

**Het Verschil tussen Woordassociaties van Kinderen en Volwassenen**

Marieke Maas

5909090

Datum: 16 april 2020

Begeleider: Dr. S. Zuckerman

Tweede beoordelaar: R.M. van Lambalgen

Opleiding: bachelor Kunstmatige Intelligentie

Verdiepingspakket: Reasoning

Institutie: Universiteit Utrecht

Studiepunten: 15 ECTS

Abstract

In deze scriptie zal het verschil in woordassociaties tussen kinderen en volwassenen onderzocht worden. Dit om meer inzicht te krijgen in de organisatie van het mentale lexicon. Hiervoor is een woordassociatie taak gedaan met een groep basisschool kinderen. Dit experiment had dezelfde basis als het experiment van Brown en Berko (1960). De resultaten van de kinderen werden vergeleken met de resultaten van het Small World of Words project (De Deyne, Navarro & Storms, 2013). Om een mogelijk verschil te kunnen ontdekken heb ik gekeken naar verschillende factoren die invloed zouden kunnen hebben op woordassociaties. Uit voorgaand onderzoek blijkt dat o.a. de volgende factoren invloed zouden kunnen hebben op woordassociaties: woordsoort, woordcombinatie, woord valentie waarde, Age of Acquisition, woordfrequentie en diversiteit. Deze factoren heb ik ook in mijn experiment onderzocht. Uit de resultaten is gebleken dat kinderen en volwassenen op het gebied van woordfrequentie en diversiteit van elkaar verschillen kijkend naar woordassociaties. Bij de kinderen werd een correlatie gevonden tussen de frequentie van het stimuluswoord en die van de associatie. Ook waren de kinderen meer divers in hun associaties dan de volwassenen. Op de andere factoren bleken de kinderen en de volwassenen niet te verschillen.

Voor de connectie met kunstmatige intelligentie is er een klein vergelijkingsexperiment uitgevoerd tussen een kunstmatig model dat gebruik maakt van het Word2Vec model en de resultaten van het experiment van Brown en Berko (1960). Hieruit blijkt dat mensen (kinderen en volwassenen) en één kunstmatig model qua woordassociaties niet op elkaar lijken.

Inhoudsopgave

[1. Introductie 5](#_Toc37887458)

[2. Achtergrond Literatuur 7](#_Toc37887459)

[2.1 Factoren die invloed kunnen hebben op woordassociaties 7](#_Toc37887460)

[2.1.1 Woordsoort/Woordcombinatie 7](#_Toc37887461)

[2.1.2 Woord Valentie Waarde 9](#_Toc37887462)

[2.1.3 Age of Acquisition 11](#_Toc37887463)

[2.1.4 Woordfrequentie 12](#_Toc37887464)

[2.1.5 Diversiteit 13](#_Toc37887465)

[2.2 Word2Vec Model 15](#_Toc37887466)

[2.3 Het Small of World of Words project 16](#_Toc37887467)

[2.4 Korte Samenvatting 17](#_Toc37887468)

[2.5 Onderzoeksvraag en hypothese 18](#_Toc37887469)

[3. Methode 20](#_Toc37887470)

[3.1 Participanten 20](#_Toc37887471)

[3.2 Materiaal en Taak 20](#_Toc37887472)

[3.3 Aanpassing na pilottest 21](#_Toc37887473)

[3.4 Design 21](#_Toc37887474)

[3.5 Procedure 22](#_Toc37887475)

[3.6 Metingen 22](#_Toc37887476)

[3.7 A Small World of Words project 22](#_Toc37887477)

[4. Resultaten 23](#_Toc37887478)

[4.1 Algemeen 23](#_Toc37887479)

[4.2 Woordsoort 24](#_Toc37887480)

[4.3 Woordcombinatie 26](#_Toc37887481)

[4.4 Woord valentie waarde 27](#_Toc37887482)

[4.5 Age of Acquisition 28](#_Toc37887483)

[4.6 Frequentie 30](#_Toc37887484)

[4.7 Diversiteit 31](#_Toc37887485)

[4.8 Overlap 32](#_Toc37887486)

[4.9 Korte Samenvatting 33](#_Toc37887487)

[4.10 Bijzondere bevindingen 33](#_Toc37887488)

[5. Discussie 34](#_Toc37887489)

[5.1 Conclusie 34](#_Toc37887490)

[5.2 Discussie van de resultaten 37](#_Toc37887491)

[5.3 Vergelijking met kunstmatig model 38](#_Toc37887492)

[5.4 Beperkingen en toekomstig onderzoek 40](#_Toc37887493)

[6. Referenties 42](#_Toc37887494)

[6.1 Literatuur 42](#_Toc37887495)

[6.2 Websites 44](#_Toc37887496)

[7. Appendix 45](#_Toc37887497)

[A. Age of Acquisition en Frequentie waarden van (100 willekeurig gekozen woorden uit) A Small World of Words 45](#_Toc37887498)

[A1. Age of Acquisition waarden 45](#_Toc37887499)

[A2. Frequentie waarden 47](#_Toc37887500)

[B. Informatiebrief en toestemmingsverklaring voor schoolleiding en ouders 50](#_Toc37887501)

[B1. Informatiebrief ouders 50](#_Toc37887502)

[B2. Toestemmingsverklaring ouders 53](#_Toc37887503)

[B3. Informatiebrief schoolleiding 54](#_Toc37887504)

[C. Reacties van kinderen en volwassenen 57](#_Toc37887505)

[C1. Top 3 reacties van de kinderen. 57](#_Toc37887506)

[C2. Top 3 reacties van de volwassenen (gebaseerd op het Small World of Words project) 58](#_Toc37887507)

[D. Reacties van het kunstmatige model 60](#_Toc37887508)

# 1. Introductie

Mens vs. machine, een veelbesproken onderwerp van de laatste tijd. Tegenwoordig verslaat een Artificial Intelligence (AI) de allerbeste menselijke spelers in schaken (Deep Blue), de tv-quiz Jeopardy! (Watson) en Go (AlphaGo). Ook maken statistische modellen betere keuzes dan menselijke experts. AI heeft ook een invloed op de cognitieve wetenschap. Om specifiek te zijn, AI heeft een aanzienlijke impact gehad op de voortdurende verschuiving van een orthodoxe naar embodied-embedded cognitieve wetenschap (Foese, 2007). AI heeft het mogelijk gemaakt om filosofische geschillen op een experimentele manier aan te pakken. Omgekeerd kan enactivism een sterke invloed hebben op AI vanwege het biologisch gegronde verslag van autonomous agency en sense-making. De ontwikkeling van zo’n AI zou mogelijk de problemen kunnen verhelpen die op dit moment de vooruitgang in embodied-embedded AI in de weg staan.

Een relatief nieuw onderwerp binnen de cognitieve wetenschap is het mentale lexicon. Dit is ons innerlijk woordenboek dat alle woorden bevat die iemand kent. Ook bevat het alle connecties tussen deze woorden. Uit verschillende onderzoeken is gebleken dat het mentale lexicon gerepresenteerd kan worden als een netwerk (De Deyne, Kennet, Anaki, Faust & Navarro, 2016; Kennet, Kennet, Ben-Jacon & Faust, 2011; De Deyne, Verheyen & Storms, 2016; Vitevich, Goldstein, Siew & Castro, 2014). Vanuit deze netwerkbenadering zou het mentale lexicon uit verschillende clusters van woorden bestaan. Woorden die zich binnen hetzelfde cluster bevinden zullen sneller worden geactiveerd dan woorden uit een ander cluster. Binnen een cluster zijn er dus sterkere verbindingen tussen woorden dan tussen verschillende clusters. Via deze netwerk benadering zou er een koppeling gemaakt kunnen worden met AI. Zo zouden semantische netwerken bv. gebruikt kunnen worden voor het representeren van kennis in AI (Martin, 2019).

Om dit te verwezenlijken zal er eerst meer duidelijkheid moeten ontstaan over de organisatie van het mentale lexicon. Dit kan o.a. gedaan worden met behulp van woordassociaties. Woordassociaties zijn de woorden waar wij aan denken bij het horen van een bepaald stimulus woord. Bij het stimulus woord ‘appel’ zou je bijvoorbeeld de associaties ‘fruit’, ‘peer’ of ‘rood’ kunnen hebben. Bij woordassociaties gaat het dus ook om bepaalde verbindingen die aanwezig zijn tussen woorden. Aan de hand hiervan zou bepaald kunnen worden wat voor woordclusters in het mentale lexicon aanwezig zijn. Het onderzoeken van het mentale lexicon door middel van woordassociaties wordt o.a. gedaan in het onderzoek van Precosky (2011), De Deyne, Verheyen en Storms (2014) en De Deyne, Navarro en Storms (2013). In dit onderzoek werd er getracht meer duidelijkheid te vinden over de organisatie van het mentale lexicon van mensen die Engels als moedertaal hebben en mensen waarvan de moedertaal niet Engels is.

Om te kijken naar de mogelijke ontwikkeling van het mentale lexicon door de jaren heen wil ik de volgende onderzoeksvraag beantwoorden: Is er een verschil in woordassociaties tussen volwassenen en kinderen? Om deze vraag te beantwoorden heb ik een aantal deelvragen bedacht: Is er een verschil in woordassociaties tussen kinderen en volwassenen kijkend naar het wisselen van woordsoort bij de reacties? Is er een verschil in woordassociaties tussen kinderen en volwassenen kijkend naar de diversiteit van de reacties? De rest van de deelvragen staan in de Achtergrond Literatuur.

Voor dit onderzoek zal ik een experiment uitvoeren dat gebruik maakt van een Word Associations Test (WAT). Ik zal vragen naar de eerste associatie die een participant heeft bij een bepaald stimulus woord. De stimuluswoorden die ik ga gebruiken zullen van verschillende woordsoorten zijn, met o.a. zelfstandig naamwoorden en werkwoorden. Dit experimentele deel van het onderzoek zal gedaan worden met kinderen. De resultaten van kinderen zullen vergeleken worden met de Nederlandse database van het Small World of Words project (De Deyne, Navarro & Storms, 2013). Dit is een grootschalige wetenschappelijke studie waarin het mentale lexicon wordt onderzocht met de woorden die een gemiddelde persoon kent. Voor het project zijn o.a. van de Nederlandse en Engelse taal woordassociaties verzameld. De Nederlandse database bestaat uit meer dan 5 miljoen antwoorden. Het project wordt verder toegelicht in de Achtergrond Literatuur sectie. Uiteindelijk zal ik met de resultaten daarvan een vergelijking maken met een kunstmatig model.

Voor de verbinding met kunstmatige intelligentie zal ik een klein vergelijkingsexperiment doen met een kunstmatig model dat gebruik maakt van het Word2Vec model (Mikolov, Chen, Corrado & Dean, 2013). Dit is een kunstmatig model dat wordt gebruikt bij het eenvoudig weergeven van natuurlijke taalverwerking en bij het ophalen van informatie. Ik zal dit model gaan vergelijken met de resultaten van het experiment van Brown en Berko (1960) die woordassociaties van Engelstalige kinderen verzameld hebben. Zou het bijvoorbeeld kunnen zijn dat een kunstmatig model qua woordassociaties vooral lijkt op een kind of juist op een volwassenen. Of misschien lijkt het wel helemaal niet op beide.

Deze scriptie is als volgt ingedeeld. Ik zal beginnen met een stukje achtergrondliteratuur over een aantal belangrijke begrippen binnen mijn scriptie. Hierbij zal ik een toelichting geven over een aantal factoren die van invloed zouden kunnen zijn op verschillen in woordassociaties tussen volwassenen en kinderen. Verder zal ik nog iets uitleggen over het Word2Vec model en over het Small World of Words project. Na het verzamelen van deze informatie zal ik de onderzoeksvraag en alle deelvragen benoemen. Hierbij zal ik een hypothese opstellen die gebaseerd is op de achtergrond literatuur. Vervolgens zal ik de methode van het experiment toelichten met o.a. de participanten, procedure en het materiaal. Verder zal ik de resultaten van dit onderzoeken presenteren in de vorm van een statische analyse. De resultaten zullen bestaan uit verschillende onderdelen, waarbij ieder onderdeel ingaat op een andere deelvraag. Ook zal ik een aantal bijzondere bevinden benoemen. Tot slot zal ik in de discussie een conclusie trekken uit de resultaten die ik heb verzameld. Ook zal ik een vergelijkingsexperiment uitvoeren met een kunstmatig model. Als laatste zal ik nog beperkingen van dit onderzoek en mogelijk toekomstig onderzoek benoemen.

# 2. Achtergrond Literatuur

In deze sectie zal ik uitleg geven over een aantal belangrijke begrippen binnen mijn scriptie. Ik zal een aantal factoren benoemen die invloed zouden kunnen hebben op woordassociaties. Dit zijn o.a. woordsoort en woord valentie waarde. Deze factoren zal ik toelichten aan de hand van het benoemen van een aantal belangrijke onderzoeken naar deze onderwerpen. Hierbij focus ik me op de verschillen tussen volwassenen en kinderen. Verder zal ik nog een korte uitleg geven over het Word2Vec model en het Small World of Words project. Ik zal deze sectie eindigen met het benoemen van de onderzoeksvraag en het opstellen van een hypothese.

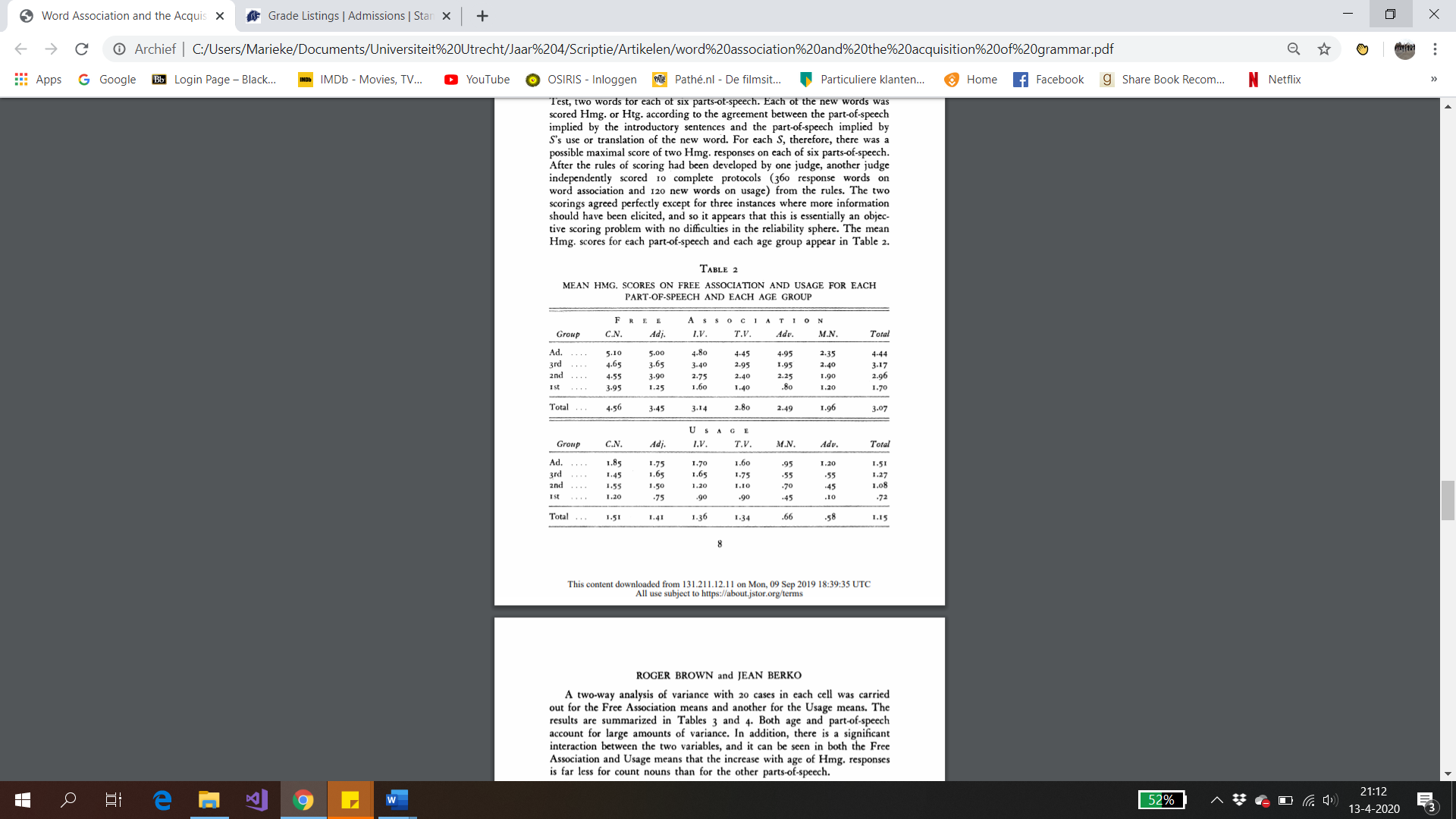
## 2.1 Factoren die invloed kunnen hebben op woordassociaties

### 2.1.1 Woordsoort/Woordcombinatie

Alle woorden vallen onder een bepaalde woordsoort. Zo valt het woord ‘tafel’ onder de woordsoort zelfstandig naamwoord en het woord ‘fietsen’ onder de woordsoort werkwoord. Voor mijn eigen onderzoek heb ik me beperkt tot zelfstandige naamwoorden, werkwoorden en bijvoeglijke naamwoorden. Uit een experiment van Woodrow en Lowell uit 1916 is al gebleken dat kinderen en volwassenen van elkaar verschillen met betrekking tot hun woordassociaties. Het bleek namelijk dat kinderen over het algemeen veel contiguïteit (bv. vuur-brand) en geheel-deel (bv. computer-scherm) associaties hadden. Volwassen hadden daarentegen veel coördinatie (behorend tot dezelfde groep, bv. appel-peer), contrast (bv. zwart-wit) en gelijkenis (bv. huilen-wenen) associaties. Soms spraken Woodrow en Lowell van “adjective-noun” associaties (bv. wit-kleur) en “verb-object” associaties (bv. verven-schilderij) (beide komen vaker voor bij kinderen dan bij volwassenen). Deze classificaties op basis van woordsoort die kinderen en volwassenen van elkaar onderscheiden bleken uiteindelijk te generaliseren te zijn naar een bepaald principe die voor het eerst werd benoemd door Ervin (1957). Het principe stelt dat de associaties van volwassenen vaker tot dezelfde woordsoort behoren als de woordsoort van de stimulus dan de associaties van kinderen.

