0001

\*correlatie is significant ( $p < 0.05$ )

In groep 1 zijn de correlaties tussen fonologische bureen en de scores hoger dan die tussen orthografische bureen en de scores. Het verschil is het grootst tussen beide soorten hoger frequente bureen. Daarnaast zijn de correlaties tussen alle bureen en de scores altijd hoger dan die tussen hoger frequente bureen en de scores en zijn de correlaties met de detectiescore zonder volgorde hoger dan met de score met volgorde, hoewel dat verschil soms erg klein is. In groep 2 zijn de correlaties met de score zonder volgorde over het algemeen ook hoger dan die met de score met volgorde. De score met volgorde heeft een negatieve correlatie met orthografische hoger frequente bureen en een extreem zwakke correlatie met de fonologische bureen. Beide scores hebben hogere correlaties met orthografische bureen dan met fonologische bureen.

#### 5.2.4. Conclusie

De gevonden correlaties tussen afhankelijke en onafhankelijke variabelen zijn zwak, vooral binnen de tweetalige groep. De correlaties binnen de controlegroep zijn in tabel 1 allemaal significant en er kan dus worden aangenomen dat er correlaties zijn tussen de desbetreffende variabelen en de scores. Binnen de tweetalige groep geldt dit alleen voor woordfrequentie, *age of acquisition* en *concreteness*. De correlaties in de eerste tabel voldoen echter wel allemaal netjes aan de verwachtingen: negatieve correlaties corresponderen met variabelen waarvan verwacht wordt dat er een negatief verband is met

woordherkenning en positieve correlaties met variabelen waarvan verwacht wordt dat er een positief verband is met woordherkenning.

Omdat in de tweede tabel correlaties van beide scores met alle burens hoger zijn dan met alleen hoger frequente burens, en bovendien de correlaties met alle burens positief zijn, sluiten deze resultaten aan bij één van de hypothesen: als er sprake is van facilitatie, zijn eigenschappen van alle burens (aantal fonologische of orthografische burens en gemiddelde frequentie daarvan) een betere voorspeller van woordherkenning dan eigenschappen van hoger frequente burens. De correlaties tussen eigenschappen van alle burens en de scores zijn echter binnen de tweetalige groep niet significant en er kan dus niet worden aangenomen dat er daadwerkelijk correlaties zijn tussen deze variabelen en woordherkenning.

In de derde tabel sluiten de gevonden correlaties binnen de controlegroep weer aan bij de hypothese dat als er sprake is van facilitatie, eigenschappen van alle burens een betere voorspeller zijn van woordherkenning dan eigenschappen van hoger frequente burens. Correlaties tussen eigenschappen van alle burens en de scores zijn namelijk weer sterker dan die tussen eigenschappen van alleen hoger frequente burens en de scores en bovendien significant. Opvallend is dat er binnen de controlegroep dus correlaties zijn gevonden tussen Engelse buurwoorden en woordherkenning, terwijl de proefpersonen een minimaal niveau van Engels hadden. Binnen de tweetalige groep zijn er juist geen significante correlaties te vinden tussen Engelse buurwoorden en woordherkenning, terwijl er binnen deze groep juist een sterkere invloed van Engelse buurwoorden te verwachten is. Correlaties zijn binnen deze groep echter in het algemeen veel zwakker en er kan dus ook iets anders aan de hand zijn.

### **5.3.1 Verdere analyse van correlaties: cognaten en woorden met deelwoorden**

Het is mogelijk dat bepaalde woorden de correlaties omhoog of omlaag halen. Een voorbeeld hiervan zijn woorden die bestaan uit deelwoorden: het is goed mogelijk dat als er een woord als “hoofdweg” in een puzzel voorkomt, een deel van de proefpersonen het woord “hoofd” opschrijft. Uit de literatuur blijkt namelijk dat (niet te) korte woorden sneller gevonden worden dan lange woorden. In groep 1 is 58 procent van de woorden zonder deelwoorden minstens één keer gevonden tegen 41 procent van de woorden met deelwoorden en in groep 2 69 procent van de woorden zonder deelwoorden tegen 59 procent van de woorden met deelwoorden. Hetzelfde geldt voor cognaten. Uit de literatuur blijkt dat cognaten sneller herkend worden en dit kan de correlaties beïnvloeden. Binnen de controle groep werd 61 procent van de cognaten gevonden tegen 64 procent van de niet-cognaten en binnen de tweetalige groep 43 procent van de cognaten tegen 48 procent van de niet-cognaten. Cognaten lijken dus juist minder vaak gevonden te worden, maar het aantal cognaten was

heel klein (23 in groep 1 en 28 in groep 2) waardoor met name de percentages van gevonden cognaten kunnen afwijken van de werkelijkheid.

In tabel 4 zijn opnieuw berekende correlaties te vinden gesplitst voor woorden die wel en geen deelwoorden van vier letters of meer bevatten. Omdat de tabel zo groot is, staat deze apart op de volgende pagina. In de tabel is te zien dat binnen de controlegroep vooral de correlaties tussen Nederlandse burenen en de scores sterker sterk veranderen: alle correlaties berekend met woorden die niet uit deelwoorden van vier letters of meer bestaan zijn sterker dan de correlaties berekend over de hele groep woorden (vetgedrukt) en alle correlatie berekend over woorden die wel uit deelwoorden van vier letters of meer bestaan zijn zwakker. Binnen groep 2 worden de correlaties met de detectiescore met volgorde in de meeste gevallen sterker, hoewel vaker als berekend over de groep woorden zonder deelwoorden. Opvallend is dat alle correlaties tussen eigenschappen van Engelse buurwoorden en de detectiescore met volgorde sterker worden, zowel voor de groep woorden met als zonder deelwoorden. Correlaties tussen eigenschappen van Engelse burenen en beide scores worden altijd negatiever als berekend over woorden zonder deelwoorden en als berekend over woorden met deelwoorden worden correlaties in de meeste gevallen positiever. Doordat de negatievere correlaties vaak al negatief waren en de positievere correlaties al positief, zijn de correlaties over de gehele groep woorden berekend zo zwak.

In tabel 5 zijn opnieuw berekende correlaties te vinden waarbij cognaten uit de data verwijderd zijn. Ook in deze tabel zijn de correlaties gesplits voor woorden met en zonder deelwoorden. Deze tabel is op de pagina na tabel 4 te vinden. Het aantal cognaten in groep 1 voor de groep woorden zonder en met deelwoorden is respectievelijk 19 en 4 en in groep 2 23 en 5. Het verwijderen van de cognaten heeft als gevolg dat binnen de controlegroep met name de correlaties tussen eigenschappen van Engelse buurwoorden en de beide scores sterker worden: alle deze correlaties worden positiever en daardoor in de meeste gevallen sterker. Het lijkt er dus op dat cognaten minder voordeel hebben van meer of frequentere Engelse buurwoorden en daardoor de correlaties tussen eigenschappen van Engelse burenen en de scores afzwakken. In tegenstelling tot bij de controlegroep zijn er bij de tweetalige groep geen verbeteringen in de correlaties tussen eigenschappen van Engelse burenen en de scores nadat cognaten uit de data verwijderd zijn. Correlaties met de detectiescore met volgorde worden gemiddeld wel sterker, maar er is geen regelmatig patroon te vinden: correlaties lijken willekeurig iets sterker of iets zwakker te worden, waardoor de gemiddelde verandering klein is. Dit geldt ook voor de verandering in correlaties met de detectiescore zonder volgorde: veranderingen zijn daar gering, doordat weer de ene helft van de correlaties iets sterker wordt en de andere iets zwakker.

Tabel 4

Overzicht van Spearmans rho tussen twee scores en alle variabelen. Links steeds de correlatie als alleen woorden meegerekend worden **zonder deelwoorden** van vier letters of meer en rechts de correlatie als alleen woorden meegerekend worden **met deelwoorden** van vier letters of meer.

	Groep 1 (N = 119 / 187)				Groep 2 (N = 131 / 174)			
	Detectiescore zonder volgorde		Detectiescore met volgorde		Detectiescore zonder volgorde		Detectiescore met volgorde	
Woordfreq	<b>.301*</b>	.166*	<b>.288*</b>	.100	.129	.096	.118	<b>.191*</b>
AoA	<b>-.275*</b>	-.175*	<b>-.323*</b>	-.119	-.109	-.118	-.079	<b>-.157*</b>
Concreteness	.142	.117	<b>.231*</b>	.080	<b>.299*</b>	.119	<b>.226*</b>	.112
Aantal letters	-.166	-.027	-.140	<u>.015</u>	<u>.060</u>	<b>-.127</b>	<b>.146</b>	<b>-.028</b>
Aantal fonemen	<b>-.299*</b>	-.076	<b>-.247*</b>	-.032	<u>.009</u>	<b>-.175*</b>	<b>.110</b>	<b>-.100</b>
Aantal syllaben	-.295*	-.180*	<b>-.273*</b>	-.141	-.017	-.077	<b>.054</b>	-.032
<b>NEDERLANDSE BUREN</b>								
Nr Orth buren	<b>.237*</b>	.055	<b>.272*</b>	.007	<b>.134</b>	.035	<b>.085</b>	.005
Freq Orth buren	<b>.313*</b>	.064	<b>.294*</b>	.024	.003	.033	.011	.014
Nr HF Orth buren	<b>.100</b>	<u>-.074</u>	<b>.080</b>	<b>-.086</b>	.035	<u>-.010</u>	-.028	<b>-.066</b>
Freq HF O buren	<b>.174*</b>	<u>-.043</u>	<b>.180*</b>	-.047	.016	<b>-.033</b>	<b>-.011</b>	<b>-.070</b>
Nr Fon buren	<b>.286*</b>	.055	<b>.265*</b>	.016	<b>.132</b>	.023	<b>.107</b>	<u>-.027</u>
Freq Fon buren	<b>.248*</b>	.053	<b>.215*</b>	.014	.034	.024	<b>.082</b>	<u>-.033</u>
Nr HF Fon buren	<b>.174*</b>	<u>-.062</u>	<b>.147</b>	<u>-.059</u>	.049	<u>-.003</u>	<b>-.024</b>	<b>-.079</b>
Freq HF F buren	<b>.142</b>	<u>-.027</u>	<b>.166</b>	<u>-.022</u>	.007	<u>-.003</u>	<b>-.012</b>	<b>-.066</b>
<b>ENGELSE BUREN</b>								
Nr Orth buren	.181*	.017	.136	.013	<u>-.025</u>	<b>.146</b>	<u>-.139</u>	<b>.143</b>
Freq Orth buren	<b>.223*</b>	.025	.134	.016	<u>-.014</u>	<b>.151*</b>	<u>-.151</u>	<b>.145</b>
Nr HF Orth buren	.028	.008	<u>-.061</u>	.009	<u>-.040</u>	<b>.099</b>	<b>-.206*</b>	<b>.090</b>
Freq HF O buren	.056	.015	<u>-.065</u>	.014	.022	<b>.104</b>	<b>-.191*</b>	<b>.095</b>
Nr Fon buren	.224*	.160*	<b>.247*</b>	.098	<u>-.037</u>	<b>.044</b>	<u>-.069</u>	<b>.036</b>
Freq Fon buren	.209*	.163*	<b>.236*</b>	.102	<u>-.013</u>	<b>.047</b>	<u>-.050</u>	<b>.036</b>
Nr HF Fon buren	.183*	.097	.171*	.125	.010	<u>-.015</u>	<b>-.049</b>	<b>-.014</b>
Freq HF F buren	.182*	.097	.183*	.124	.039	<u>-.016</u>	<u>-.036</u>	<u>-.013</u>

\*correlatie is significant ( $p < 0.05$ )

**Sterkere correlatie dan in tabel 1, 2 of 3**

Teken omgedraaid vergeleken met tabel 1, 2 of 3



Tabel 5

Overzicht van Spearmans rho tussen twee scores en alle variabelen. Links steeds de correlatie als alleen woorden meegerekend worden **zonder deelwoorden** van vier letters of meer en rechts de correlatie als alleen woorden meegerekend worden **met deelwoorden** van vier letters of meer. Cognaten zijn eruit gehaald.