Bij een onderzoek van Brown en Berko (1960) over woordassociaties wordt de manier van woordassociaties van volwassenen “homogenous-by-part-of-speech” (Hmg.) genoemd. De manier van associëren van kinderen wordt “heterogeneous-by-part-of-speech” (Htg.) genoemd. Brown en Berko onderzochten of dit principe daadwerkelijk klopte. Oftewel, ze onderzochten of bij volwassenen de woordsoort van het stimuluswoord en de reactie over het algemeen gelijk waren en bij kinderen juist verschillend. Voor het onderzoek werd gebruik gemaakt van 4 groepen participanten. Iedere groep bestond uit 20 deelnemers. De eerste drie groepen waren kinderen uit respectievelijk groep 1, 2 en 3 (leeftijd 6-9 jaar) en de laatste groep bestond uit studenten en volwassenen. Het experiment bevatte 36 stimulus woorden die waren onder te verdelen in 6 woordsoorten. Vervolgens werd aan de participanten hun eerste associatie bij ieder stimuluswoord gevraagd. Uit de resultaten bleek dat volwassenen inderdaad een neiging hadden tot Hmg. associaties. Bij kinderen kwam dit principe veel minder vaak voor. Wat ook is gebleken is dat hoe ouder kinderen worden hoe groter de neiging werd om een Hmg. associatie te geven. De resultaten waren bepaald aan de hand van een bepaalde score met een maximum van 6. Hoe hoger de score hoe meer homogeen de reacties waren. In tabel 1 is duidelijk te zijn dat volwassenen veel meer homogeen waren en dat de kinderen naarmate ze ouder werden meer homogeen werden.

Tabel : Gemiddelde Hmg. scores voor iedere woordsoort en groep.



Dat kinderen naarmate ze ouder worden meer gebruik gaan maken van het Hmg. principe met betrekking tot woordassociaties bleek ook uit het onderzoek van Ervin (1961). Bij dit onderzoek werden kinderen van de kleuterschool (23 deelnemers, 4-6 jaar) en uit de eerste (10 deelnemers, 6-7 jaar), derde (52 deelnemers, 8-9 jaar) en zesde (99 deelnemers, 11-12 jaar) klas onderzocht aan de hand van een free en two-choice association test. Uit de resultaten bleek dat er een toename was van het gebruik van dezelfde woordsoort in de reactie als de woordsoort van het stimuluswoord wanneer de kinderen ouder werden. In tabel 2 is te zien dat naarmate de kinderen ouder werden het percentage homogene reacties omhoog ging.

Tabel : Percentage homogene reacties per woord per groep.

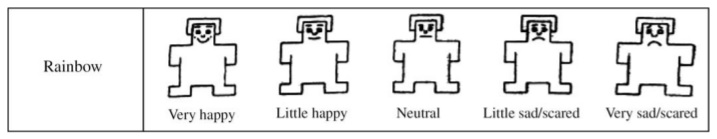


In de discussie gingen Brown en Berko (1960) iets dieper in op de reacties van de kinderen. Het bleek namelijk dat ze er niet expliciet voor kozen om een reactie te geven met een woordsoort die niet gelijk is aan de woordsoort van de stimulus. Het bleek namelijk dat de reacties van zowel de kinderen als de volwassenen semantisch gerelateerd zijn aan het stimuluswoord. In het artikel noemen ze als voorbeeld het stimuluswoord ‘to send’. Kinderen hadden op dit woord de volgende reacties: ‘away’, ‘letter’, ‘a card’, ‘mail’, etc. Volwassenen hadden als reacties: ‘to receive’, ‘to get’, ‘to deliver’, to bring’, ‘to mail’, etc. Hieruit kun je concluderen dat kinderen veel phrase completions hadden als reactie. Dat terwijl de reacties van de volwassenen over het algemeen nooit ‘to send’ zullen opvolgen. Bij volwassenen kwamen de reacties overeen in dat het allemaal transitieve werkwoorden waren. Dus het is misschien beter om te zeggen dat kinderen bij het geven van een associatie veel gebruik maken van phrase completions dan om te zeggen dat kinderen als reactie er expliciet voor kiezen om een reactie te geven die een andere woordsoort heeft dan het stimuluswoord. Voor de rest van deze scriptie zal ik het begrip woordcombinatie definiëren als een phrase completion.

### 2.1.2 Woord Valentie Waarde

Binnen de psychologie is valentie een bepaalde positieve of negatieve waarde van een gebeurtenis, object of situatie. Zo worden ook woorden geassocieerd met een positieve of negatieve emotionele waarde. Dit wordt de woord valentie waarde genoemd. Over het algemeen geldt hoe hoger de waarde, hoe positiever de emotionele waarde. Andersom geldt dan, hoe lager het getal, hoe negatiever de emotionele waarde. Uit een onderzoek van Moors et al. (2013), die voor 4300 Nederlandse woorden de valentie waarde hebben bepaald, blijkt bijvoorbeeld dat het woord ‘geluk’ een waarde heeft van 6,39 uit 7 en het woord ‘dood’ een waarde van 1,38 uit 7. Voor het bepalen van de emotionele waarde van een woord wordt over het algemeen uitgegaan van een schaal van 1 tot 7, waarbij 1 de meest negatieve en 7 de meest positieve waarde. Aan het onderzoek hebben 224 studenten meegedaan. Deze studenten kwamen van 2 universiteiten uit België en 3 universiteiten uit Nederland. De 4300 woorden waren gekozen vanuit verschillende bronnen zodat iedere woordsoort ongeveer evenveel woorden had. Iedere participant kreeg alle 4300 woorden te zien en moest op een schaal van 1 tot 7 aangeven hoe positief/negatief het woord was. Hier werd uiteindelijk het gemiddelde van genomen.

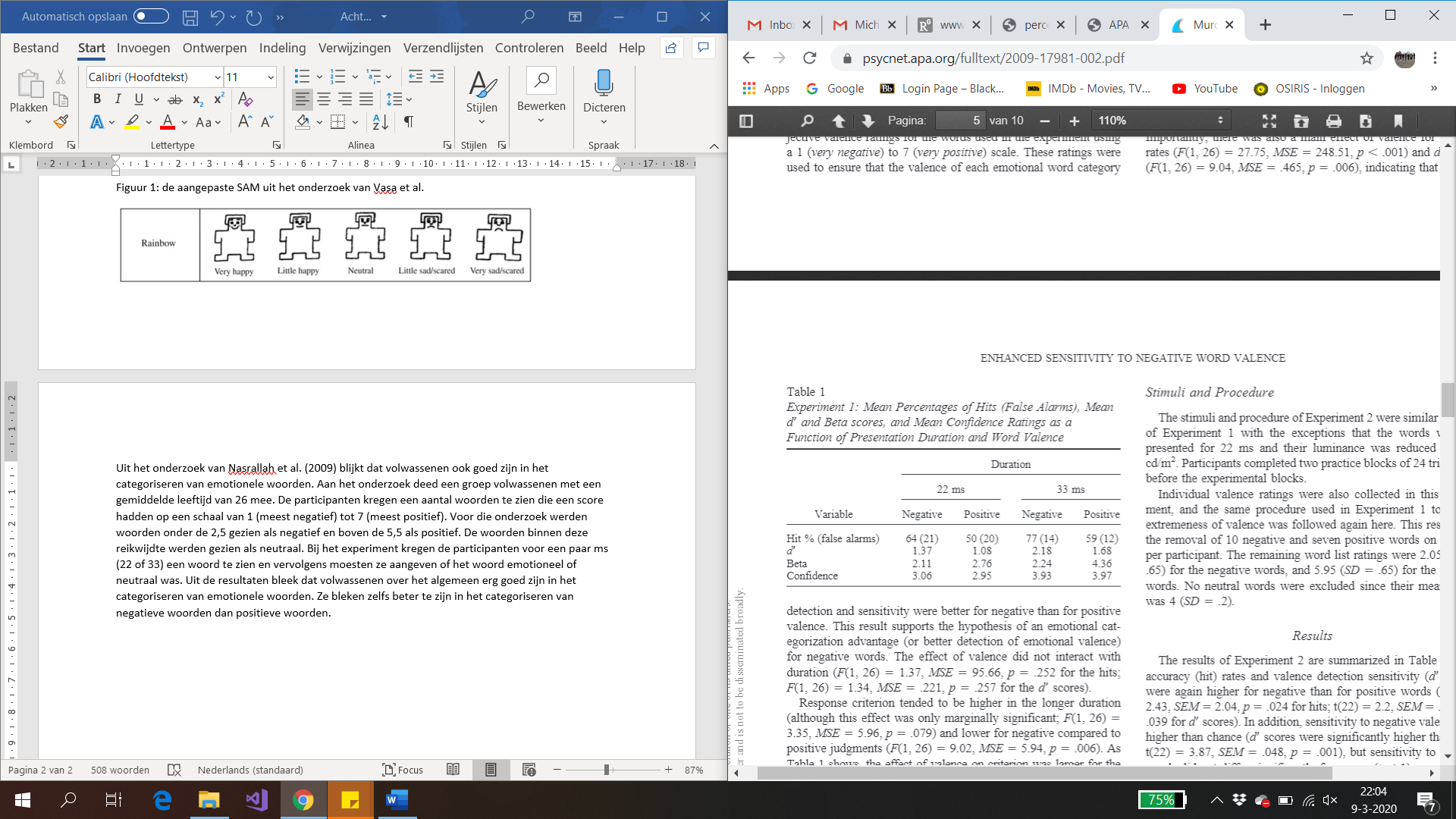
In een onderzoek van Vasa, Carlino, London en Min (2006) werden de valentie waarden van kinderen voor verschillende types emotionele woorden verzameld. Het ging om 149 kinderen tussen de 9 en 11 jaar oud. Het onderzoek bevatte drie woord categorieën: bedreiging, positief en neutraal. De kinderen kregen een boekje met daarin 81 woorden waarvan ze de valentiewaarde moesten bepalen aan de hand van een aangepast Self-Assesment Manikin (SAM). De originele SAM was aangepast op 2 punten: (1) de zinnen “how happy?” en “how sad/scared?” werden gebruikt in plaats van ‘pleasant’ of ‘unpleasant’; (2) in plaats van een schaal van 9 punten werd er een van 5 punten gebruikt. In figuur 1 is een voorbeeld te zien van hoe de aangepaste SAM er uit zag. Uit de resultaten bleek dat kinderen de drie woordcategorieën goed weten te onderscheiden aan de hand van hun gegeven valentie waarden. Zo waren er voor de positieve woorden vooral hoge valentiewaarden en voor de bedreigingen vooral lage valentiewaarden gegeven. Ook voor de neutrale woorden gaven de kinderen een neutrale valentiewaarde. Wat bovendien opviel was dat de meisjes meer extreme valentiewaarden gaven dan de jongens. Kinderen lijken dus woorden goed te kunnen categoriseren op de emotionele waarde van een woord.



Figuur : de aangepaste SAM uit het onderzoek van Vasa et al.

Uit het onderzoek van Nasrallah, Carmel en Lavie (2009) bleek dat volwassenen ook goed zijn in het categoriseren van emotionele woorden. Aan het onderzoek deden 27 volwassenen met een gemiddelde leeftijd van 26 mee. De participanten kregen een aantal woorden te zien die een score hadden op een schaal van 1 (meest negatief) tot 7 (meest positief). Voor die onderzoek werden woorden onder de 2,5 gezien als negatief en boven de 5,5 als positief. De woorden binnen deze reikwijdte werden gezien als neutraal. Bij het experiment kregen de participanten voor een paar ms (22 of 33) een woord te zien en vervolgens moesten ze aangeven of het woord emotioneel of neutraal was. Uit de resultaten bleek dat volwassenen over het algemeen erg goed zijn in het categoriseren van emotionele woorden. Ze bleken zelfs beter te zijn in het categoriseren van negatieve woorden dan positieve woorden. In de tabel hieronder is bij het percentage hits te zien dat volwassenen over het algemeen meer dan de helft van de woorden goed weten te categoriseren. Wanneer ze het woord langer te zien kregen werden zo ook beter. Dus net als kinderen zijn ook volwassenen goed in het categoriseren van emotionele woorden.

Tabel : Gemiddelde percentage hits.



In het onderzoek van Bahn, Kauschke, Vesker en Schwarzer (2018) werden volwassenen met kinderen vergeleken. Er werd onderzocht of volwassenen en kinderen de emotionele valentie en arousal van woorden op de zelfde manier waarnemen. Ik ga verder alleen in op de delen van het onderzoek die over de emotionele valentiewaarde gaan. Aan het onderzoek deden 3 groepen mee: een groep volwassen en een groep van 9 jaar oude kinderen die getest werden in een lab omgeving, en een groep volwassenen die werden getest via een online vragenlijst. De stimuli bestonden uit 48 positieve en negatieve Duitse woorden. Deze werden gekozen uit de BAWL-R database. De participanten moesten aan de hand van de Self-Assesment Manikin (SAM) scale aangeven hoe positief of negatief een woord was (7-punt Likert schaal). Uit de resultaten bleek dat er sterke positieve correlaties waren tussen de 2 groepen volwassenen en de data van de BAWL-R database. Wanneer de resultaten van de kinderen met de volwassenen wordt vergeleken blijkt dat hiertussen ook hoge correlaties zijn. Dit zou kunnen suggereren dat kinderen met een leeftijd van 9 jaar zich lijken te gedragen als volwassenen kijkend naar het bepalen van de emotionele waarde van woorden.

### 2.1.3 Age of Acquisition

De Age of Acquisition (AoA) is de leeftijd waarop over het algemeen een bepaald woord wordt geleerd. Het woord ‘mama’ (AoA: 2,04) en ‘papa’ (AoA: 2,76) wordt vaak al vroeg aangeleerd. Terwijl woorden zoals ‘verzekering’ (AoA: 10,55) en ‘universiteit’ (AoA: 11,28) vaak een veel hogere Age of Acquisition hebben. Deze AoA waarden komen uit het onderzoek van Brysbaert, Stevens, De Deyne, Voorspoels en Storms (2014), waarin de Age of Acquisition waarden van 30.000 Nederlandse woorden was bepaald. De waardes zijn berekend aan de hand van een experiment met 74 studenten en onderzoekers van de universiteit van Ghent. Iedere student kreeg 6500 woorden te zien. Voor ieder woord kregen ze de vraag om te beoordelen op welke leeftijd ze het woord hadden aangeleerd. Voor het woord ‘banaan’ werd bv. vaak 3 jaar ingevuld en voor het woord ‘boekhouder’ 11. Van de antwoorden werd uiteindelijk het gemiddelde genomen.

In een onderzoek van Garlock, Walley en Metsala (2001) werden de effecten van Woordfrequentie, Age of Acquisition en Neighborhood Density op gesproken woord herkenning gemeten. Voor deze sectie zal ik me beperken tot de effecten van Age of Acquisition. In deze studie werden 3 groepen onderzocht: kleuters, basisschool kinderen en volwassenen. De gemiddelde leeftijd was respectievelijk 5.6, 7.6 en 25.6 en iedere groep bestond uit 64 deelnemers. Het experiment werd gedaan door middel van een word repitition task. De stimuluswoorden werden gekozen zodat er een variatie aan woorden aanwezig was van lage en hoge AoA woorden. Uit de resultaten bleek dat zowel kinderen als volwassenen sneller waren in het herkennen van vroeg verworven woorden. Kinderen waren echter sneller in het herkennen van vroeg verworven woorden dan volwassenen. Bij later verworven woorden werd het omgekeerde waargenomen. Volwassenen waren sneller in het herkennen van later verworven woorden dan de kinderen. Dit tegenovergestelde effect was erg bijzonder en dit onderzoek was de eerste die dit effect aantoonden. AoA lijkt dus een positief effect te hebben op de woordherkenning van zowel volwassenen als kinderen.

Verdere effecten van Age of Acquisition werden onderzocht in het onderzoek van Juhasz (2005). Dit is een samenvattend paper over de effecten van AoA op verschillende taken. In dit onderzoek woerd o.a. de invloed van AoA onderzocht voor picture naming en word naming. Qua picture naming werd er voor de meeste onderzoeken een effect gevonden van AoA. Uit resultaten blijkt dat hoe vroeger het afgebeelde is geleerd hoe sneller het wordt benoemd. Voor experimenten met betrekking tot het benoemen van objecten bleek dat er significante effecten zijn op de snelheid van het benoemen van het object. Ook voor het herkennen van woorden is er een significant effect. Er geldt dat hoe lager de AoA waarde hoe sneller het object wordt benoemd en hoe sneller het woord wordt herkend.

Waar ik echter geen onderzoek over heb kunnen vinden is of de AoA van het stimuluswoord invloed heeft op de AoA van het reactiewoord. Laat staan of er een verschil hierin zou zijn tussen kinderen en volwassenen. Zou het bijvoorbeeld kunnen zijn dat wanneer het stimuluswoord een lage AoA heeft mensen automatisch reageren met een reactie die ook een lage AoA heeft. Voor dit mogelijke effect heb ik alvast gekeken naar 100 willekeurig gekozen woorden uit het Small World of Words project. Voor deze 100 woorden en de meest voorkomende reactie heb ik de AoA waarde opgezocht aan de hand van het onderzoek van Brysbaert et al. (2014). Ieder woord kwam voor in de database. Wanneer de meest voorkomende reactie niet voorkwam in de database heb ik het daaropvolgende woord gekozen. Vervolgens heb ik via SPSS een correlatie test uitgevoerd.

De gemiddelde AoA waarde van de stimuluswoorden was 7,55 en de stdv was 1,87 . Bij de reacties was de gemiddelde AoA waarde 6,01 en de stdv was 1.79. Een Pearson Correlation test was uitgevoerd om de AoA waarden van de stimuluswoorden te vergelijken met de AoA waarden van de reacties. Er was een significante correlatie tussen de stimuluswoorden en de reacties betreft AoA waarde, Pearson Correlation = 0.458, p < 0.001. Dit betekent dat er tussen de AoA waarden van de stimuluswoorden en de reacties een verband is wat niet op toeval berust. Dit resultaat zou verklaard kunnen worden door mogelijke verbindingen in het mentale lexicon met betrekking tot AoA waarden. Het zou bv. kunnen zijn dat woorden met een vergelijkbare AoA waarde bij elkaar in een cluster zitten. Deze worden dus sneller geactiveerd wanneer een bepaald woord in dit cluster geactiveerd wordt dan woorden met een hogere AoA waarden die weer in een andere cluster zitten.

De resultaten van dit experiment zijn te zien in appendix A1.

### 2.1.4 Woordfrequentie

Woordfrequentie geeft aan hoe vaak een woord voorkomt binnen een bepaalde context. Een functiewoord zoals een lidwoord heeft over het algemeen een veel hogere frequentie dan een inhoudswoord zoals een zelfstandig naamwoord. De frequentie van het woord ‘frequentie’ is bijvoorbeeld 5,95 miljoen. Het woord met de hoogste frequentie uit de database is ‘ik’ en dit woord heeft een frequentie van 39883,03 miljoen. Deze frequenties waarden zijn gebaseerd op de subtext NLdatabase van Keuleers, Brysbaert en New (2010). Deze database is gebaseerd op 44 miljoen verschillende Nederlandse woorden uit de ondertiteling van films en series.

In het onderzoek van Cofer en Shevitz (1952) werd er onderzocht of er een relatie is tussen woordfrequentie en woordassociaties. Dit werd onderzocht met behulp van 2 groepen studenten, een groep mannen en een groep vrouwen. De participanten kregen een aantal woorden te zien die een hoge of een lage frequentie hadden. Vervolgens kregen ze de opdracht om zoveel mogelijk associaties bij dat woord op te schrijven. Uit de resultaten bleek dat er geen significante verschillen waren tussen de mannen en de vrouwen. Als er werd gekeken naar het totaal aantal gegeven associaties werden er significant meer associaties gegeven bij de hoge frequentie woorden dan bij de lage frequentie woorden. Deze resultaten laten duidelijk zien dat er een relatie is tussen de frequentie van een stimuluswoord en het aantal associaties die bij dat woord worden gegeven.

Het onderzoek van Morrison en Ellis (1995) ging dieper in op de invloed van Age of Acquisition en woordfrequentie op Word Naming en Lexical Decision. Voor deze sectie zal ik me focussen op de invloeden van woordfrequentie. Het wordt over het algemeen geaccepteerd dat woord frequentie een significante impact heeft op de snelheid van lexical decision. Hoge frequentie woorden worden over het algemeen veel sneller herkend als echte woorden dan lage frequentie woorden. In dit onderzoek vond een experiment plaats om te onderzoeken of deze onderzoekers ook dezelfde resultaten zouden krijgen. Hiervoor werden 16 studenten gebruikt als participanten. Deze participanten kregen een mix van hoge en lage frequentie woorden en non-woorden te zien. Aan de studenten de taak om aan te geven of het woord wat ze zagen een echt woord of een non-woord was. Uit de resultaten bleek dat er inderdaad een significant verschil was in reactiesnelheid tussen hoge en lage frequentiewoorden. Op hoge frequentie woorden werd er veel sneller een reactie gegeven dat het echte woorden waren dan bij lage frequentie woorden. Woordfrequentie blijkt dus een significante impact te hebben op de reactiesnelheid van mensen met betrekking tot lexical decision.

Voor woordfrequentie geldt hetzelfde als voor AoA. Ook hiervoor heb ik geen onderzoek kunnen vinden of de frequentie waarde van het stimuluswoord invloed heeft op de frequentie waarde van het reactiewoord. Laat staan of er een verschil hierin zou zijn tussen kinderen en volwassenen. Zou het bijvoorbeeld kunnen zijn dat wanneer het stimuluswoord een hoge frequentie waarde heeft mensen automatisch reageren met een reactie die ook een hoge frequentie waarde heeft. Voor dit mogelijke effect heb ik alvast gekeken naar 100 willekeurig gekozen woorden uit het Small World of Words project. Het gaat hier om dezelfde woorden als die voor de AoA waarden. Voor deze woorden en de meest voorkomende reactie heb ik de frequentie waarde opgezocht aan de hand van de database van Keuleers et al. (2010). Ieder stimuluswoord kwam voor in de database. Wanneer de meest voorkomende reactie niet voorkwam in de database heb ik het daaropvolgende woord gekozen. Vervolgens heb ik via SPSS een correlatie test uitgevoerd.

De gemiddelde frequentie waarde van de stimuluswoorden was 32,84 en de stdv was 89,65. Bij de reacties was de gemiddelde frequentie waarde 126,78 en de stdv was 234,79. Een Pearson Correlation test was uitgevoerd om de frequentie waarden van de stimuluswoorden te vergelijken met de frequentie waarden van de reacties. Er was geen significante correlatie tussen de stimuluswoorden en de reacties betreft frequentie waarde, Pearson Correlation = 0.193, p =0.055. Dit betekent dat er tussen de frequentie waarden van de stimuluswoorden en de reacties geen verband is.

De resultaten van dit experiment zijn te zien in appendix A2.

### 2.1.5 Diversiteit

In deze scriptie zal ik diversiteit definiëren als de hoeveelheid **verschillende** reacties die er worden gegeven op een stimuluswoord. Stel er is een experiment waaraan 5 volwassenen en 5 kinderen meedoen. Aan de participanten wordt gevraagd om hun eerste associatie te geven bij het woord ‘banaan’. De reacties van de 5 volwassenen zijn respectievelijk: fruit, peer, fruit, fruit en appel. De kinderen geven de volgende reacties: fruit, fruit, appel, fruit en fruit. In dit geval zijn de volwassenen meer divers dan de kinderen. De volwassenen hebben namelijk in totaal 3 verschillende reacties gegeven: appel, peer en fruit. Terwijl de kinderen in totaal maar 2 verschillende reacties hebben gegeven, namelijk: fruit en appel.

In het onderzoek van Zortea en de Salles (2012) werd er o.a. ingegaan op de diversiteit van kinderen en volwassenen bij een woordassociatie taak. Aan deze studie deden een groot aantal kinderen (leeftijd tussen de 7 en 11 jaar oud) en volwassenen (leeftijd tussen 16 en 49 jaar oud) mee. De participanten werden getest op forward associative strength en set size door middel van een semantische woordassociatie taak. Een semantische woordassociatie taak verschilt van een free association taak in het feit dat bij de semantische variant wordt gevraagd om een associatie die een betekenis heeft die gerelateerd is aan het stimuluswoord. Bij free association mag het eerste woord wat in je op komt genoemd worden. Voor deze sectie zal ik me beperken tot de resultaten met betrekking tot set size, omdat dit vooral te maken heeft met de diversiteit. Met meaning set size wordt het aantal verschillende reacties op een stimuluswoord van 2 of meer mensen bedoeld. Meaning set size is klein wanneer er maar weinig verschillende associaties zijn en groot wanneer er veel verschillende associaties zijn bij een stimuluswoord. Om een algemene maat te hebben voor de diversiteit hebben de onderzoekers Shannon’s H statistic gebruikt. Dit is een response diversity index voor een groep mensen. Uit de resultaten bleek dat volwassenen een hogere response diversity index hadden dan de kinderen. Dit betekent dat de volwassenen meer divers waren in hun reacties. De onderzoekers van deze studie verwachten dat dit komt omdat kinderen nog een relatief kleine woordenschat hebben. De woordenschat van iemand groeit naarmate je ouder wordt.

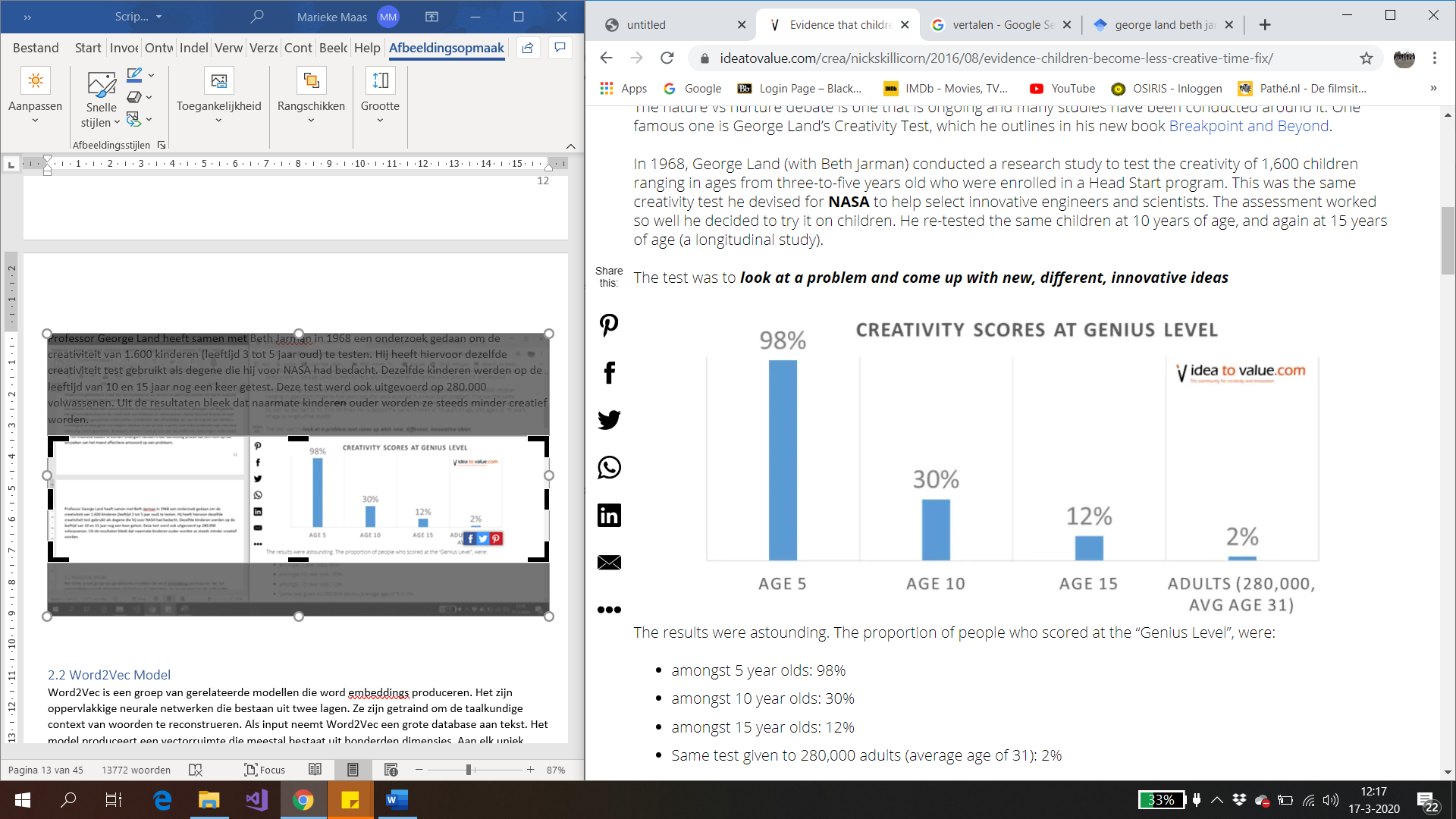
In het onderzoek van Zortea en de Salles (2012) wordt ook een ander onderzoek genoemd dat de diversiteit van kinderen en volwassenen heeft onderzocht (Coronges, Stacy & Valentie, 2007). In dit onderzoek werd de meaning set size van meer dan duizend kinderen van 12 en 13 jaar vergeleken met de resultaten van een onderzoek van Nelson, McEvoy en Schreiber (2004) waar meer dan zesduizend volwassenen aan mee deden. Uit het onderzoek bleek dat de meaning set size bij de kinderen kleiner was dan bij de volwassenen. Dit betekent dat de kinderen minder divers waren dan de volwassenen in hun reacties. Volgens Coronges et al. zijn deze resultaten te verklaren doordat kinderen hun reacties minder filteren en beschikken over meer uiteenlopende gedachtes. Een probleem bij dit onderzoek was echter dat de kinderen een semantische associatie taak gedaan hebben terwijl de volwassenen uit het onderzoek van Nelson et al. (2004) een free association taak gedaan hebben. Door een semantische association taak wordt je al beperkt in welke reacties je kunt geven, namelijk alleen associaties die een betekenis hebben die gerelateerd is aan het stimuluswoord. Dit verschil in taken zou kunnen verklaren waarom de kinderen minder divers waren dan de volwassenen.

Een verschil in diversiteit tussen kinderen en volwassenen zou o.a. kunnen afhangen van de grootte van de woordenschat van kinderen en volwassenen. Ook creativiteit kan een rol spelen. Dit zijn 2 aparte factoren die onafhankelijk van elkaar zijn. Creativiteit kan afhankelijk zijn van de manier van denken, convergent of divergent. Convergent denken is een proces waarbij voor ieder probleem een concrete oplossing wordt gevonden. Divergent denken is een proces dat verschillende oplossingen onderzoekt om tot creatieve ideeën te komen. Bij convergent denken wordt er gericht op één ding, terwijl bij divergent denken meerdere opties worden opgehouden.

Professor George Land heeft samen met Beth Jarman in 1968 een onderzoek gedaan om de creativiteit van 1.600 kinderen (leeftijd 3 tot 5 jaar oud) te testen. Ze hebben hiervoor dezelfde creativiteit test gebruikt als degene die ze voor NASA hadden bedacht om de creativiteit van raket wetenschappers en ingenieurs te testen. Over wat deze test precies inhield is echter niks te vinden. Dezelfde kinderen werden op de leeftijd van 10 en 15 jaar nog een keer getest. Deze test werd ook uitgevoerd op 280.000 volwassenen. Uit de resultaten bleek dat naarmate kinderen ouder werden ze steeds minder creatief werden. In figuur 2 zijn de resultaten van dit onderzoek te zien. Volgens Land komt dit grote verschil omdat er 2 verschillende manier van denken zijn, namelijk convergent en divergent denken. Kinderen maken in eerste instantie veel gebruik van divergent denken. Zodra ze op school komen wordt hen convergent denken aangeleerd. Hierdoor verdwijnt het divergent denken, met als gevolg minder creativiteit (Skillicorn, 2016).

Aan de hand van dit onderzoek kun je dus stellen dat jonge kinderen over het algemeen gebruik maken van divergent denken en volwassenen van convergent denken. Volwassenen zijn dus minder creatief dan kinderen. Omdat kinderen creatiever zijn zou je kunnen verwachten dat ze meer verschillende reacties weten te benoemen bij een woordassociatie taak.

Figuur : Scores op de creativiteitstest van George Land.



Een andere factor die invloed zou kunnen hebben op de diversiteit is de grootte van de woordenschat. Volwassenen in Nederland met een leeftijd van 30 jaar hebben een gemiddelde woordenschat van 71.000 woorden (Brysbaert, Keuleers, Mandera & Stevens, 2013). Naarmate iemand ouder wordt zal gemiddelde zelfs een woordenschat bereikt kunnen worden van 80.000 woorden. Nederlandse kinderen van 4 jaar hebben een gemiddelde woordenschat hebben van 3.200 woorden en kinderen van 8 jaar 5.600 woorden (Appel, Kuiken & Vermeer, 1995). Volwassenen kennen dus veel meer woorden dan kinderen. Hierdoor zou je kunnen verwachten dat volwassenen meer verschillende reacties hebben bij een woordassociatie taak, puur omdat ze ook meer woorden kennen.