	Groep 1 (N = 116, 186)				Groep 2 (N = 108, 169)			
	Detectiescore zonder volgorde		Detectiescore met volgorde		Detectiescore zonder volgorde		Detectiescore met volgorde	
Woordfreq	<b>.291*</b>	.168*	<b>.289*</b>	.101	<b>.163</b>	.093	<b>.186</b>	<b>.196*</b>
AoA	-.220*	-.185*	<b>-.285*</b>	-.130	<b>-.182</b>	-.118	<b>-.156</b>	-.141
Concreteness	<b>.171</b>	.133	<b>.255*</b>	.095	<b>.325*</b>	.123	<b>.216*</b>	.099
Aantal letters	<b>-.210*</b>	-.021	-.167	.025	.000	<b>-.110</b>	<b>.094</b>	<b>-.034</b>
Aantal fonemen	<b>-.328*</b>	-.069	<b>-.274*</b>	-.021	-.039	<b>-.155*</b>	<b>.055</b>	<b>-.093</b>
Aantal syllaben	<b>-.329*</b>	-.170*	-.293*	-.128	-.002	-.062	<b>.055</b>	<b>-.044</b>
Nr Orth buren	<b>.275*</b>	.041	<b>.294*</b>	<u>-.008</u>	<b>.175</b>	.021	<b>.135</b>	<u>-.002</u>
Freq Orth buren	<b>.323*</b>	.038	<b>.295*</b>	<u>-.006</u>	.025	.014	.050	.015
Nr HF Orth buren	<b>.119</b>	<u>-.096</u>	<b>.078</b>	<u>-.110</u>	.020	<u>-.014</u>	<b>-.046</b>	<b>-.057</b>
Freq HF O buren	<b>.200*</b>	<u>-.072</u>	<b>.195*</b>	<u>-.078</u>	<u>-.006</u>	<b>-.040</b>	<b>-.010</b>	<b>-.058</b>
Nr Fon buren	<b>.339*</b>	.047	<b>.295*</b>	.007	<b>.142</b>	.026	<b>.139</b>	<u>-.015</u>
Freq Fon buren	<b>.282*</b>	.031	<b>.234*</b>	<u>-.012</u>	.035	.022	<b>.126</b>	<u>-.013</u>
Nr HF Fon buren	<b>.196*</b>	<u>-.074</u>	<b>.149</b>	<u>-.073</u>	.015	<u>-.003</u>	<b>-.039</b>	<b>-.062</b>
Freq HF F buren	<b>.164</b>	<u>-.049</u>	<b>.182</b>	<u>-.046</u>	<u>-.019</u>	<u>-.007</u>	-.003	<b>-.045</b>
Nr Orth buren	<b>.262*</b>	.013	<b>.198*</b>	.003	.038	<b>.125</b>	<u>-.083</u>	<b>.161*</b>
Freq Orth buren	<b>.283*</b>	.024	.173	.011	.002	<b>.131</b>	<u>-.139</u>	<b>.159*</b>
Nr HF Orth buren	<b>.104</b>	.025	<u>-.012</u>	.025	.010	<b>.102</b>	<b>-.217*</b>	<b>.091</b>
Freq HF O buren	<b>.134</b>	.033	<u>-.015</u>	.031	.028	<b>.107</b>	<b>-.216*</b>	<b>.096</b>
Nr Fon buren	<b>.269*</b>	.127	<b>.308*</b>	.056	<u>-.019</u>	<u>-.002</u>	<u>-.047</u>	<b>.031</b>
Freq Fon buren	<b>.249*</b>	.130	<b>.289*</b>	.060	<u>-.006</u>	.001	<u>-.036</u>	<b>.031</b>
Nr HF Fon buren	<b>.232*</b>	.046	<b>.239*</b>	.060	<u>-.015</u>	<b>-.036</b>	<b>-.058</b>	<b>.019</b>
Freq HF F buren	<b>.232*</b>	.046	<b>.248*</b>	.070	.011	<u>-.037</u>	<u>-.046</u>	<b>.020</b>

\*correlatie is significant ( $p < 0.05$ )

**Sterkere correlatie dan in tabel 1, 2 of 3**

Teken omgedraaid vergeleken met tabel 1, 2 of 3

### 5.3.2 Conclusie

Het uit de data verwijderen van woorden met deelwoorden van vier letters of meer heeft binnen de controlegroep een duidelijk positief effect op de correlaties. Binnen groep 2 is dit alleen zo bij de correlaties met de detectiescore met volgorde en bij de correlaties tussen de detectiescore zonder volgorde en eigenschappen van Engelse buurwoorden. De belangrijkste verandering binnen de correlaties tussen de Engelse buurwoorden en de scores, is het negatiever worden van correlaties berekend over woorden zonder deelwoorden en het positiever worden van correlaties berekend over woorden met deelwoorden. In de meeste gevallen worden correlaties daardoor sterker.

Nadat ook cognaten uit de data verwijderd zijn, veranderen binnen groep 1 vooral de correlaties tussen Engelse buurwoorden en de beide scores. Binnen groep 2 is er geen veranderingspatroon te vinden: de veranderingen zijn gering en lijken vrij willekeurig hoger of lager te worden.