## 2.2 Word2Vec Model

Word2Vec is een groep van gerelateerde modellen die word embeddings produceren. Het model is gecreëerd door een team van onderzoekers onder leiding van Mikolov (Mikolov, Chen, Corrado & Dean, 2013). Het zijn oppervlakkige neurale netwerken die bestaan uit twee lagen. Ze zijn getraind om de taalkundige context van woorden te reconstrueren. Als input neemt Word2Vec een grote database aan tekst. Het model produceert een vectorruimte die meestal bestaat uit honderden dimensies. Aan elk uniek woord in de database wordt een overeenkomstige vector in de vectorruimte toegewezen. Woordvectoren worden in de vectorruimte geplaatst zodat woorden die een gemeenschappelijk context hebben in de database dicht bij elkaar staan in de vectorruimte.

Met voldoende gegevens kan het Word2Vec model zeer nauwkeurige inschattingen maken over de betekenis van een woord op basis van eerdere voorkomens. Deze inschattingen kunnen gebruikt worden om de associatie van een woord met een ander woord vast te stellen (bv. het woord ‘man’ is tot het woord ‘jongen’, wat het woord ‘vrouw’ tot het woord ‘meisje’ is). Ook kan het documenten samenbundelen en ze op onderwerp classificeren. Deze bundels kunnen de basis vormen voor analyses en aanbevelingen in de meest uiteenlopende gebieden, zoals wetenschappelijk onderzoek, juridische ontdekkingen en klantrelatiebeheer (Nicholson, 2019).

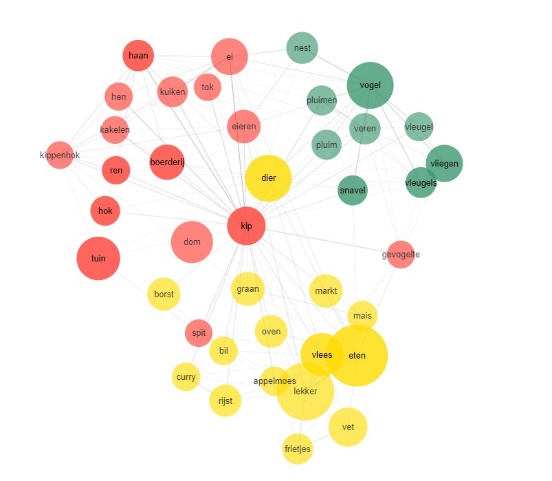
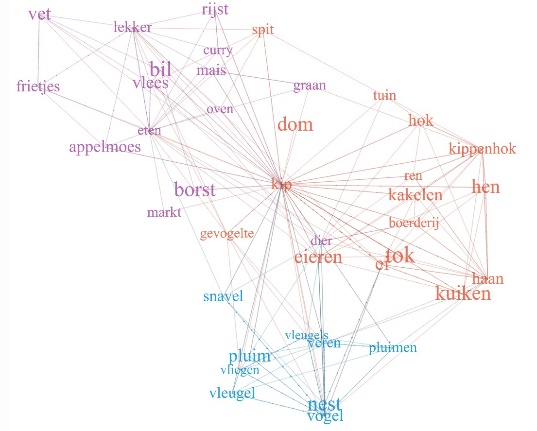
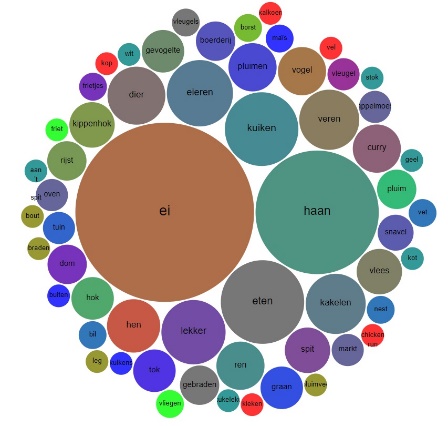
Zoals ik in de introductie al had aangegeven zal ik een klein vergelijkingsexperiment uitvoeren tussen de resultaten van het onderzoek van Brown en Berko (1960) en een kunstmatig model. Voor het kunstmatige model heb ik gebruik gemaakt van een online demo van The Turku NLP Group. Dit is een groep onderzoekers van zowel de Turku Universiteit als de UTU graduate school (UTUGS). Ze focussen zich vooral op onderzoek met betrekking tot het verwerken van natuurlijke taal en taal technologie. Voor het onderzoek hebben ze o.a. een online demo gemaakt die gebaseerd is op word embeddings en gebruik maakt van het Word2Vec model. Het was origineel getraind op 4,5 biljoen Finse woorden, maar tegenwoordig is het ook getraind op Engelse woorden. Voor de Engelse woorden is als database het Engelse Google nieuws gebruikt. De reden waarom ik ervoor heb gekozen om de resultaten van het kunstmatige model te vergelijken met het onderzoek van Brown en Berko (1960) in plaats van mijn eigen experiment is omdat de demo niet op Nederlandse woorden is getraind. Ik doe mijn eigen experiment namelijk in het Nederlands. Het onderzoek van Brown en Berko (1960) is wel in het Engels gedaan.

## 2.3 Het Small of World of Words project

Het Small World of Words project is een grootschalige wetenschappelijke studie met als doel het in kaart brengen van het menselijke lexicon voor de belangrijkste talen in de wereld. Het project is gestart in 2003 op de afdeling Experimental Psycholinguistics van de universiteit in Leuven. Sindsdien heeft het project een groot netwerk van woordassociaties gemaakt voor het Nederlands (meer dan 5 miljoen antwoorden) en het Engels (meer dan 1 miljoen antwoorden). Ook de volgende talen krijgen in dit project steeds meer reacties: Frans, Duits, Spaans, Italiaans, Cantonees, Mandarijns en Russisch. Iedereen kan deelnemen aan het project en mag de resultaten uit het project gebruiken voor eigen onderzoek.

Iedereen kan meedoen aan het project, ongeacht leeftijd of opleiding. Om mee te doen selecteer je je eigen taal. Vervolgens wordt er om leeftijd, geslacht, opleidingsniveau en moedertaal gevraagd. Hierna volgt een korte instructie van het experiment. Bovenaan in het scherm komt telkens een woord te staan. De bedoeling is om de eerste 3 associaties die spontaan in je opkomen te noteren. Wanneer je het woord niet kent kun je dat aangeven. Ook wanneer je minder dan 3 associaties hebt bij het woord kun je dat aangeven. In totaal komen er 18 woorden voorbij en het zal ongeveer 5 minuten duren. De resultaten worden automatisch opgenomen in de database van de desbetreffende taal.

Voor mijn onderzoek ga ik de Nederlands database gebruiken. Deze bestaat uit woordassociaties en deelnemer informatie van de 100 eerste, tweede en derde associaties op 12.571 stimuluswoorden. Dit is de data die tot en met november 2010 verzameld is (De Deyne, Navarro & Storms, 2013). Ook kunnen er via de website vroegere databases gevonden worden.

Naast de database kun je ook via de website de associaties opzoeken van verschillende stimuluswoorden. Wanneer je bijvoorbeeld het woord “kip” invoert krijg je als resultaat dat er 124 keer de reactie “ei” is gegeven en 59 keer de reactie “haan”. Ook kun je een visuele representatie van de associaties opzoeken. Dit kan op verschillende manieren, namelijk in de vorm van een direct netwerk, een 3D netwerk en een bubbel grafiek.

Figuur : Visuele representatie mogelijkheden van het Small World of Words project. Links is een direct netwerk, in het midden een 3D netwerk en rechts een bubbel grafiek

Hier zijn voorbeelden met het woord “kip” te zien van de visuele representatie mogelijkheden van het Small World of Words project. Links is een direct netwerk, midden een 3D netwerk en rechts een bubbel grafiek. In het directe netwerk en 3D netwerk zijn ook de connecties aangegeven tussen de associaties van “kip” mochten die aanwezig zijn.

## 2.4 Korte Samenvatting

In deze sectie heb ik een aantal belangrijk elementen uit mijn scriptie toegelicht. Ik begon met het toelichten van een aantal factoren die mogelijk invloed hebben op een verschil in woordassociaties tussen volwassenen en kinderen. Deze factoren heb ik nader toegelicht aan de hand van het benoemen van een aantal onderzoeken.

Voor het wisselen van woordsoort en woordcombinatie stellen Brown en Berko (1960) dat volwassenen en kinderen van elkaar verschillen. Zo zouden volwassenen over het algemeen met hun associaties bij de zelfde woordsoort blijven als de woordsoort van het stimuluswoord. Kinderen daarentegen zullen vaak van woordsoort wisselen bij hun associaties. Ook met betrekking tot woordcombinaties zou er een verschil zijn. Kinderen hebben vooral veel phrase completions terwijl dit bij volwassenen bijna niet voorkwam.

Qua woord valentie waarde blijken kinderen en volwassenen veel op elkaar te lijken. Uit onderzoeken is gebleken dat zowel kinderen als volwassenen goed zijn in het inschatten van de emotionele waarde van een woord. In het experiment van Bahn et al. (2018) waarbij kinderen met volwassenen werden vergeleken, blijkt dat volwassenen en kinderen bijna niet verschillen met betrekking tot hun inschatting van de emotionele waarde van woorden.

Age of Acquisition lijkt vooral een effect te hebben op de snelheid van woordherkenning. Er is echter geen onderzoek gedaan naar de mogelijke invloed van de AoA van het stimuluswoord op de AoA van het reactiewoord. Uit een klein experiment wat ik heb uitgevoerd aan de hand van het Small World of Words project blijkt dat er een correlatie is tussen AoA van de stimulus en de reactie.

Ook de woordfrequentie waarde lijkt vooral een effect te hebben op de snelheid van o.a. woordherkenning. Ook hiervoor geldt echter dat er geen onderzoek is gedaan naar de mogelijke invloed van de frequentie waarde van het stimuluswoord op de frequentiewaarde van het reactiewoord. Uit een klein experiment wat ik heb uitgevoerd aan de hand van het Small World Of Words project blijkt dat er geen correlatie is tussen de frequentie waarde van de stimulus en de reactie.

Tot slot heb ik nog gekeken naar de factor diversiteit. Het blijkt dat volwassenen over het algemeen meer divers zijn dan kinderen in het geven van associaties bij een stimuluswoord. Er kan hierover gediscussieerd worden. Volwassenen hebben namelijk een grotere woordenschat dan kinderen waardoor ze waarschijnlijk meer divers zijn. Kinderen zijn echter vaak creatiever waardoor je ook zou kunnen verwachten dat juist kinderen meer divers zijn dan volwassenen.

Naast de toelichtingen over de verschillende factoren geef ik ook een uitleg over het Word2Vec model. Dit is een model wat een grote tekst als database krijgt voor input en vervolgens ieder uniek woord een vector geeft in de vectorruimte. Deze woorden worden in de vectorruimte geplaatst zodat woorden die een gemeenschappelijk context hebben in de database dicht bij elkaar staan in de vectorruimte. In de discussie zal ik een vergelijkingsexperiment uitvoeren tussen de resultaten van het onderzoek van Brown en Berko (1960) en een kunstmatig model. Het kunstmatige model waar ik voor heb gekozen is een online demo die gebruik maakt van het Word2Vec model.

Vervolgens geef ik nog een korte toelichting over het Small World of Words project. Dit is een grootschalig wetenschappelijk project met als doel het in kaart brengen van het menselijke lexicon voor de belangrijkste talen in de wereld. Iedereen kan via de website deelnemen aan het project. Wanneer je besluit mee te doen krijg je voor een specifieke taal, een voor een, een aantal woorden te zien. Vervolgens moet je op een woord je eerste 3 associaties geven. De resultaten van het project zijn vrij toegankelijk voor gebruik voor eigen onderzoek.

## 2.5 Onderzoeksvraag en hypothese

De hoofdonderzoeksvraag die ik wil beantwoorden luidt als volgt:

*Is er een verschil tussen de woordassociaties van kinderen en volwassenen?*

Ook wil ik graag nog een tweede onderzoeksvraag beantwoorden. Deze vraag luidt als volgt:

*Is er een verschil in woordassociaties tussen mensen (kinderen en volwassenen) en kunstmatige modellen?*

De eerste onderzoeksvraag wil ik gaan beantwoorden met behulp van een aantal deelvragen. Betreft de verschillen tussen kinderen en volwassenen wil ik de volgende deelvragen beantwoorden om een antwoord te geven op de hoofd onderzoeksvraag:

* *Is er een verschil kijkend naar het wisselen van woordsoort?*
* *Is er een verschil kijkend naar woordcombinaties?*
* *Is er een verschil kijkend naar de woord valentie waarde?*
* *Is er een verschil de Age of Acquisition van de stimuluswoorden en reacties?*
* *Is er een verschil kijkend naar de woordfrequentie van de stimuluswoorden en reacties?*
* *Is er een verschil kijkend naar de diversiteit van de reacties?*

Aan de hand van bovenstaande literatuur stel ik de volgende hypotheses op voor de deelvragen:

* Ja, er zal een verschil zijn kijkend naar woordsoort tussen kinderen en volwassenen. De reacties van de kinderen zullen qua woordsoort vaker afwijken van de woordsoort van het stimuluswoord dan bij de volwassenen.
* Ja, er zal een verschil zijn kijkend naar woordcombinaties tussen kinderen en volwassenen. Kinderen zullen meer woordcombinaties hebben dan volwassenen gebaseerd op de manier waarop ik woordcombinatie heb gedefinieerd.
* Nee, er zal geen verschil zijn kijkend naar de woord valentie waarde. Kinderen en volwassenen zullen hierin relatief gelijk zijn.
* Nee, er zal geen verschil zijn kijkend naar Age of Acquisition. Gebaseerd op het kleine experiment wat ik heb uitgevoerd verwacht ik dat de reacties van zowel de kinderen als de volwassenen een correlatie zullen hebben met de stimuluswoorden.
* Nee, er zal geen verschil zijn kijkend naar de woord frequentie. Gebaseerd op het kleine experiment wat ik heb uitgevoerd verwacht ik dat de reacties van zowel de kinderen als de volwassenen geen correlatie hebben met de stimuluswoorden.
* Ja, er zal een verschil zijn kijkende naar de diversiteit van kinderen en volwassenen. Volwassenen zullen meer divers zijn in hun reacties dan kinderen omdat volwassenen meer woorden kennen dan kinderen.

Aan de hand van deze hypotheses voor de deelvragen stel ik volgende hypothese op voor de hoofdonderzoeksvraag:

*Ja, er is een verschil in woordassociaties tussen volwassenen en kinderen. Dit verschil heeft betrekking op het wisselen van woordsoort, woordcombinaties en diversiteit van de reacties. Qua woord valentie waarde, AoA waarde en frequentie waarde zal er geen verschil zijn tussen kinderen en volwassenen.*

Voor de tweede onderzoeksvraag stel ik de volgende hypothese op:

*Ik verwacht dat er een verschil zal zijn tussen mensen en het model. Dit verschil zal vooral betrekking hebben op de woorden die deze twee groepen kennen. Het model is namelijk beperkt tot woorden uit het Engelse Google nieuws. Terwijl mensen een veel uitgebreidere database hebben, namelijk alle woorden die ze lezen of horen.*

# 3. Methode

In deze sectie zal ik de methode van het experiment toelichten. Ik zal het onder andere hebben over de participanten, het materiaal en de taak, en de procedure. Wat ook wordt uitgelegd is de aanpassingen die ik heb gedaan aan het experiment aan de hand van een pilot test.

## 3.1 Participanten

Aan het experiment hebben 46 kinderen (23 meisjes, 23 jongens) meegedaan met een leeftijd variërend van 4 tot en met 8 jaar oud. De gemiddelde leeftijd was 5,9 jaar oud. Deze deelnemers zijn geworven op de Basisschool Puur Sang – Locatie Sint Lucia. De kinderen zijn afkomstig uit groep 1 tot en met groep 4. Het onderzoek was geheel vrijwillig en er kon op ieder moment gestopt mee worden. Als beloning voor het deelnemen aan het onderzoek kregen de kinderen een sticker.

Van te voren heb ik de ouders van de kinderen en de schoolleiding duidelijk geïnformeerd wat het onderzoek inhoud. Dit door middel van een informatiebrief (zie appendix B). Aan alle ouders heb ik toestemming gevraagd door middel van een toestemmingsformulier (zie appendix B). Ik heb alleen de kinderen onderzocht waarvan de ouders het toestemmingsformulier hebben ondertekend. Het onderzoek. Dit onderzoek, de informatiebrief en het toestemmingsformulier zijn goedgekeurd door de FETC (reference number: zucke102-03-03-2019)

## 3.2 Materiaal en Taak

Qua opzet zal dit onderzoek lijken op het experiment wat is uitgevoerd door Brown en Berko (1960).

Voor het onderzoek heb ik gebruik gemaakt van een word associaitions test (WAT). In totaal heb ik de kinderen op 36 woorden getest. Deze 36 woorden zijn onder te verdelen in 6 woordsoorten zodat van iedere woordsoort 6 woorden zijn getest. De 6 woordsoorten waren de volgende: telbare zelfstandige naamwoorden (*count nouns*), ontelbare zelfstandige naamwoorden (*mass nouns*), overgankelijke werkwoorden (*transitive verbs*), onovergankelijke werkwoorden (*intransitive verbs*), bijvoeglijke naamwoorden (*adjectives*) en bijwoorden (*adverbs)*.

In het Nederlands kunnen zelfstandige naamwoorden ondergebracht worden in twee categorieën: telbaar en ontelbaar. Telbare zelfstandige naamwoorden kun je, zoals de naam al zegt, tellen: één kip, twee kippen. Ontelbare zelfstandige naamwoorden kun je niet tellen. Je kunt bijvoorbeeld niet zeggen één zand, twee zand.

In het Nederlands is er bij werkwoorden een verschil tussen overgankelijke en onovergankelijke werkwoorden. Een overgankelijk werkwoord heeft naast een onderwerp ook een of meerdere objecten. Deze objecten kunnen een lijdend voorwerp, meewerkend voorwerp of voorzetsel voorwerp zijn. Bijvoorbeeld: Hij leest een boek. Een onovergankelijk werkwoord heeft geen direct of indirect object: Hij snurkt.

De woorden die ik uiteindelijk hebt gebruikt zijn de volgende:

* Telbare zelfstandige naamwoorden: kip, computer, rugzak, gitaar, spiegel, rivier.
* Ontelbare zelfstandige naamwoorden: zand, melk, suiker, liefde, water, lucht.
* Overgankelijke werkwoorden: drinken, bouwen, spelen, lezen, schrijven, zoeken.
* Onovergankelijke werkwoorden: snurken, huilen, groeien, springen, schaken, denken.
* Bijvoeglijke naamwoorden: rood, klein, moeilijk, zacht, donker, warm.
* Bijwoorden: gisteren, wanneer, vanmiddag, altijd, enorm, nergens.

Deze woorden zijn grotendeels gekozen uit een lijst van 1000 meest voorkomende Nederlandse woorden. De rest is gekozen als onderdeel van een project van Dr. S. Zuckerman dat onderzoek doet naar een nieuwe manier om het begrijpen van taal te onderzoeken.

## 3.3 Aanpassing na pilottest

Na het testen van drie kinderen heb ik besloten om de taak aan te passen. De kinderen bleken erg veel moeite te hebben met de bijwoorden. Ze wisten er eigenlijk geen reactie op te geven. Ik heb daarom besloten om de bijwoorden uit de taak te laten.

Verder bleken de kinderen ook veel moeite te hebben met het geven van in totaal 3 reacties. Om 3 reacties te kunnen geven moet het kind het stimuluswoord actief houden in het werkgeheugen. Dit blijkt dus nog moeilijk te zijn voor jonge kinderen. Omdat de kinderen hier zo veel moeite mee hadden duurde de test erg lang. Daarom heb ik besloten om maar om één reactie te vragen. Ook heb ik het totaal aantal woorden verminderd. Bij alle woordsoorten heb ik er twee woorden uit gehaald.

Na deze aanpassing bestond de taak nog uit 20 woorden en de volgende woordsoorten: telbare zelfstandige naamwoorden, ontelbare zelfstandige naamwoorden, overgankelijke werkwoorden, onovergankelijke werkwoorden en bijvoeglijke naamwoorden. Van iedere woordsoort zijn 4 woorden getest. De volgende woorden zijn getest:

* Telbare zelfstandige naamwoorden: kip, computer, rugzak, gitaar.
* Ontelbare zelfstandige naamwoorden: zand, melk, suiker, liefde.
* Overgankelijke werkwoorden: drinken, bouwen, spelen, lezen.
* Onovergankelijke werkwoorden: snurken, huilen, groeien, springen.
* Bijvoeglijke naamwoorden: rood, klein, moeilijk, zacht.

## 3.4 Design

Dit experiment maakt gebruik van een between-groups design. De resultaten van de kinderen worden namelijk vergeleken met de resultaten van de volwassenen uit het Small World of Words project. Uit het project worden dezelfde woorden gepakt als degene die voor het experiment van de kinderen zijn gebruikt.

## 3.5 Procedure

Het onderzoek heeft plaats gevonden op twee dagen. Deze twee dagen zijn van te voren afgesproken met de schoolleiding. Dit omdat er een onderbreking plaats heeft gevonden van de normale les. Alleen de kinderen waarvan de ouders een toestemmingsformulier hadden ondertekent werden onderzocht. De leraren hadden voor mij de formulieren verzameld. Deze formulieren heb ik aan het begin van het experiment in beslag genomen.

Ik werd aan het begin van het onderzoek voorgesteld aan de klas door de leraar. In het kort heb ik uitgelegd wie ik was en dat ik een onderzoek kwam doen. Daarna werd het experiment gedaan op de gang of in een ongebruikt lokaal. De kinderen gingen een voor een met mij mee. Na het experiment bracht ik het kind terug naar de klas en nam ik de volgende mee. Aan het begin van het experiment vertelde ik het kind dat we een spelletje gingen doen genaamd ‘zeg een woord’. Vervolgens legde ik uit wat de bedoeling was. Ik zou een woord noemen en vervolgens moest het kind een woord noemen die als eerste bij het kind opkwam. Dit mocht alleen niet hetzelfde woord zijn als het woord wat ik had gezegd. Voor dat we aan het daadwerkelijke experiment begonnen heb ik eerst een korte oefening gedaan om te kijken of het kind de taak begreep. Bij de oefening vroeg ik het kind wat het eerste was waar hij/zij aan dacht als ik het woord ‘appel’ zei. Vervolgens vroeg ik hetzelfde bij het woord ‘boom’. Als bleek dat het kind het hierna begreep ging ik door naar het daadwerkelijke experiment. De reacties op deze oefeningen werden niet meegenomen in de resultaten. Het kwam wel eens voor dat kinderen voorwerpen uit de omgeving opnoemden als reactie op de stimulus (die gebeurde vooral bij de jonge kinderen van 4 jaar). Wanneer ze dit deden herhaalde ik nog een keer dat ze het eerste ding moesten noemen waaraan ze dachten en niet iets uit de omgeving. Hierna ging het wisselend wel of niet goed. Als een kind moeite had met het geven van een associatie stelde ik voor dat we alvast verder ging naar een volgende woord en we aan het eind zouden terugkeren naar het woord waar het kind moeite mee had. Wanneer het kind op alle 20 woorden een antwoord had gegeven kreeg het als beloning een sticker en werd het terug gestuurd naar de klas. Halverwege het experiment gaf ik vaak aan dat het kind goed bezig was om hem/haar gemotiveerd te houden.

## 3.6 Metingen

De metingen in dit experiment waren de associaties van de kinderen op het stimulus woord wat ik heb genoemd. Ieder kind heeft 1 associatie gegeven. Deze heb ik vervolgens allemaal in een Excel document genoteerd. Wat ik verder heb genoteerd is de leeftijd en het geslacht van het kind.

Ook zal ik voor dezelfde woorden als bij de kinderen opzoeken welke associaties mensen als antwoord hebben gegeven bij het A Small World of Words project. Vanuit dit project heb ik alle associaties meegenomen die als reactie gegeven zijn op de stimuluswoorden.

## 3.7 A Small World of Words project

Ik heb de data van het A Small World of Words project gedownload. Uit deze data heb ik, met behulp van een programma van Jakub Dotlacil, de woorden gepakt die overeenkomen met de woorden die ik heb gebruikt voor de kinderen. Uit het overzicht bleek dat er op ieder stimuluswoord 100 reacties waren gegeven. Dit hoefden echter niet dezelfde 100 volwassenen te zijn. Iedere deelnemer kreeg namelijk maar 18 woorden in het experiment.

# 4. Resultaten

In de resultaten zal ik aan de hand van de deelvragen stap voor stap het resultaat uitleggen. Eerst leg ik kort uit hoe ik de resultaten hebt verwerkt. Vervolgens komen de deelvragen aan bod in deze volgorde: woordsoort, woordcombinatie, woord valentie waarde, Age of Acquisition, woordfrequentie diversiteit en tot slot nog enkele bijzondere bevindingen.

## 4.1 Algemeen

De reacties van de kinderen heb ik genoteerd in een excl. bestand. Dit is echter nogal onoverzichtelijk om resultaten uit af te lezen. Met behulp van Jakub Dotlacil, die een algoritme heeft gemaakt in R, heb ik de resultaten overzichtelijker kunnen maken. Dit is ook gedaan voor de resultaten van het A Small World of Words project. In het overzichtelijke bestand is duidelijk te zien hoe vaak een bepaalde reactie voorkomt bij ieder stimuluswoord.

Alle kinderen hebben de hele test afgerond. Op ieder stimuluswoord hebben de kinderen één reactie gegeven. Hier was echter een uitzondering op die ik later in deze sectie zal bespreken.

In de appendix C is de top 3 reacties van de kinderen en de volwassenen te zien voor ieder stimuluswoord. De daadwerkelijke aantallen en de percentages zijn te zien. Ook wordt er bij ieder stimuluswoord en bij alle reacties de woordsoort gegeven.

In het algemeen bleek dat kinderen en volwassenen een overlap in reacties hadden van 37%. Dit percentage is gebaseerd op de top 3 reacties van de kinderen en volwassen. Ook bleek dat kinderen meer divers waren in hun reacties dan de volwassenen. Verdere overeenkomsten en verschillen worden in onderstaande secties toegelicht beginnend met woordsoort.

## 4.2 Woordsoort

Voor dit experiment heb ik me voor de stimuluswoorden beperkt tot de volgende woordsoorten: telbaar zelfstandig naamwoord, ontelbaar zelfstandig naamwoord, overgankelijk werkwoord, onovergankelijk werkwoord en bijvoeglijknaamwoord. Bij de reacties kwamen deze vijf woordsoorten ook voor, evenals het bijwoord.

Om de kinderen met de volwassenen te vergelijken heb ik gekeken naar de 3 meest voorkomende reacties van zowel de kinderen als de volwassenen. In tabel 1 is een overzicht te zien van het aantal kinderen en volwassenen dat van woordsoort is gewisseld bij het geven van zijn/haar reactie. Hiervan zijn de daadwerkelijke aantallen en de percentages gegeven. Deze percentages en aantallen heb ik op de volgende manier bepaald. Ik heb als totaal aantal kinderen (100%) de kinderen genomen die een reactie hadden binnen de top 3 reacties. Dit heb ik gedaan voor ieder stimuluswoord. Zo waren er bv. op het woord ‘bouwen’ in totaal 22 reacties die in de top 3 hoorden bij de kinderen. Bij 19 van deze reacties was er een wisseling van woordsoort. Daarom staat er een percentage van 86% bij het woord ‘bouwen’ (19/22 \* 100%). Bij de volwassenen waren er in totaal 80 reacties die bij de top 3 hoorden. Bij alle 80 reacties was er een wisseling van woordsoort. Daarom staat er een percentage van 100% bij het woord ‘bouwen’ (80/80 \* 100%). Ook wordt er in deze tabel op een ‘strenge’ manier naar de woordsoorten gekeken, vergelijkbaar met heet onderzoek van Brown en Berko (1960). Hiermee bedoel ik dat ik verschil maak tussen een telbaar en ontelbaar zelfstandig naamwoord en tussen een overgankelijk en onovergankelijk werkwoord.

Het gemiddelde percentage wisselen bij de kinderen was 70% en de stdv 0.31. Bij de volwassenen was het gemiddelde percentage wisselen 72% en de stdv 0,32. Een paired samples t-test was uitgevoerd in SPSS om de hoeveelheid van het wisselen van woordsoort van volwassenen en kinderen te vergelijken. Er was geen significant verschil in woordsoort wisselen tussen kinderen en volwassenen SD = 0.274, t(19) = -0.237, p = 0.815.

Er is echter wel een significante correlatie tussen volwassenen en kinderen voor de strenge woordsoorten, Pearson Correlation = 0.633, p = 0.003. Dit betekent dat er een verband is tussen de kinderen en volwassenen op het gebied van wisselen van woordsoort wat niet berust op toeval.

Vervolgens heb ik ook nog op een minder ‘strenge’ manier naar woordsoort gekeken. Hierbij maak ik alleen verschil tussen zelfstandig naamwoord, werkwoord en bijvoeglijk naamwoord. Hierbij heb ik de aantallen en percentages op dezelfde manier bepaald als bij de strenge manier van het kijken naar woordsoort. In tabel 2 is een overzicht te zien van het aantal kinderen en volwassenen dat van woordsoort is gewisseld bij het geven van zijn/haar reactie. Te zien zijn de daadwerkelijke aantallen en de percentages.

Het gemiddelde percentage wisselen bij de kinderen was 55% en de stdv was 0.36. Bij de volwassenen was het gemiddelde percentage wisselen ook 55% en de stdv was 0.39. Een paired samples t-test was uitgevoerd in SPSS om de hoeveelheid van het wisselen van woordsoort van volwassenen en kinderen te vergelijken. Er was geen significant verschil in woordsoort wisselen tussen kinderen en volwassenen SD = 0.062, t(19) = -0.080, p = 0.937. Hierbij heb ik een minder ‘strenge’ manier naar de woordsoorten gekeken. Hiermee bedoel ik dat ik alleen verschil maakten tussen zelfstandige naamwoorden, werkwoorden en bijvoeglijke naamwoorden.

Er is echter wel een significante correlatie tussen kinderen en volwassenen voor de minder strenge woordsoorten, Pearson Correlation = 0.727, p = < 0.001. Dit betekent dat er een verband is tussen de volwassenen en kinderen op het gebied van wisselen van woordsoort wat niet berust op toeval.

Tabel : Het percentage en de daadwerkelijk aantallen bij het wisselen van woordsoort van zowel de kinderen als de volwassenen. Hier wordt op de strenge manier naar de woordsoort gekeken.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Stimulus | Aantal Kinderen | %Gewisseld Kinderen | Aantal Volwassenen | %Gewisseld Volwassenen |
| Bouwen | 19/22 | 86% | 80/80 | 100% |
| Computer | 6/17 | 35% | 13/34 | 38% |
| Drinken | 20/20 | 100% | 57/57 | 100% |
| Gitaar | 23/27 | 85% | 49/69 | 71% |
| Groeien | 15/19 | 79% | 33/33 | 100% |
| Huilen | 18/18 | 100% | 64/64 | 100% |
| Kip | 0/26 | 0% | 6/41 | 15% |
| Klein | 8/19 | 42% | 16/52 | 31% |
| Lezen | 18/22 | 82% | 76/80 | 95% |
| Liefde | 18/18 | 100% | 27/32 | 85% |
| Melk | 27 | 89% | 66/69 | 95% |
| Moeilijk | 20 | 30% | 14/36 | 39% |
| Rood | 20 | 50% | 31/43 | 72% |
| Rugzak | 27 | 30% | 9/30 | 30% |
| Snurken | 34/34 | 100% | 36/42 | 86% |
| Spelen | 8/8 | 100% | 38/38 | 100% |
| Springen | 22/22 | 100% | 31/31 | 100% |
| Suiker | 13 | 46% | 50/56 | 89% |
| Zacht | 24 | 50% | 22/27 | 82% |
| Zand | 22/22 | 100% | 4/76 | 5% |
| Gemiddeldes | **-** | **70%** | **-** | **72%** |
| STDV | **-** | **0,31** | **-** | **0,32** |

Tabel : Het percentage en de daadwerkelijke aantallen bij het wisselen van woordsoort van zowel de kinderen als de volwassenen. Hier wordt op de minder strenge manier naar de woordsoort gekeken.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Stimulus | Aantal Kinderen | %Gewisseld Kinderen | Aantal Volwassene | %Gewisseld Volwassenen |
| Bouwen | 19/22 | 86% | 80/80 | 100% |
| Computer | 6/17 | 35% | 0/34 | 0% |
| Drinken | 20/20 | 100% | 57/57 | 100% |
| Gitaar | 5/27 | 18% | 0/69 | 0% |
| Groeien | 15/19 | 79% | 33/33 | 100% |
| Huilen | 18/18 | 100% | 38/64 | 59% |
| Kip | 0/26 | 0% | 0/41 | 0% |
| Klein | 8/19 | 42% | 16/52 | 31% |
| Lezen | 18/22 | 82% | 76/80 | 95% |
| Liefde | 18/18 | 100% | 14/32 | 44% |
| Melk | 14/27 | 52% | 30/69 | 43% |
| Moeilijk | 6/20 | 30% | 14/36 | 39% |
| Rood | 10/20 | 50% | 31/43 | 72% |
| Rugzak | 0/27 | 0% | 9/30 | 30% |
| Snurken | 7/34 | 21% | 7/42 | 17% |
| Spelen | 8/8 | 100% | 38/38 | 100% |
| Springen | 22/22 | 100% | 31/31 | 100% |
| Suiker | 3/13 | 23% | 50/56 | 89% |
| Zacht | 12/24 | 50% | 22/27 | 82% |
| Zand | 5/22 | 23% | 0/76 | 0% |
| Gemiddeldes | **-** | **55%** | **-** | **55%** |
| STDV | **-** | **0,36** | **-** | **0,39** |

## 4.3 Woordcombinatie

Zoals al eerder gesteld definieer ik een woordcombinatie als een ‘phrase completion’: woorden die over het algemeen het stimuluswoord voorafgaan of volgen. Zowel de kinderen als de volwassenen hadden als reactie op het ‘melk’ het woord ‘drinken’. Dit is een woord wat vaak achter ‘melk’ komt. Vandaar dat ik dit als een woordcombinatie heb geteld. Op het stimuluswoord ‘klein’ werd vaak met ‘groot’ gereageerd. Deze twee woorden volgen elkaar echter niet vaak op. Dit zie ik dus niet als een woordcombinatie.