### 5.4.1 Partiële correlaties en regressie-analyse

Simpele correlaties tussen onafhankelijke en afhankelijke variabelen houden geen rekening met mogelijk derde variabelen die de relatie beïnvloeden. Omdat veel van de onderzochte variabelen sterk met andere variabelen correleren, zoals het aantal letters met het aantal fonemen ( $r = 0.908$  in groep 1 en  $r = 0,944$  in groep 2), is het belangrijk ook partiële correlaties te berekenen of een regressie-analyse uit te voeren. Omdat bijna alle onafhankelijke variabelen wel enigszins onderling met elkaar correleren (en vaak zelfs sterk), zou een regressie-analyse de beste optie zijn. Er zijn echter vier belangrijke voorwaarden waaraan voldaan moet zijn voordat er zo'n analyse wordt uitgevoerd: de relatie tussen onafhankelijke en afhankelijke variabelen moet lineair zijn, variabelen moeten normaal verdeeld zijn, de metingen moeten betrouwbaar zijn en er moet sprake zijn van homoscedasticiteit. Als er aan één of meer van de voorwaarden niet voldaan wordt, zullen de ware relaties onderschat worden en in het geval van multiële regressie-analyse zullen daardoor variabelen die correleren met de onderschatte variabelen juist overschat worden. Hoe meer onafhankelijke variabelen er toegevoegd worden aan het model, hoe twijfelachtiger de resultaten van de analyse worden (Osborne & Waters, 2002). Omdat mijn data niet voldoen aan alle voorwaarden, heb ik er voor gekozen enkele partiële correlaties te berekenen (geen normale verdelingen, mogelijk geen betrouwbare scores en soms heteroscedasticiteit), zodat mogelijke onder- en overschattingen beperkt blijven. Bovendien heb ik er voor gekozen cognaten en woorden met deelwoorden uit de data te verwijderen, omdat beiden de correlaties beïnvloeden, in ieder geval duidelijk in de controlegroep. Het nadeel hiervan is dat de groep woorden kleiner wordt met als gevolg instabieler correlaties,

maar omdat correlaties vooral sterker worden als gevolg van het verwijderen van een deel van de woorden, heb ik dit toch gedaan.

#### **5.4.2. Woordfrequentie en *Age of acquisition***

De volgens de literatuur belangrijkste twee variabelen, woordfrequentie en *age of acquisition*, correleren sterk met elkaar ( $r = -0,722$  in groep 1 en  $r = -0,701$  in groep 2). Daarom heb zowel de partiële correlaties berekend tussen de scores en woordfrequentie en gecontroleerd voor *age of acquisition*, als tussen de scores en *age of acquisition* en gecontroleerd voor woordfrequentie. Binnen de controlegroep zijn de partiële correlaties met woordfrequentie heel veel lager dan de simpele correlaties als er gecontroleerd wordt voor *age of acquisition*: nieuwe correlaties zijn 0,019 en -0,031 voor respectievelijk de score zonder en met volgorde. De partiële correlaties met *age of acquisition* zwakken daarentegen maar heel iets af als er gecontroleerd wordt voor woordfrequentie: de nieuwe correlaties zijn -0,212 en -0,225. *Age of acquisition* lijkt dus het sterkste verband te hebben met de scores en als gevolg van de sterke negatieve correlatie met woordfrequentie lijken daardoor de correlaties tussen woordfrequentie en de scores hoger dan ze in werkelijkheid zijn als er simpele correlaties worden berekend. Binnen de tweetalige groep is er een zelfde effect, maar minder sterk: woordfrequentie neemt sterker af dan *age of acquisition*, maar het verschil is niet zo extreem als binnen de controlegroep.

#### **5.4.3 Woordlengte**

Het aantal letters, fonemen en syllaben correleert sterk met elkaar. De gemiddelde correlatie is 0,845 (SD = 0,047). Binnen de controlegroep lijkt het aantal fonemen het sterkste effect te hebben, want partiële correlaties tussen het aantal fonemen en de scores veranderen het minste en als er gecontroleerd wordt voor het aantal fonemen zwakken de partiële correlaties het meeste af. De partiële correlatie tussen het aantal letters en de scores worden zelfs positief als er gecontroleerd wordt voor het aantal fonemen, maar dit zal een gevolg zijn van overschatting van het effect, onder andere doordat de scores niet normaal verdeeld zijn. Binnen de tweetalige groep zijn er minder goede patronen te vinden: de simpele correlaties zijn erg zwak en de partiële correlaties gaan alle kanten op. Waarschijnlijk ligt de oorzaak in de sterkere samenhang tussen woordlengte en buurwoorden en beïnvloeden buurwoorden de simpele correlaties met woordlengte in groep 2 sterker of op een andere manier waardoor de correlaties zo zwak zijn.

#### **5.4.4 Woordlengte en aantal buurwoorden**

Het aantal buuren correleert sterk met het aantal letters, fonemen en syllaben (gemiddelde correlaties liggen tussen -0,62 en -0,71). Als er gecontroleerd wordt voor het aantal buuren,

heeft dit binnen de controlegroep altijd een positief effect: de negatieve woordlengtecorrelaties worden minder negatief. Binnen de tweetalige groep heeft controle voor de Nederlandse burenen ook een positief effect op de woordlengtecorrelaties, maar controle voor de Engelse burenen zorgt juist voor iets negatievere woordlengtecorrelaties. Het verschil in effect van controle op Engelse buurwoorden tussen de twee groepen kan erop wijzen dat Engels burenen in de tweetalige groep een eigen effect hebben en in de controlegroep alleen invloed lijken te hebben door de correlatie tussen Nederlandse en Engelse burenen. Ook verklaart het tegengestelde effect van Engelse en Nederlandse buurwoorden in groep 2 de lage correlaties tussen de woordlengtevariabelen en de scores. In groep 1 hebben Engelse en Nederlandse buurwoorden eenzelfde effect en zijn de correlaties tussen woordlengtevariabelen en de scores veel hoger.

#### **5.4.5 Alle burenen en hoger frequente burenen**

Voor Nederlandse burenen geldt dat na controle van hoger frequente burenen de correlaties tussen alle burenen en de scores positief zijn en na controle van alle burenen de correlaties tussen hoger frequente burenen en de scores negatief zijn. Alleen binnen de controlegroep veranderen de correlaties met alle orthografische burenen na controle voor hoger frequente burenen minder sterk dan de correlaties met hoger frequente burenen na controle voor alle burenen. In de andere gevallen zijn de veranderingen (ongeveer) even sterk. Binnen groep 1 worden de correlaties niet altijd sterker, maar binnen groep 2 wel. De simpele correlaties binnen groep 2 zijn namelijk erg zwak, mogelijk doordat eigenschappen van alle burenen en van hoger frequente burenen elkaars onderdrukkers zijn: door de correlaties tussen alle en hoger frequente burenen lijken de simpele correlaties met de scores lager dan ze in werkelijkheid zijn. Om te bepalen of alle burenen of alleen hoger frequente burenen de sterkste voorspellers zijn van de scores is regressie-analyse nodig. De verschillen in sterkte van correlaties zijn niet erg groot en bovendien correleren bijvoorbeeld het aantal orthografische burenen niet alleen met het aantal hoger frequente orthografische burenen, maar ook met de frequentie van orthografische burenen en het aantal fonologische burenen. Er zijn te veel sterke correlaties onderling tussen de variabelen die te maken hebben met buurwoorden, dus ik kan geen uitsluitel geven welke burenen de beste voorspellers zijn van woordherkenning. Wel blijkt uit de partiële correlaties dat alle burenen in totaal een positieve correlatie hebben met de scores en de hoger frequente burenen apart een negatieve correlatie.

#### **5.4.7 Conclusie**

In deze paragraaf heb ik enkele variabelen verder onder de lope genomen door partiële correlaties te berekenen. Daaruit zijn de volgende conclusies voortgekomen: *age of acquisition* lijkt een belangrijker voorspeller te zijn van woordherkenning dan

woordfrequentie; het aantal fonemen lijkt de sterkste voorspeller te zijn van woordfrequentie binnen de drie woordlengtevariabelen; controle voor Engelse buurwoorden heeft negatieve correlaties tussen woordlengte en woordherkenning tot gevolg in de tweetalige groep en positieve in de controlegroep. Dit kan een hint zijn van een per groep verschillend effect van Engelse buurwoorden op woordherkenning. Bovendien verklaart het de lagere correlaties tussen woordlengte en woordherkenning in de tweetalige groep, omdat controle voor Nederlandse buurwoorden altijd positieve correlaties tussen woordlengte en woordherkenning ten gevolg hebben. Tot slot is de laatste conclusie dat correlaties tussen eigenschappen van alle buurwoorden en woordherkenning positief zijn en tussen eigenschappen van hogerfrequente buurwoorden en woordherkenning negatief. Antwoord op de vraag welke van de twee een betere voorspeller is van woordfrequentie kan ik niet geven, omdat daar regressie-analyse voor nodig is. Kijkend naar de simpele correlaties lijken echter alle buurwoorden een betere voorspeller te zijn. Ook heb ik geen bewijs kunnen vinden voor een sterker effect van Engelse buurwoorden op woordherkenning binnen de tweetalige groep, omdat ook daar regressie-analyse voor nodig is.

## 6. Discussie

### 6.1 De woordzoeker als methode

De eerste vraag die ik moet beantwoorden is of de woordzoeker een goede methode is. Immers, als dit niet het geval is, kan ik ook geen antwoord vinden om mijn andere onderzoeksvragen. De woordzoekermethode zou een goede methode zijn als er met deze methode dezelfde effecten op woordherkenning gevonden worden als met andere methoden die woordherkenning testen. De belangrijkste van die effecten zijn woordfrequentie, *age of acquisition*, concreteness, woordlengte en buurwoorden.

Binnen de controlegroep heb ik significante correlaties gevonden tussen al deze variabelen en woordherkenning (de scores) en ook in de juiste richting en binnen de tweetalige groep tussen woordherkenning en de variabelen woordfrequentie, *age of acquisition* en *concreteness*. De correlaties in deze groep waren over het algemeen zwakker dan in de controlegroep. Nadat cognaten en woorden met deelwoorden uit de data verwijderd waren werden de correlaties sterker, vooral in de controlegroep, maar bleven het zwakke tot matige correlaties.

Vervolgens heb ik enkele partiële correlaties berekend en bleek dat *age of acquisition* onafhankelijk van woordfrequentie een sterker effect heeft op woordherkenning dan woordfrequentie onafhankelijk van *age of acquisition*. Dit resultaat is al eens eerder gevonden, bijvoorbeeld in het onderzoek van Garlock, Walley en Metsala (2001). Verder bleek dat het aantal fonemen de belangrijkste representant van woordlengte is, maar dit is

tegenstrijdig met eerder onderzoek dat ik besproken heb in het theoretisch kader. De lagere correlaties tussen woordlengte en woordherkenning binnen de tweetalige groep bleek veroorzaakt te worden door het tegengestelde effect van Engelse en Nederlandse buren op woordlengte. Een verklaring voor de lagere correlaties tussen buurwoorden en woordherkenning binnen de tweetalige groep heb ik niet kunnen vinden met behulp van partiële correlaties, maar mogelijk heeft dit ook te maken met de invloed van Engelse buren.

Omdat de belangrijkste factoren die normaal gesproken gevonden worden in onderzoek naar woordherkenning in ieder geval binnen de controlegroep gevonden zijn, ga ik er vanuit dat de woordzoeker daadwerkelijk woordherkenning onderzoekt en dus een goede methode is.

## **6.2 Invloed van (Engelse) buurwoorden**

In mijn onderzoek ben ik op zoek gegaan naar invloed van Engelse buurwoorden op woordherkenning. De simpele correlaties laten een vreemd patroon zien: binnen de controlegroep zijn er significante correlaties gevonden tussen Engelse buurwoorden en woordherkenning en binnen de tweetalige groep niet. Een regressie-analyse zou moeten uitwijzen wat de echte bijdragen van Engelse buurwoorden zijn, want ik heb met behulp van partiële correlaties geen oplossing kunnen vinden voor het vreemde resultaat. Wel is het opvallend dat nadat cognaten en woorden met deelwoorden uit de data verwijderd waren, veel van de correlaties tussen Engelse buurwoorden en de scores negatief werden binnen de tweetalige groep, in tegenstelling tot in de controlegroep. Verschillen tussen de twee groepen zijn er dus wel te vinden.