Om de kinderen met de volwassenen te vergelijken heb ik gekeken naar de 3 meest voorkomende reacties van zowel de kinderen als volwassenen. In tabel 3 is een overzicht te zien van het aantal woordcombinaties bij kinderen en volwassenen per stimuluswoord. De daadwerkelijke aantallen en het percentage is gegeven. Deze aantallen heb ik bepaald op een vergelijkbare manier als bij woordsoort. Ik heb als totaal aantal kinderen (100%), de kinderen genomen die een reactie hadden binnen de top 3 reacties. Dit heb ik gedaan voor ieder stimuluswoord. Zo waren er bv. op het woord ‘computer’ in totaal 17 reacties die in de top 3 hoorden bij de kinderen. 11 van deze reacties vormden een combinatie met het stimuluswoord (dit waren bv. computer-scherm en computer-spelletjes). Daarom staat er een percentage van 65% bij het woord ‘computer’ (11/17 \* 100%). Bij de volwassenen waren er in totaal 34 reacties die bij de top 3 hoorden. 12 van deze reacties vormden een combinatie met het stimuluswoord. Daarom staat er een percentage van 35% bij het woord ‘computer’ (12/34 \* 100%).

Het gemiddelde percentage woordcombinaties bij kinderen was 31% en de stdv was 0.37. Bij de volwassenen bij het gemiddelde percentage woordcombinaties 25% en de stdv was 0.33. Een paired samples t-test was uitgevoerd in SPSS om de hoeveelheid woordcombinaties van kinderen en volwassenen met elkaar te vergelijken aan de hand van het stimuluswoord. Er was geen significant verschil in de hoeveelheid woordcombinaties van kinderen en volwassenen SD = 0.229, t(19) = 1.210, p = 0.241.

Er was wel een significante correlatie tussen kinderen en volwassenen kijkend naar de hoeveelheid woordcombinaties, Pearson Correlation = 0.790, p = < 0.001. Dit bekent dat er een verband is tussen kinderen en volwassenen op het gebied van woordcombinaties dat niet berust op toeval.

Tabel : Het percentage en de daadwerkelijke aantallen van de hoeveelheid combinaties bij kinderen en volwassenen.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Stimulus | Aantal Kinderen | %Combinatie Kinderen | Aantal Volwassenen | %Combinatie Volwassenen |
| Bouwen | 19/22 | 86% | 71/80 | 89% |
| Computer | 11/17 | 65% | 12/34 | 35% |
| Drinken | 16/20 | 80% | 42/57 | 74% |
| Gitaar | 5/27 | 18% | 20/69 | 29% |
| Groeien | 6/19 | 32% | 0/33 | 0% |
| Huilen | 0/18 | 0% | 0/64 | 0% |
| Kip | 0/26 | 0% | 6/41 | 15% |
| Klein | 0/19 | 0% | 0/52 | 0% |
| Lezen | 18/22 | 82% | 76/80 | 95% |
| Liefde | 0/18 | 0% | 0/32 | 0% |
| Melk | 24/27 | 89% | 39/69 | 57% |
| Moeilijk | 0/20 | 0% | 0/36 | 0% |
| Rood | 0/20 | 0% | 19/44 | 44% |
| Rugzak | 0/27 | 0% | 0/30 | 0% |
| Snurken | 0/34 | 0% | 0/42 | 0% |
| Spelen | 5/8 | 62% | 0/38 | 0% |
| Springen | 18/22 | 82% | 16/31 | 52% |
| Suiker | 3/13 | 23% | 0/56 | 0% |
| Zacht | 0/24 | 0% | 0/27 | 0% |
| Zand | 0/22 | 0% | 4/76 | 5% |
| Gemiddeldes | **-** | **31%** | **-** | **25%** |
| STDV | **-** | **0,37** | **-** | **0,33** |

## 4.4 Woord valentie waarde

Voor dit onderzoek ben ik uitgegaan van de woord valentie waarden van de universiteit van Gent, die voor 4300 Nederlands woorden de valentie waarde heeft bepaald (Moors et al., 2013).

Om de kinderen met de volwassenen te vergelijken heb ik gekeken naar de meest voorkomende reactie voor zowel de kinderen als de volwassenen. De stimuluswoorden ‘computer’ en ‘rugzak’ kwamen niet voor in de lijst van de universiteit van Gent. Deze woorden heb ik daarom niet meegenomen voor dit deel van de resultaten. Wanneer de eerste reactie niet in de lijst stond, heb ik de eerstvolgende reactie genomen die wel in de lijst stond. Ook stond er in de lijst soms alleen de enkelvoud vorm van een woord terwijl de reactie of stimulus meervoud was. In dat geval ben ik uitgegaan van de valentie van de enkelvoud vorm van het woord. In tabel 4 is een overzicht te zien die weergeeft wat de valentiewaarden zijn van zowel het stimuluswoord als de reacties van de kinderen en de volwassenen.

De gemiddelde valentie waarde van de kinderen was 4.70 en de stdv was 0.97. Bij de volwassenen was de gemiddelde valentie waard 4.60 en de stdv was 0.98. Een Pearson Correlation test was uitgevoerd om de woord valentie waarden van de stimuluswoorden te vergelijken met de valentie waarden van de reacties van de kinderen. Er was een significante correlatie tussen de stimuluswoorden en de reacties van de kinderen betreft valentie waarde, Pearson Correlation = 0.470, p = 0.049. Dit betekent dat er tussen de valentiewaarden van de stimuluswoorden en die van de reacties van de kinderen een verband is dat niet berust op toeval.

Ook is er een Pearson Correlation test uitgevoerd om de woord valentie waarden van de stimuluswoorden te vergelijken met de valentie waarden van de reacties van de volwassenen. Hier was net geen significante correlatie aanwezig, Pearson Correlation = 0.456, p = 0.057.

Tabel : De woord valentie waarde van het stimuluswoord en de meest voorkomende reactie van de kinderen en de volwassenen.

|  |  |  |  |
| --- | --- | --- | --- |
| Stimulus | Valentie Stimulus | Valentie Meest Voorkomende Reactie Kinderen | Valentie Meest Voorkomende Reactie Volwassenen |
| Bouwen | 4,78 | 3,52 | 4,94 |
| Drinken | 4,78 | 4,95 | 4,53 |
| Gitaar | 5,22 | 5,75 | 5,75 |
| Groeien | 5,31 | 4,78 | 5,05 |
| Huilen | 2,38 | 2,17 | 2,55 |
| Kip | 4,39 | 4,44 | 4,44 |
| Klein | 3,70 | 4,78 | 4,78 |
| Lezen | 4,81 | 4,89 | 4,89 |
| Liefde | 6,53 | 6,38 | 4,39 |
| Melk | 4,63 | 4,78 | 4,31 |
| Moeilijk | 2,80 | 4,59 | 3,09 |
| Rood | 4,39 | 4,89 | 2,83 |
| Snurken | 2,59 | 5,50 | 5,50 |
| Spelen | 5,41 | 5,94 | 5,05 |
| Springen | 4,48 | 4,34 | 4,02 |
| Suiker | 4,66 | 5,27 | 5,58 |
| Zacht | 5,48 | 3,64 | 5,89 |
| Zand | 4,41 | 4,00 | 5,23 |
| Gemiddeldes | **4,49** | **4,70** | **4,60** |
| STDV | **1,06** | **0,98** | **0,97** |

## 4.5 Age of Acquisition

In dit onderzoek ben ik voor de Age of Acquistion waarden uitgegaan van het onderzoek van Brysbaert et al. (2014), die van 30.000 Nederlandse woorden de Age of Acquisition norm hebben onderzocht.

Om de kinderen met de volwassenen te vergelijken heb ik gekeken naar de meest voorkomende reactie voor zowel de kinderen als de volwassenen. Niet alle stimuluswoorden kwamen voor in de lijst van Brysbaert et al., maar over het algemeen was er wel een woord wat er sterk op leek. Zo kwam het woorden ‘blokken’ niet voor. In plaats daar van heb ik het woord ‘blok’ genomen. In tabel 5 is een overzicht te zien die de AoA van het stimuluswoord en de reacties van de kinderen en volwassenen weergeeft.

De gemiddelde Age of Acquisition van het stimuluswoord is 5.14 en de stdv 1.25. Voor de kinderen was de gemiddelde Age of Acquisition 4.85 en de stdv was 1.03. Bij de volwassenen was de gemiddelde Age of Acquisition 5.30 en de stdv 1.71. Er was een Pearson Correlation test uitgevoerd om de AoA waarden van de stimuluswoorden te vergelijken met de AoA waarden van de reacties van de kinderen. Er was een significante correlatie tussen de AoA waarden van de stimuluswoorden en de reacties van de kinderen, Pearson Correlation = 0.473, p = 0.035. Dit wil zeggen dat er een verband is tussen de AoA waarden van de stimuluswoorden en de AoA van de reacties van de kinderen dat niet op toeval berust.

Ook was er een Pearson Correlation test uitgevoerd om de AoA waarden van de stimuluswoorden te vergelijken met de AoA waarden van de reacties van de volwassenen. Er was geen significante correlatie tussen de AoA waarden van de volwassenen en die van stimuluswoorden, Pearson Correlation = 0.375, p = 0.103. Er was echter wel een significante correlatie gevonden wanneer ik 100 woorden nam, te zien in de sectie Achtergrond Literatuur. Een mogelijke verklaring hiervoor zal ik bespreken in de discussie

Tabel : De AoA van het stimuluswoord en de meest voorkomende reactie van de kinderen en de volwassenen.

|  |  |  |  |
| --- | --- | --- | --- |
| Stimulus | AoA Stimulus | AoA Reactie Kinderen | AoA Reactie Volwassenen |
| Bouwen | 5,53 | 5,15 | 4,24 |
| Computer | 8,51 | 6,78 | 9,7 |
| Drinken | 3,56 | 3,67 | 7,38 |
| Gitaar | 6,26 | 4,84 | 4,84 |
| Groeien | 5,09 | 4,28 | 4,17 |
| Huilen | 4,4 | 5,4 | 4,15 |
| Kip | 4,37 | 4,75 | 4,49 |
| Klein | 4,03 | 4,28 | 4,28 |
| Lezen | 5,65 | 4,55 | 4,55 |
| Liefde | 7,09 | 6,03 | 3,91 |
| Melk | 3,69 | 3,56 | 4,17 |
| Moeilijk | 5,96 | 6,22 | 9,15 |
| Rood | 3,91 | 4,4 | 5,01 |
| Rugzak | 5,83 | 3,61 | 5,65 |
| Snurken | 5,96 | 3,57 | 3,57 |
| Spelen | 3,9 | 4,34 | 4,17 |
| Springen | 4,59 | 6,94 | 6,94 |
| Suiker | 5,33 | 4,19 | 5,9 |
| Zacht | 4,65 | 5,55 | 5,04 |
| Zand | 4,41 | 4,88 | 4,78 |
| Gemiddeldes | **5,14** | **4,85** | **5,30** |
| STDV | **1,25** | **1,03** | **1,71** |

## 4.6 Frequentie

Voor dit onderzoek ben ik qua woord frequentie waarden uitgegaan van het onderzoek van Keuleers et al. (2010). Bij dit onderzoek hoort een database van de frequentie van Nederlandse woorden gebaseerd op 44 miljoen woorden van de ondertiteling van films en series.

Om de kinderen met de volwassenen te vergelijken heb ik naar de meest voorkomende reactie gekeken voor zowel de kinderen als de volwassenen. In de database van Keuleers et al. (2010) kwamen de stimuluswoorden ‘ei’/ ‘eieren’ en ‘groot’ niet voor. Deze heb ik voor dit deel van de resultaten dan ook niet meegenomen. In tabel 6 is een overzicht te zien van de frequentie waarden van de stimuluswoorden en de meest voorkomende reacties van de volwassenen en de kinderen.

De gemiddelde frequentie bij de kinderen was 147,55 miljoen en de stdv was 147,37 miljoen. Bij de volwassenen was de gemiddelde frequentie 160,69 miljoen en de stdv 218,99 miljoen. Een Pearson Correlation test was uitgevoerd om de frequentie waarden van de stimuluswoorden te vergelijken met de frequentie waarden van de reacties van de kinderen. Er was een significante correlatie tussen de frequentie waarden van de stimuluswoorden en de reacties van de kinderen, Pearson Correlation = 0.600, p = 0.014. Dit wil zeggen dat er een verband is tussen de frequentie waarden van de stimuluswoorden en de frequentie waarden van kinderen dat niet op toeval berust.

Ook was er een Pearson Correlation test uitgevoerd om de frequentie waarden van de stimuluswoorden te vergelijken met de frequenties waarden van de reacties van de volwassenen. Er is geen significant correlatie tussen de frequentie waarden van de stimuluswoorden en van de reacties van de volwassenen, Pearson Correlation = 0.028, p = 0.918.

Tabel : De frequentie waarden in miljoenen van de stimuluswoorden en de reacties van kinderen en volwassenen.

|  |  |  |  |
| --- | --- | --- | --- |
| Stimulus | Frequentie Stimulus (miljoenen) | Frequentie Reactie Kinderen (miljoenen) | Frequentie Reactie Volwassenen (miljoenen) |
| Bouwen | 44,20 | 6,22 | 818,90 |
| Drinken | 193,67 | 244,50 | 53,35 |
| Gitaar | 11,46 | 107,46 | 107,46 |
| Groeien | 25,36 | 237,51 | 333,30 |
| Huilen | 54,52 | 17,84 | 9,33 |
| Lezen | 107,80 | 150,93 | 150,93 |
| Liefde | 208,90 | 107,16 | 48,59 |
| Melk | 39,70 | 193,67 | 18,55 |
| Moeilijk | 207,21 | 211,23 | 10,95 |
| Rood | 48,59 | 39,68 | 185,32 |
| Snurken | 1,35 | 209,22 | 209,22 |
| Spelen | 247,77 | 595,57 | 474,49 |
| Springen | 46,86 | 63,23 | 26,25 |
| Suiker | 21,85 | 12,55 | 10,98 |
| Zacht | 22,85 | 159,46 | 45,62 |
| Zand | 19,99 | 4,62 | 67,80 |
| Gemiddeldes | **81,85** | **147,55** | **160,69** |
| STDV | **83,41** | **147,37** | **218,99** |

## 4.7 Diversiteit

Met diversiteit bedoel ik hoeveel verschillende reacties er in totaal zijn gegeven op een bepaald stimulus woord. Wanneer er bijvoorbeeld veel verschillende de reacties zijn gegeven op een stimuluswoord is de diversiteit hoog. Wanneer er veel dezelfde reacties gegeven worden is de diversiteit laag.

Om de kinderen met de volwassenen te vergelijken heb ik gekeken naar de hoeveelheid verschillende reacties die gegeven zijn op een stimuluswoord. In tabel 7 is een overzicht te zien die weergeeft hoeveel diversiteit er in de reacties was van zowel volwassenen en kinderen per stimuluswoord. Ik heb de percentages berekent door het aantal verschillende reacties te delen door het totaal aantal reacties. Zo hadden de kinderen op het woord ‘bouwen’ in totaal 46 reacties gegeven. Hierbij waren er 23 verschillende reacties. Dus de diversiteit voor het woord ‘bouwen’ bij de kinderen is 50% (23/46 \* 100%). De volwassenen hadden in totaal 100 reacties op het woord ‘bouwen’. Hierbij waren er 18 verschillende reacties (heel veel volwassenen hadden de reactie ‘huis’). Dus de diversiteit van de volwassenen voor het woord ‘bouwen’ is 18% (18/100 \* 100%).

De gemiddelde diversiteit bij de kinderen was 51% en de stdv was 0.11. Bij de volwassenen was de gemiddelde diversiteit 38% en de stdv was 0.12. Een paired samples t-test was uitgevoerd om de diversiteit van de volwassenen te vergelijken met de diversiteit van de kinderen aan de hand van het stimuluswoord. Er was een significant verschil in diversiteit tussen kinderen en volwassenen SD = 0.037, t(19) = 3.457, p = 0,003. De kinderen waren meer divers dan de volwassenen.

Tabel : De percentages van de diversiteit van de reacties van de kinderen en de volwassen per stimuluswoord.

|  |  |  |
| --- | --- | --- |
| Stimulus | Diversiteit Kinderen | Diversiteit Volwassenen |
| Bouwen | 50% | 18% |
| Computer | 54% | 51% |
| Drinken | 48% | 27% |
| Gitaar | 37% | 29% |
| Groeien | 52% | 62% |
| Huilen | 54% | 25% |
| Kip | 37% | 45% |
| Klein | 54% | 31% |
| Lezen | 50% | 20% |
| Liefde | 54% | 51% |
| Melk | 43% | 28% |
| Moeilijk | 59% | 44% |
| Rood | 57% | 34% |
| Rugzak | 43% | 44% |
| Snurken | 28% | 42% |
| Spelen | 83% | 42% |
| Springen | 50% | 53% |
| Suiker | 61% | 34% |
| Zacht | 46% | 53% |
| Zand | 50% | 23% |
| Gemiddeldes | **51%** | **38%** |
| STDV | **0,11** | **0,12** |

## 4.8 Overlap

Met overlap bedoel ik hoe vaak kinderen en volwassenen dezelfde reactie hebben gegeven op een stimuluswoord. Hiervoor heb ik gekeken naar de 3 meest voorkomende reacties. Hieruit blijkt dat kinderen en volwassenen een overlap van 37% hebben op hun woordassociaties.

In tabel 8 is het percentage van overlap tussen de reacties van kinderen en volwassenen per stimuluswoord te zien.

Tabel : De percentages van de hoeveelheid overlap in de reacties van de kinderen en volwassenen per stimuluswoord.

|  |  |
| --- | --- |
| Stimulus | %Overlap |
| Bouwen | 33% |
| Computer | 33% |
| Drinken | 33% |
| Gitaar | 33% |
| Groeien | 33% |
| Huilen | 33% |
| Kip | 67% |
| Klein | 67% |
| Lezen | 100% |
| Liefde | 0% |
| Melk | 33% |
| Moeilijk | 33% |
| Rood | 33% |
| Rugzak | 0% |
| Snurken | 33% |
| Spelen | 0% |
| Springen | 33% |
| Suiker | 67% |
| Zacht | 67% |
| Zand | 0% |
| Gemiddelde | **37%** |
| STDV | **0,26** |

## 4.9 Korte Samenvatting

De resultaten laten zien dat kinderen en volwassenen verschillen op het gebied van diversiteit, kinderen waren meer divers in hun reacties dan de volwassenen. Er was een overlap tussen de twee groepen van 37%. Op het gebied van woordsoort en woordcombinatie bleek dat de kinderen en de volwassenen niet van elkaar verschillen. Er was zelfs een verband gevonden tussen beiden. Hierbij maakte het niet uit over er op de strenge of minder strenge manier naar woordsoort werd gekeken. Op het gebied van woord valentie waarden bleken kinderen een correlatie te hebben met de waarden van de stimuluswoorden, voor de volwassen was er net geen significante correlatie. Voor de Age of Acquisition gold hetzelfde, bij de kinderen was er een correlatie met de AoA van de stimuluswoorden en bij de volwassenen niet. Bij de woordfrequentie waarde was het ook hetzelfde geval. Tussen de stimuluswoorden en de reacties was er een correlatie. Tussen de stimuluswoorden en de reacties van de volwassenen niet.

## 4.10 Bijzondere bevindingen

**Het woord moeilijk**

Zoals ik al eerder stelde hebben de kinderen op ieder stimuluswoord een reactie gegeven. Er was hier echter een uitzondering op en dat was het woord ‘moeilijk’. Bij dit stimuluswoord is het meerdere keren voorgekomen dat een kind er geen associatie bij had. Zelfs na het woord overslaan en later er naar terugkeren wisten kinderen er geen associatie bij te geven. Of de kinderen hier zo veel moeite mee hadden omdat het woord toevallig ‘moeilijk’ was is mij niet helemaal duidelijk. Het was iets wat mij was opgevallen.

**Typische woorden voor kinderen en volwassenen**

Wat ook opviel was dat sommige reacties duidelijk van kinderen of volwassenen afkwamen. Zo was de meest voorkomende reactie op het stimuluswoord ‘drinken’ bij de volwassenen het woord ‘bier’. Een kind tussen de 4 en 8 jaar oud zal niet zo snel aan bier denken bij het horen van het woord drinken, aangezien kinderen van deze leeftijd nog ver van de legale leeftijd van bier drinken zitten.

Bij het stimuluswoord ‘bouwen’ is bij kinderen de populairste reactie het woord ‘blokken’. Dit waarschijnlijk omdat kinderen blokken zien als speelgoed waarmee je kunt bouwen. Aangezien volwassenen niet vaak meer in aanraking komen met speelgoed zal een reactie als ‘blokken’ bij volwassenen niet zo snel voorkomen.

Andere voorbeelden hiervan is de reactie ‘spelletjes’ op het stimuluswoord ‘computer’ bij kinderen en de reactie ‘kinderen’ op het stimuluswoord ‘spelen’ bij de volwassenen. Dit omdat bij kinderen de computer nog vooral bekend staat als een apparaat om spelletjes op te spelen en kinderen vooral bekend staan om bijvoorbeeld buiten te spelen.

# 5. Discussie

In de discussie zal ik allereest een conclusie trekken uit de resultaten. Per deelvraag zal ik kort mijn conclusie toelichten. Ik zal aangeven of mijn opgestelde hypothese juist of onjuist zijn. Dan volgt een korte discussie van de resultaten. Vervolgens bespreek ik het vergelijkingsexperiment met het kunstmatige model. Tot slot benoem ik een aantal beperkingen van het onderzoek en mogelijkheden voor toekomstig onderzoek.

## 5.1 Conclusie

**Woordsoort**

Uit de resultaten valt de conclusie te trekken dat kinderen en volwassenen op het gebied van het wisselen van woordsoort niet van elkaar verschillen. Er was zelfs een correlatie gevonden tussen de hoeveelheid wisselen van de kinderen en de volwassenen. Dit resultaat gold voor zowel de strenge als de mindere strenge manier van het kijken naar de woordsoort. De hypothese die ik had opgesteld was dus onjuist.

Dat er een significante correlatie is gevonden tussen de kinderen en de volwassenen op het gebied van wisselen van woordsoort betekent dat er een verband is tussen kinderen en volwassenen dat niet berust op toeval. Qua wisselen van woordsoort zijn kinderen en volwassenen dus vergelijkbaar, ondanks het verschil in leeftijd en de verdere ontwikkeling van taal bij volwassenen.

Dit resultaat is dus tegenstrijdig met de resultaten van het onderzoek van Brown en Berko (1960), die stellen dat kinderen en volwassenen juist wel verschillen in het wisselen van woordsoort. Volgens hen zouden volwassenen over het algemeen bij dezelfde woordsoort blijven terwijl kinderen veel van woordsoort wisselen bij het geven van een associatie.

**Woordcombinatie**

Uit de resultaten valt de conclusie te trekken dat kinderen en volwassen op het gebied van woordcombinaties niet van elkaar verschillen. Er was zelfs een correlatie gevonden tussen de hoeveelheid woordcombinatie bij de kinderen en de hoeveelheid bij de volwassenen. Kinderen en volwassenen hebben dus gemiddeld evenveel woordcombinaties bij het geven van associaties. De hypothese die ik had opgesteld was dus onjuist.

Dat er een significante correlatie is gevonden tussen de kinderen en de volwassenen op het gebied van woordcombinaties betekent dat er een verband is tussen kinderen en volwassenen dat niet berust op toeval. Qua woordcombinaties zijn kinderen en volwassenen dus vergelijkbaar, ondanks het verschil in leeftijd en verdere ontwikkeling van taal bij volwassenen.

Ook dit resultaat is tegenstrijdig met de resultaten van het onderzoek van Brown en Berko (1960), die stellen dat kinderen en volwassenen juist wel verschillen in hoeveelheid woordcombinaties. Volgens hen zouden kinderen meer woordcombinaties (lees phrase completions) maken dan volwassenen.

**Woord valentie waarde**

Uit de resultaten valt de conclusie te trekken dat kinderen en volwassenen op het gebied van woord valentie waarde niet van elkaar verschillen. Voor de kinderen was er een significantie relatie en voor de volwassenen net niet. Deze twee waarden liggen echter zo dicht bij elkaar dat ik de conclusie trek dat er geen verschil is tussen kinderen en volwassenen. Om het zeker te weten heb ik nog een correlatie test uitgevoerd via SPSS op de 100 willekeurige woorden die ik ook voor de AoA en frequentie waarden heb gebruikt (volwassenen). Uit deze test blijkt dat er een significante correlatie is tussen de woord valentie waarde van het stimuluswoord en die van de reactie, Pearson Correlation = 0.652, p < 0.001. De woord valentie waarden kwamen weer uit het onderzoek van Moors et al. (2013). De hypothese die ik had opgesteld was dus juist.

Dat er een significante correlatie is gevonden tussen de kinderen en volwassenen op het gebied van woord valentie waarde betekent dat er een verband is tussen kinderen en volwassenen dat niet berust op toeval. Zowel kinderen als volwassenen lijken een goed begrip van de emotionele waarde van woorden te hebben. Vooral voor jonge kinderen is dit bijzonder aangezien hun taalontwikkeling nog in volle gang is.

Deze correlatie zou aanwezig kunnen zijn door mogelijke clusters in het mentale lexicon, waarin woorden met dezelfde woord valentie waarde sterkere verbindingen hebben met elkaar. Wanneer er dan bv. een stimuluswoord is met een lage valentie waarde, zal die sneller andere woorden met een lage valentie waarde activeren dan woorden met een hogere valentie waarde.

**Age of Acquisition**

Uit de resultaten valt de conclusie te trekken dat kinderen en volwassenen op het gebied van de AoA waarde niet van elkaar verschillen. Voor de kinderen was er een significante correlatie tussen de stimuluswoorden en de reactie, maar voor de volwassenen niet. Dit kan echter liggen aan de kleine hoeveelheid woorden die ik heb gebruikt en dat deze woorden over het algemeen een lage AoA waarde hebben (lage waarde zodat ook de kinderen alle woorden zouden kennen). Wanneer er een grotere database wordt gepakt zoals ik in de Achtergrond Literatuur sectie heb gedaan, blijkt dat er ook voor volwassenen een significante correlatie is tussen de stimuluswoorden en de reacties. Daarom trek ik de conclusie dat zowel kinderen als volwassenen een reactie met een vergelijkbare AoA waarde als het stimuluswoord geven. De hypothese die ik had opgesteld was dus juist.

Dat er een significante correlatie is gevonden bij de kinderen en bij de volwassenen (bij een grotere database) tussen de AoA waarden van het stimuluswoord en de reactie betekent dat er een verband is tussen stimulus- en reactiewoord dat niet berust op toeval. Zowel kinderen als volwassenen lijken een begrip te hebben van de AoA waarde van woorden. Deze correlatie zou aanwezig kunnen zijn door mogelijk clusters in het mentale lexicon, waarin woorden met dezelfde AoA waarde dichter bij elkaar liggen (dus een sterkere verbinding hebben). Wanneer er een stimuluswoord is met een lage AoA waarde, zal die eerst andere woorden met een lage AoA waarde. Dit omdat deze woorden in hetzelfde cluster liggen.

**Frequentie**

Uit de resultaten valt de conclusie te trekken dat kinderen en volwassenen op het gebied van de woordfrequentie waarde van elkaar verschillen. Voor de kinderen was er een significante correlatie tussen de waarde van de stimuluswoorden en de reacties. Voor de volwassenen was dit niet. Ook wanneer er gekeken werd naar een grotere database (Achtergrond Literatuur) is er geen significante correlatie te vinden. De kinderen geven dus over het algemeen een associaties met een frequentie waarde die vergelijkbaar is met de frequentie waarde van het stimuluswoord. De hypothese die ik heb opgesteld was dus onjuist.

Dat er een significante correlatie is gevonden bij de kinderen tussen de frequentie waarden van het stimuluswoord en de reactie betekent dat er een verband is tussen stimulus- en reactiewoord dat niet berust op toeval. Voor volwassenen was deze correlatie niet aanwezig. Dit zou mogelijk kunnen betekenen dat er bij kinderen clusters in het mentale lexicon aanwezig zijn die woorden met vergelijkbare frequentie waarden bij elkaar clusteren. Bij volwassenen zijn deze clusters aan de hand van deze resultaten niet aanwezig. Wanneer er dan bv. een stimuluswoord is met een lage frequentie, zal er bij de kinderen sneller andere woorden met een lage frequentie geactiveerd worden.

**Diversiteit**

Uit de resultaten valt de conclusie te trekken dat kinderen en volwassenen op het gebied van diversiteit van elkaar verschillen. Kinderen zijn meer divers dan volwassenen. De kinderen geven dus meer verschillende associaties bij een stimuluswoord dan volwassenen. De hypothese die ik heb opgesteld is dus onjuist.

Dat er een significant verschil is tussen kinderen en volwassenen betekent dat er een verschil is dat niet berust op toeval. Dit zou verklaart kunnen worden door creatievere denkwijze van kinderen.

**Algemeen**

Uit de verschillende conclusies van de deelvragen valt de algehele conclusie te trekken dat kinderen en volwassenen qua woordassociaties van elkaar verschillen op het gebied van woordfrequentie waarde en diversiteit van de reacties. De hypothese die ik heb opgesteld is dus maar voor een klein deel juist. In de hypothese had ik namelijk gesteld dat er een verschil zou zijn op gebied van diversiteit. Alleen had ik verwacht dat juist de volwassenen meer divers zouden zijn dan de kinderen.

## 5.2 Discussie van de resultaten

Een deel van de resultaten is bijzonder. De resultaten met betrekking tot het wisselen van woordsoort en de hoeveelheid woordcombinaties gaan namelijk in tegen de resultaten van het onderzoek van Brown en Berko (1960). Voor een goede verklaring hiervan zal er toekomstig onderzoek uitgevoerd moeten worden om te kijken of er vergelijkbare resultaten aangetoond kunnen worden.

Ook bijzonder zijn de resultaten met betrekking tot de diversiteit van de reacties. Ondanks dat kinderen veel minder woorden kennen dan volwassenen hebben ze toch een grotere diversiteit in hun associaties. Dit zou dan mogelijk toch te verklaren kunnen zijn door de verschillende denkwijzen van kinderen en volwassenen. Kinderen hebben over het algemeen een divergente denkwijzen, terwijl volwassenen een convergente denkwijze hebben. Door de divergente denkwijze zijn kinderen vaak creatiever dan volwassenen. Hierdoor zou het kunnen dat kinderen meer associaties weten te bedenken bij het horen van een stimuluswoord. Dit zou kunnen betekenen dat bij de diversiteit van het benoemen van woorden een creatieve denkwijze een positievere invloed heeft dan de grootte van de woordenschat. Uit de resultaten blijkt namelijk dat kinderen, die creatiever zijn maar een kleinere woordenschat hebben, meer verschillende reacties hebben dan de volwassenen die juist een grotere woordenschat hebben maar over het algemeen minder creatief zijn.

Er was nog niet eerder onderzoek gedaan naar de relatie tussen de woordfrequentie en AoA waarden van het stimuluswoord en die van de associaties. De resultaten die ik heb gevonden waren dus nieuwe resultaten. Om deze resultaten te kunnen bevestigen zou er toekomstig onderzoek uitgevoerd moeten worden om te kijken of dezelfde resultaten aangetoond kunnen worden.

De resultaten met betrekking tot de woord valentie waarde waren naar verwachting. De resultaten die ik heb gevonden waren vergelijkbaar met voorgaand onderzoek.

Dat er correlaties zijn gevonden bij zowel de kinderen en volwassenen op het gebied van woord valentie waarde en de Age of Acquisition waarde zou meer inzicht kunnen geven in de organisatie van het mentale lexicon. Het lijkt aan de hand van deze resultaten zo te zijn dat woorden o.a. op woord valentie waarde en Age of Acquisition waarde bij elkaar geclusterd worden. Dit geldt voor het mentale lexicon van kinderen en volwassenen. Woorden in een clusters worden eerder geactiveerd dan woorden buiten het cluster. Er lijken geen clusters te zijn die gebaseerd zijn op woordfrequentie waarden, aangezien hier een significant verschil voor is gevonden.

Dit onderzoek heeft een mogelijke bijdrage kunnen leveren aan het beter begrijpen van het mentale lexicon en de mogelijke verschillen hiervan tussen kinderen en volwassenen. Om deze resultaten te bevestigen is er echter vervolgonderzoek nodig om te kijken of dezelfde resultaten aangetoond kunnen worden.

## 5.3 Vergelijking met kunstmatig model

Voor de connectie met kunstmatige intelligentie heb ik gekeken naar de woordassociaties van een kunstmatig model. Dit om te onderzoeken of een kunstmatig model vergelijkbaar is met hoe volwassenen of kinderen reageren. Het kunstmatige model waar ik voor heb gekozen is een online demo (<http://bionlp-www.utu.fi/wv_demo/>) gebaseerd op “word embeddings” afgeleid van de word2vec methode (Mikolov, Chen, Corrado & Dean, 2013). Het model is getraind op Engels Google nieuws. Omdat dit model alleen in het Engels is, heb ik het vergeleken met de resultaten van het onderzoek van Brown en Berko (1960).

In het model heb ik gebruik gemaakt van de Nearest Words tool. Deze tool laat, gegeven een woord, een lijst van andere woorden zien die vergelijkbaar zijn met het gegeven woord. Oftewel woorden die nabijgelegen zijn in de vector space.

Omdat ik de resultaten van het kunstmatig model ga vergelijken met het experiment van Brown en Berko ga ik ook dezelfde stimuluswoorden gebruiken. Het gaat om te volgende stimuluswoorden die zijn onderverdeeld in zes woordsoorten:

Stimuluswoorden

* Count nouns (C.N.): table, house, foot, needle, apple, doctor.
* Mass nouns (M.N.): milk, water, sand, sugar, air, cheese.
* Transitive verb (T.V.): to send, to bring, to find, to take, to hit, to invite.
* Intransitive verb (I.V.): to skate, to come, to live, to laugh, to stand, to walk.
* Adjectives (Adj.): dark, soft, cold, white, sweet, hard.
* Adverbs (Adv.): quickly, slowly, sadly, now, softly, gently.

In het experiment van Brown en Berko werd gewerkt met 4 verschillende groepen die ieder 20 personen bevatte. Groep 1, 2 en 3 waren respectievelijk kinderen uit groep 1, 2 en 3. De vierde groep bestond uit volwassenen en studenten. Aan de participanten werd de eerste reactie gevraagd die ze hadden bij een stimuluswoord. Van alle participanten werd uiteindelijk het gemiddelde genomen. Omdat ik maar één kunstmatig model heb ga ik kijken naar de eerste 3 reacties en daarvan het gemiddelde nemen.

Ik ga hetzelfde score systeem gebruiken als in het experiment van Brown en Berko. Namelijk 1 punt wanneer de reactie homegenous-by-part-of-speech (Hmg.) is. Per woordsoort viel er dus een maximale score van 6 punten te halen per persoon te behalen. Voor de woordsoort C.N. kon je bv. 1 punt krijgen als op ‘table’ reageerde met een C.N. Er waren in totaal 6 C.N. dus vielen er ook 6 punten in totaal per woordsoort te verdienen. Uiteindelijk hebben ze qua punten aantal het gemiddelde genomen van alle participanten per groep.