Omdat ik geen sterker effect van Engelse buren op woordherkenning binnen de tweetalige groep heb kunnen vinden, kan ik geen verdere conclusies trekken of de hypothesen van Perea en Pollatsek (1998) en Carreiras, Perea en Grainger (1997) binnen een cross-linguïstische context kloppen. Wel kan ik deze testen binnen een eentalige context, namelijk door te kijken of alle buren in totaal een faciliterend effect hebben op woordherkenning en het aantal hoger frequente buren een inhibiterend effect. Ook zou ik kunnen bekijken welke van de twee groepen buurwoorden de beste voorspeller is voor woordherkenning. Kijkend naar de simpele correlaties tussen buurwoorden en woordherkenning lijkt het totale aantal buren de beste voorspeller te zijn van woordherkenning en deze correlaties zijn bovendien positief. Ik kan ik hieruit echter niet helemaal met zekerheid concluderen dat het totale aantal buurwoorden de beste voorspeller is, omdat er teveel sterke onderlinge correlaties tussen variabelen zijn. Mocht dit we leen juiste conclusie zijn, dan betekent dat, dat proefpersonen die de woordzoeker maken gebruikmaken van een gokmechanisme in plaats van dat er sprake is van de identificatie van unieke woorden. Opvallend is dat nadat er gecontroleerd is voor hoger frequente buren de

correlaties tussen alle buren en de scores positief zijn en na controle van alle buren de correlaties tussen hoger frequente buren en de scores negatief zijn. Dit is mogelijk een aanwijzing dat de onderzochte hypothese klopt, maar uit deze partiële correlaties blijkt nog niet welke van de twee groepen buren in dit onderzoek de beste voorspeller is voor woordherkenning.

### **6.3 Vervolgonderzoek**

Er kan nog veel verbeterd worden aan dit onderzoek. Ten eerste is het aan te raden dat in vervolgonderzoek de individuele frequenties van woorden in de puzzels hoger zijn, waardoor er meer kans is dat de scores normaal verdeeld zijn. Dit betekent dat er ofwel meer puzzels ingevuld moeten worden, of er meer dezelfde woorden in de puzzels moeten voorkomen, of de puzzels meer woorden moeten bevatten. Een combinatie van de eerste en laatste methode heeft mijn voorkeur, omdat een grote hoeveelheid verschillende woorden van belang is voor het berekenen van correlaties. Verder zou het goed zijn meteen al te zorgen dat er geen cognaten in de puzzel voorkomen en het liefst ook geen samengestelde woorden. Er zou daarnaast ook meer aandacht besteed kunnen worden aan de manier waarop proefpersonen zoeken en daar op te controleren. Verder zou er met behulp van de woordzoekermethode onderzocht kunnen worden wat de rol van het aantal buren ten opzichte van de frequentie van buren is: in hoeverre meten deze variabelen hetzelfde en zijn beiden nodig binnen een model? Ook zou er onderzocht kunnen worden wat de rol van fonologische buren en woordlengte in fonemen is binnen woordherkenning: zijn deze variabelen net zo belangrijk als orthografische buren en woordlengte in letters?

### **6.4 Huidig onderzoek in een breder kader**

Mijn eigen onderzoek naar woordherkenning heeft een kleine stap gezet in de richting van meer begrip over woordherkenning, maar de hypothese die ik heb willen testen zal in vervolgonderzoek nog nader getest moeten worden: ik ben er niet achter gekomen of deze klopt en of deze van toepassing is binnen een meertalige context. Zelfs het onderzoek naar geschreven woordherkenning als geheel zorgt echter voor slechts een klein deel van het begrip van hoe mensen taal verwerken. Geschreven taal is immers maar een klein onderdeel van het totale taalgebruik van mensen, en bovendien gaat woordherkenning over het herkennen van één enkel woord, terwijl we taal altijd binnen een context van andere woorden lezen of horen. Het is belangrijk te beseffen dat onderzoek naar de herkenning van geïsoleerde woorden niet zomaar te generaliseren is naar woordherkenning tijdens het lezen of tijdens het verwerken van taal in het algemeen. Een combinatie van onderzoek naar woorden en onderzoek naar zinnen en meer is nodig om de verwerking van taal te leren begrijpen. Het voordeel van onderzoek naar woordherkenning is echter dat woorden kleine,

makkelijk te manipuleren eenheden zijn die veel interessante eigenschappen hebben. Door te begrijpen hoe één enkel woord herkend wordt, leren we uiteindelijk ook meer over taalverwerking in het algemeen, omdat woorden belangrijke basiseenheden in taal zijn. Onderzoek naar woordherkenning is dus zeker van belang en ik hoop daar iets aan bijgedragen te hebben.



# Bijlage 1: Voorbeeldpuzzel

f	a	m	h	x	u	o	p	t	a	b	n	y	z	d	x	o	g	p	j	p	r	u	d
q	u	n	q	u	f	q	p	b	w	j	k	o	a	f	r	a	n	s	e	l	e	n	t
a	x	n	i	w	v	o	g	e	l	w	f	j	v	q	a	g	l	r	p	h	r	o	g
i	b	i	n	h	v	v	k	n	a	c	h	t	z	o	e	n	f	r	n	g	v	n	b
b	a	p	p	e	l	p	t	j	g	y	f	d	v	z	u	s	j	e	z	o	c	h	w
m	q	c	y	v	e	r	g	r	o	t	i	n	g	n	j	d	o	o	l	d	y	g	r
r	i	d	d	e	r	e	n	i	s	a	h	g	a	a	n	i	k	d	h	f	h	t	z
d	h	r	o	m	r	e	k	m	f	e	v	h	b	k	e	h	s	u	k	x	d	g	o
b	r	x	y	e	g	m	j	u	e	s	s	e	n	e	u	s	y	k	f	q	v	f	u
g	q	w	b	w	n	a	k	k	o	e	k	m	r	k	h	i	o	g	u	g	p	q	x
i	y	c	z	w	e	m	l	e	s	c	v	b	f	d	o	o	r	s	t	o	t	e	n
u	o	t	c	m	r	r	p	f	p	g	m	i	e	r	k	w	e	c	w	t	p	o	z
s	b	h	a	t	q	w	v	e	e	g	k	t	k	p	p	q	n	f	c	w	a	a	f
z	o	l	d	e	r	q	e	q	y	w	s	p	u	i	t	e	n	o	t	s	x	d	f
w	s	j	a	l	o	t	y	i	d	k	u	z	d	u	r	e	n	a	n	g	o	e	q
f	t	q	p	t	f	c	m	e	e	r	e	k	e	n	e	n	f	l	a	z	u	v	x
u	w	r	a	t	e	l	e	n	g	s	e	e	o	v	e	r	z	w	e	l	g	e	n
s	f	l	i	t	i	f	p	h	y	d	r	v	e	r	v	r	o	e	g	e	n	d	o
x	f	m	e	e	w	e	r	k	e	n	d	r	v	e	r	b	e	t	e	r	e	n	t
e	s	n	y	s	o	b	p	k	e	g	e	l	p	q	l	h	a	w	f	k	o	m	l
u	i	j	k	y	y	m	o	o	r	d	e	n	d	b	t	j	k	b	n	m	y	b	w
h	g	u	z	a	a	t	x	p	m	u	b	l	o	e	m	l	t	p	d	m	z	q	c
d	x	e	w	e	l	d	o	e	n	i	c	c	n	s	c	h	u	d	d	e	n	e	z
l	y	f	v	r	m	o	c	c	j	l	o	p	f	o	k	k	e	n	j	y	s	v	j

- 1 \_\_\_\_\_
- 2 \_\_\_\_\_
- 3 \_\_\_\_\_
- 4 \_\_\_\_\_
- 5 \_\_\_\_\_
- 6 \_\_\_\_\_
- 7 \_\_\_\_\_
- 8 \_\_\_\_\_
- 9 \_\_\_\_\_
- 10 \_\_\_\_\_
- 11 \_\_\_\_\_
- 12 \_\_\_\_\_
- 13 \_\_\_\_\_
- 14 \_\_\_\_\_
- 15 \_\_\_\_\_
- 16 \_\_\_\_\_
- 17 \_\_\_\_\_
- 18 \_\_\_\_\_
- 19 \_\_\_\_\_
- 20 \_\_\_\_\_
- 21 \_\_\_\_\_
- 22 \_\_\_\_\_
- 23 \_\_\_\_\_
- 24 \_\_\_\_\_
- 25 \_\_\_\_\_
- 26 \_\_\_\_\_
- 27 \_\_\_\_\_
- 28 \_\_\_\_\_
- 29 \_\_\_\_\_
- 30 \_\_\_\_\_
- 31 \_\_\_\_\_
- 32 \_\_\_\_\_
- 33 \_\_\_\_\_

Vragenlijst:

datum: \_\_\_\_\_

tijd: \_\_\_\_\_

Naam: \_\_\_\_\_

Student nr. \_\_\_\_\_

moedertaal: \_\_\_\_\_

tweede taal: \_\_\_\_\_  
niveau (1 redelijk -10 moedertaal niveau\_\_

meerdere talen (met niveau): \_\_\_\_\_

leeftijd: \_\_\_\_\_

m/v: \_\_\_\_\_

onderwijs: MBO / HBO / WO / andere: \_\_\_\_\_

Wat is jou stemming op dit moment [1- slecht - 10: zeer goed]: \_\_\_\_\_

aantal uren sinds laatst gegeten: \_\_\_\_\_

Beroep (studierichting): \_\_\_\_\_

Hobbies (sport, andere): \_\_\_\_\_

Email\* : \_\_\_\_\_

Heb je een strategie gebruikt bij het zoeken? :  
\_\_\_\_\_

\* wordt gebruikt uitsluitend voor een mogelijk follow-up onderzoek

### Bijlage 3: lijst van woorden die in de puzzels voorkomen

betweterig	spreken	cluedo	huizen	omgeving
omlaag	afknappen	onderbroek	spoed	yoga
klam	kurk	bezaaid	snoet	fluiten
verhaal	sponsors	toehappen	leger	hoed
zand	kleren	gehoor	balk	duwen
flora	bakker	loslaten	vies	panel
deeg	navel	muur	uitspelen	wreed
later	lieverd	bestemmen	praatpaal	varen
hoogtepunt	koets	windstoot	blaf	knaagdier
oplichten	oplettend	tandpasta	punt	telg
vier	zandbak	racewagen	zeep	afranselen
geel	springen	gracieu	spelen	vogel
platteland	intocht	verzetten	staan	nachtzoen
tactvol	aquarel	cowboy	riddertijd	appel
trein	masker	salvo	platvoet	zusje
giechelen	blaffen	herrie	eeuwenoud	vergroting
rotvent	hond	worm	verjaardag	ridderen
azuur	aardolie	synagoge	drinken	gaan
keeper	rumoer	onbewaakt	hemd	neus
takelen	eerbaar	poes	toestel	koek
vertellen	knoop	rijstrook	willen	zwemles
lekken	levenslijn	zwak	heffen	doorstoten
grijnzen	wind	honger	veulen	mier
appelmoes	model	verlekkerd	heerlijk	veeg
keukenkast	cafe	hijskraan	opvreten	zolder
zwijmelen	vroom	fietstocht	hoofd	sputten
water	gehucht	gevangenis	groente	sjalot
schouder	pirouette	bladluis	onbeslist	duren
kroon	slapen	verhogen	kies	meerekenen
precisie	hoek	komedie	daarbinnen	ratelen
goud	rood	braaf	kleuren	verzwelgen
bijnaam	actrice	stoel	informatie	vervroegen
opdrinken	middelste	stijging	eten	meewerkend
ketel	modder	meehelpen	droog	verbeteren
betovering	verontrust	borstel	timmeren	kegel

moordend	regenwoud	wijzer	filiaal	schoonmaak
bloem	bestendig	reisgenoot	hoofddoek	koekje
weldoen	bloedeigen	oase	spruiten	melk
schudden	moeder	bakje	cake	sportman
opfokken	berekenen	woestenberg	uitstel	kind
onberoerd	twist	kerstman	papa	schrijven
meer	nieuwjaar	behangen	hoofdprijs	mysterie
volzin	sprookje	rijst	deur	klas
kitsch	speeldoos	hoentje	school	tellen
rustpunt	sabbelen	opkikker	inspannen	meepraten
geven	lolbroek	mama	beginnen	klaslokaal
isoleren	hoofdweg	rest	vlieger	niet
handen	waterkant	barbiepop	lijkbleek	buddy
ontbijten	lokeend	letterlijk	bewaker	klank
lafbek	piekeren	taak	naam	muis
buskaartje	strijdbijl	rotweer	pijn	rioolput
mond	lego	klimaat	ikke	geweldig
wippen	speelgoed	rijm	vrijdag	hallo
werpen	broer	klokje	accent	schoen
drieling	konijn	beertje	tekenen	scepter
kapot	duif	papje	voeten	sukkelen
predikant	koopjes	stokbrood	wreken	knieen
meisje	grammofoon	verzorging	muziekje	zwartmaken
vermeend	opperbevel	bezorgd	wuiven	bezoeker
televisie	ontredderd	openbaren	afmeting	winkeltje
karkas	zeemacht	vrijplaats	fontein	psychologe
luisteren	huilen	drie	stronk	ademloos
stoomschip	klembord	bejaarde	zwijgzaam	bezweet
gras	eitje	vlug	rustig	diploma
hardop	jongen	zwaar	reflex	vinden
vader	boom	hand	panisch	kluts
wegsmelten	strekken	plassen	poppenhoek	soms
wiel	woelen	spits	vanmiddag	spons
politieaan	vriend	botsen	kriebels	namiddag
bordspel	wang	kusje	voet	twee
openrijten	framboos	bukken	verzorgen	open
mooi	foei	koffiepot	buik	boeten

## Literatuur

Adelman, J. S. (2012). *Visual word recognition: models and methods, orthography and phonology* (Vol. 1). Psychology Press.

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta psychologica*, 150, 80-84.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of the orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of experimental psychology: learning, memory, and cognition*, 23(4), 857-871.

Cortese, M. J., & Schock, J. (2013). Imageability and age of acquisition effects in disyllabic word recognition. *The Quarterly Journal of Experimental Psychology*, 66(5), 946-972.

Davis, C. J., & Taft, M. (2005). More words in the neighborhood: Interference in lexical decision due to deletion neighbors. *Psychonomic bulletin & review*, 12(5), 904-910.

De Groot, A. M., Borgwaldt, S., Bos, M., & Van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language*, 47(1), 91-124.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5(03), 175-197.

Dijksta, T. (2005). Bilingual visual word recognition and lexical access. *Handbook of bilingualism*, 179-201.

Dufau, S., Lété, B., Touzet, C., Glotin, H., Ziegler, J. C., & Grainger, J. (2010). A developmental perspective on visual word recognition: New evidence and a self-organising model. *European Journal of Cognitive Psychology*, 22(5), 669-694.

Duyck, W., Desmet, T., Verbeke, L. P., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers*, 36(3), 488-499.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research*, 25(1), 57-70.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and language*, 45(3), 468-492.

Grainger, J., and Jacobs, A. M. (1996) Orthographic processing in visual word recognition: a multiple read-out model. *Psychological Review*, 103: 518–65.

Macizo, P., & Van Petten, C. (2007). Syllable frequency in lexical decision and naming of English words. *Reading and Writing*, 20(4), 295-331.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Psychology Press.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of experimental psychology*, 90(2), 227.

Midgley, K. J., Holcomb, P. J., Walter, J. B., & Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Research*, 1246, 123-135.

New, B., Ferrand, L., Pallier, C., Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic bulletin & review*, 13(1), 45-52.

Osborne, J., Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, 8(2). Retrieved 14 Juli 2015 van <http://PAREonline.net/getvn.asp?v=8&n=2>.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of experimental psychology: human perception and performance*, 24(3), 767-779.

Rastle, K. (2007). Visual word recognition. In Gaskell, M. G. (2007). *The Oxford handbook of psycholinguistics*. Oxford University Press (pp. 71-87).

Schwartz, A. I., & Van Hell, J. G. (2012). Bilingual visual word recognition in sentence context. In Adelman, J., S. (Ed.) (2012). *Visual word recognition*, volume 2: meaning and context, individuals and development. Hove, Engeland: Psychology Press (pp. 131-150).

Segalowitz, S. J., Segalowitz, N. S., & Wood, A. G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics*, 19(01), 53-67.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6), 643-662

Van Hell, J. G. & Dijkstra, T. (2002). Foreign Language Knowledge can Influence native language performance in exclusively native contexts. *Psychonomic Nulletin & Preview*, 9(4), 780-789.

Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39(3), 458-483.

Wei, L. (2000). Dimensions of Bilingualism. In Wei, L. (Ed.). (2000). *The bilingualism Reader*. Psychology Press (pp. 3-25).

Yates, M., Locker, L., & Simpson, G. B. (2003). Semantic and phonological influences on the processing of words and pseudohomophones. *Memory & Cognition*, 31, 856-866.

Yates, M., Locker, L., & Simpson, G. B. (2004). The influence of phonological neighborhood on visual word perception. *Psychonomic Bulletin & Review*, 11(3), 452-457.