Bij het kunstmatige model heb ik gekeken naar de eerste 3 reacties. In eerste instantie heb ik het gemiddelde punten aantal van de 3 reacties gepakt en vervolgens dezelfde punten berekening toegepast als bij het onderzoek van Brown en Berko. Weer met een maximale score van 6 punten per woordsoort

In tabel 10 is een overzicht te zien van de score van volwassenen, kinderen uit groep 1, 2 en 3 (uit het onderzoek van Brown en Berko), en het kunstmatige model per woordsoort. In appendix D zijn de reacties van het kunstmatige model te zien.

Tabel : De resultaten van het onderzoek van Brown en Berko (1960) en de resultaten van het kunstmatige model.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Volwassenen | Kinderen Groep 1 | Kinderen Groep 2 | Kinderen Groep 3 | Kunstmatig Model |
| C.N. | 5,10 | 3,95 | 4,55 | 4,65 | 5,66 |
| M.N. | 2,35 | 1,20 | 1,90 | 2,40 | 4,98 |
| T.V. | 4,45 | 1,40 | 2,40 | 2,95 | 6,00 |
| I.V. | 4,80 | 1,60 | 2,75 | 3,40 | 5,66 |
| Adj. | 5,00 | 1,25 | 3,90 | 3,90 | 5,66 |
| Adv. | 4,95 | 0,80 | 2,25 | 1,95 | 5,66 |
| Gemiddeldes | **4,44** | **1,70** | **2,96** | **3,21** | **5,60** |
| STDV | **1,05** | **1,13** | **1,04** | **0,99** | **0,33** |

Om te onderzoeken of er mogelijke significante correlaties of significante verschillen zijn tussen het kunstmatige model en de volwassenen en kinderen heb een statistische analyse gedaan.

Ik ga eerst het kunstmatige model vergelijken met de kinderen. Het gemiddelde puntenaantal voor de kinderen per groep was: groep 1: 1,70; groep 2: 2,96; groep 3: 3,21. De verschillende groepen hadden de volgende stdvs: groep 1: 1,13; groep 2: 1,04; groep 3: 0,99. Het kunstmatige model had een gemiddelde van 5,60 punten en een stdv van 0,33. Een paired samples t-test was uitgevoerd om de reacties van de kinderen met de reacties van het kunstmatige model te vergelijken:

* Groep 1: er was een significant verschil in reacties tussen de kinderen en het kunstmatige model, t(5) = -8.381, p < 0.001.
* Groep 2: er was een significant verschil in reacties tussen de kinderen en het kunstmatige model, t(5) = -6.547, p = 0.001.
* Groep 3: er was een significant verschil in reacties tussen de kinderen en het kunstmatige model, t(5) = -6.162, p = 0.002.

Tussen de reacties van alle groepen kinderen en de reacties van het kunstmatige model is er een verschil dat niet op het toeval berust.

Vervolgens vergelijk ik het kunstmatige model met de volwassenen. Het gemiddelde puntenaantal voor de volwassenen was 4,44 en de stdv was 1,05. Het kunstmatige model had een gemiddelde van 5,60 punten en een stdv van 0,33. Een paired samples t-test was uitgevoerd om de reacties van de volwassenen te vergelijken met de reacties van het kunstmatige model. Er was een significant verschil in reacties tussen de volwassenen en het kunstmatige model, t(5) = -3.548, p = 0.016. Dit wil zeggen dat er een verschil is tussen de volwassenen en het kunstmatige model dat niet op het toeval berust.

Ook heb ik een Pearson Correlation test uitgevoerd op de resultaten van de volwassenen en het kunstmatige model. Er was een significante correlatie tussen de reacties van de volwassenen en de reacties van het kunstmatige model, Pearson Correlation = 0.812, p = 0.049. Dit wil zeggen dat er een verband is tussen de reacties van de volwassenen en de reacties van het kunstmatige model. Woorden die een hoge score krijgen bij volwassenen hebben dus ook een hoge score bij het model.

Uit de resultaten kan de conclusie getrokken worden dat het kunstmatige model wat ik heb gebruikt niet vergelijkbaar is met mensen op het gebied van wisselen van woordsoort. Het model gaf bv. op het stimuluswoord ‘to send’ de volgende drie reacties: ‘sending’, ‘sent’ en ‘sends’ en op het stimuluswoord ‘doctor’ werden de volgende reacties gegeven: ‘physician’, ‘doctors’ en ‘gynecologist’. De manier waarop mensen en het kunstmatige model omgaan met woordassociaties is verschillend. Als er ook wordt gekeken naar de reacties die het model heeft gegeven kun je zien dat het model over het algemeen veel reageerde met vervoegingen op de stimulus werkwoorden. Op de zelfstandige naamwoorden werd als eerste reacties vaak het meervoud van het stimuluswoord gegeven. Ook werd er erg vaak een synoniem gegeven. Hierdoor bleef het model vaak binnen dezelfde woordsoort. In appendix D zijn hier meer voorbeelden van te zien. Dit zijn natuurlijk niet reacties die bij een mens veel zullen voorkomen, zoals te zien is aan de resultaten die ik bij mijn eigen experiment heb gevonden.

## 5.4 Beperkingen en toekomstig onderzoek

Ik wil dit niet zo zeer een beperking van mijn onderzoek noemen, maar een experiment uitvoeren op jonge kinderen is lastig. Ik heb dit onderschat. Ze werden snel afgeleid en de jongste kinderen van 4 jaar begrepen de opdracht vaak ook niet zo goed. Ook wilden ze me allemaal graag allerlei verhalen vertellen wat niet erg was, maar daardoor duurde het experiment wel vrij lang. Toch was het ook heel leuk om met kinderen te werken. Ze vonden het namelijk allemaal heel leuk om mee te doen. Aan de hand van mijn onderzoek naar kinderen kan ik een soort template opstellen voor toekomstige onderzoeken met betrekking tot woordassociaties naar kinderen. Wat ik merkte was dat kinderen erg veel moeite hadden met het noemen van een associatie bij woorden van de woordsoort bijwoord. Vandaar dat ik deze woordsoort in mijn uiteindelijk experiment heb weggelaten. In een onderzoek naar kinderen is het dus aan te raden om stimuluswoorden van woordsoort bijwoord buiten beschouwing te laten. Ook was mijn experiment in eerste instantie te lang. Eerst had ik in totaal 36 woorden en wilde ik de eerste 3 associaties van de kinderen noteren. Ik kwam er al snel achter dat dit te lang duurde voor de concentratie van de kinderen, vooral degene van 4 en 5 jaar. Daarom heb ik besloten om alle bijwoorden uit mijn experiment te halen en van de overige woordsoorten slechts 4 woorden te nemen in plaats van 6. Ook vroeg ik alleen maar naar de eerste associatie in plaats van de eerste 3. Door deze kortere variant bleven de kinderen er ook meer plezier inhouden. Voor een woordassociatie onderzoek naar kinderen is het dus aan te raden om het onderzoek niet te lang te laten duren, 5 tot 10 minuten moet voldoende zijn. Het belonen van de kinderen met een sticker had ook een positieve invloed. Als ik voor het experiment vertelde dat de kinderen een beloning kregen waren ze over het algemeen meteen al meer geconcentreerd.

Verder heb ik ook geen kunstmatig model kunnen vinden dat woordassociaties gaf voor het Nederlands. Daarom heb ik helaas mijn eigen resultaten niet kunnen vergelijken met een kunstmatig model. Een ander nadeel was dat Brown en Berko (1960) geen resultaten hadden van wat de reacties waren van de kinderen en de volwassenen. Er was alleen het puntenaantal gegeven wat de verschillende groepen hadden gescoord. Als de reacties ook waren gegeven had ik ook kunnen kijken naar mogelijke verschillen in bijvoorbeeld de woordvalentie waarden en de AoA waarden. Ook had ik dan kunnen kijken naar hoeveel overlap er mogelijk zou zijn tussen de mensen en het model. Hiermee had ik op nog meer gebieden het verschil tussen mensen en het kunstmatige model kunnen onderzoeken.

Het kunstmatige model wat ik heb gebruikt is wel een beperking van het onderzoek. Het was namelijk niet een heel goed model. Het voornamelijk vervoegingen en meervoudsvormen van de stimuluswoorden (zoals te zien in appendix D), waardoor het kunstmatige model over het algemeen bij dezelfde woordsoort bleef. Voor toekomstig onderzoek zou er wellicht naar een ander kunstmatig model gekeken kunnen worden, die bijvoorbeeld gebruik maakt van een andere implementatie dan het Word2Vec model.

Omdat ik alleen het wisselen van woordsoort heb controleren en het niet zo’n heel goed model was, is het vergelijkingsexperiment wat ik heb gedaan naar mijn mening niet van grote waarde voor de kunstmatige intelligentie wereld. Mocht ik meer tijd hebben zou ik graag voor een toekomstig onderzoek mijn eigen resultaten van de kinderen en volwassenen willen vergelijken met een kunstmatig model wat op Nederlandse woorden is getraind.

Wat een andere beperking zou kunnen zijn is dat ik mijn onderzoek in het Nederlands heb gedaan. Bijna al het voorgaande onderzoek waar ik mijn hypotheses op het gebaseerd zijn gebaseerd op het Engels of een andere taal. In hoeverre zijn de resultaten uit een andere taal te vergelijken met resultaten uit de Nederlandse taal.

Het experiment wat ik heb gedaan bij de kinderen heeft een mooie database opgeleverd. Deze data kan ook gebruikt worden voor toekomstig onderzoek van andere onderzoekers. Stel ze willen een specifiek onderwerp binnen woordassociaties onderzoeken van kinderen. Daar zouden ze mijn database voor kunnen gebruiken.

Aangezien mijn resultaten in gingen tegen de resultaten van Brown en Berko (1960) zou er eigenlijk een vervolgonderzoek moeten plaatsvinden. Dit zou moeten gebeuren met een grotere groep kinderen en volwassenen en met meer stimuluswoorden om te kijken of de resultaten die ik heb gevonden herhaalbaar zijn. Mochten dezelfde resultaten gevonden worden, namelijk dat kinderen niet meer wisselen van woordsoort dan volwassenen, zou het onderzoek van Brown en Berko in twijfel getrokken kunnen worden.

Er zou nog dieper op de relatie tussen het mentale lexicon en woordassociaties in gegaan kunnen worden door middel van vervolgonderzoek. Aan de hand van mijn bevindingen zou gesteld kunnen worden dat er clusters in het mentale lexicon zijn waarin woorden met vergelijkbare AoA en valentie waarden verzameld worden. Om deze bevindingen sterker te maken zou er een vervolgonderzoek gedaan kunnen worden waaruit dezelfde resultaten moeten blijken. Er zou ook nog gekeken kunnen worden of andere talen dezelfde resultaten laten zien. Misschien is bij mensen met een andere taal het mentale lexicon wel anders gestructureerd.

In deze scriptie is aangetoond dat kinderen en volwassenen qua woordassociaties verschillen van elkaar op het gebied diversiteit en woordfrequentie. Ze zijn vergelijkbaar op het gebied van wisselen van woordsoort, woordcombinaties, woord valentie waarde en AoA waarde. Deze resultaten kunnen een bijdrage leveren aan het beter begrijpen van de organisatie van het mentale lexicon. Hier dient echter ook nog vervolgonderzoek voor gedaan te worden. Tot slot is er ook nog een klein vergelijkingsexperiment gedaan met één kunstmatig model. Uit de resultaten hiervan blijkt dat de woordassociaties van mensen niet te vergelijken zijn met die van het kunstmatige model.

# 6. Referenties

## 6.1 Literatuur

Appel, R., Kuiken, F., & Vermeer, A. (1995). Nederlands als tweede taal in het basisonderwijs. Handboek voor leerkrachten in het basisonderwijs.

Bahn, D., Kauschke, C., Vesker, M., & Schwarzer, G. (2018). Perception of valence and arousal in German emotion terms: a comparison between 9-year-old children and adults. *Applied Psycholinguistics*, *39*(3), 463-481.

Brown, R., & Berko, J. (1960). Word association and the acquisition of grammar. *Child development*, 1-14.

Brysbaert, M., Keuleers, E., Mandera, P., & Stevens, M. (2013). Woordenkennis van Nederlanders en Vlamingen anno 2013: Resultaten van het Groot Nationaal Onderzoek Taal.

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.

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, *104*(2), 215-226.

Cofer, C. N., & Shevitz, R. (1952). Word-association as a function of word-frequency. *The American journal of psychology*, *65*(1), 75-79.

Coronges, K. A., Stacy, A. W., & Valente, T. W. (2007). Structural Comparison of Cognitive Associative Networks in Two Populations 1. *Journal of Applied Social Psychology*, *37*(9), 2097-2129.

De Deyne, S., Kenett, Y. N., Anaki, D., Faust, M., & Navarro, D. J. (2016). Large-scale network representations of semantics in the mental lexicon. *Big data in cognitive science: From methods to insights*, 174-202.

De Deyne, S., Navarro, D., & Storms, G. (2013). Associative strength and semantic activation in the mental lexicon: evidence from continued word associations. *Cognitive Science Society.*

De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior research methods*, *45*(2), 480-498.

De Deyne, S., Verheyen, S., & Storms, G. (2016). Structure and organization of the mental lexicon: A network approach derived from syntactic dependency relations and word associations. In *Towards a theoretical framework for analyzing complex linguistic networks* (pp. 47-79). Springer, Berlin, Heidelberg.

Ervin, S. M. (1957). Grammar and classification. *American Psychological Association, New York*.

Ervin, S. M. (1961). Changes with age in the verbal determinants of word-association. *The American journal of psychology*, *74*(3), 361-372.

Froese, T. (2007). On the role of AI in the ongoing paradigm shift within the cognitive sciences. In *50 years of artificial intelligence* (pp. 63-75). Springer, Berlin, Heidelberg.

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.

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological bulletin*, *131*(5), 684.

Kenett, Y. N., Kenett, D. Y., Ben-Jacob, E., & Faust, M. (2011). Global and local features of semantic networks: Evidence from the Hebrew mental lexicon. *PloS one*, *6*(8).

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods*, *42*(3), 643-650.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., Van Schie, K., Van Harmelen, A. L., ... & Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior research methods*, *45*(1), 169-177.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 116.

Nasrallah, M., Carmel, D., & Lavie, N. (2009). Murder, she wrote: enhanced sensitivity to negative word valence. *Emotion*, *9*(5), 609.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402-407.

Precosky, K. (2011). Exploring the mental lexicon using word association tests: How do native and non-native speakers of English arrange words in the mind.

Vasa, R. A., Carlino, A. R., London, K., & Min, C. (2006). Valence ratings of emotional and non-emotional words in children. *Personality and individual differences*, *41*(6), 1169-1180.

Vitevitch, M. S., Goldstein, R., Siew, C. S., & Castro, N. (2014, December). Using complex networks to understand the mental lexicon. In *Yearbook of the Poznan Linguistic Meeting* (Vol. 1, No. 1, pp. 119-138). De Gruyter Open.

Woodrow, H., & Lowell, F. (1916). Children's association frequency tables. *The Psychological Monographs*, *22*(5), i.

Zortea, M., & de Salles, J. F. (2012). Semantic word association: Comparative data for Brazilian children and adults. *Psychology & Neuroscience*, *5*(1), 77.

## 6.2 Websites

A list of the 1000 most commonly spoken Dutchs words. (z.d.). Geraadpleegd van <https://1000mostcommonwords.com/1000-most-common-dutch-words/>

Martin, F.R. (2019, 23 januari) How important are semantic networks in Artificial Intelligence. Geraadpleegd van <https://analyticsindiamag.com/semantic-networks-ai/>

Nicholson, C. (2019) A beginner’s guide to Word2Vec and neural word embeddings. Geraadpleegd van <https://pathmind.com/wiki/word2vec>

Skillicorn, N. (2016, 5 augustus) Evidence that children become less creative over time (and how to fix it). Geraadpleegd van <https://www.ideatovalue.com/crea/nickskillicorn/2016/08/evidence-children-become-less-creative-time-fix/>

Small World of Words: Discover what words mean for people worldwide. (z.d.). Geraadpleegd van

<https://smallworldofwords.org/nl/project/home>

Small World of Words: Help researchers discover how the meaning of words is stored in memory. (2016, 10 mei). Geraadpleegd van <https://www.scientificamerican.com/citizen-science/small-world-of-words/>

Word Embedding Demo (z.d.). Geraadpleegd van <http://bionlp-www.utu.fi/wv_demo/>

# 7. Appendix

## A. Age of Acquisition en Frequentie waarden van (100 willekeurig gekozen woorden uit) A Small World of Words

### A1. Age of Acquisition waarden

|  |  |  |  |
| --- | --- | --- | --- |
| Stimulus | AoA waarde | Meest voorkomende Reactie | AoA waarde |
| aaien | 5,59 | strelen | 6,65 |
| aal | 9,98 | paling | 8,42 |
| aambeeld | 11,32 | hamer | 6,50 |
| aanbellen | 7,39 | deur | 4,44 |
| aanbieding | 10,29 | koopje | 8,78 |
| aandacht | 8,15 | concentratie | 9,35 |
| aandoening | 10,55 | ziekte | 5,96 |
| aanduiden | 7,11 | wijzen | 6,15 |
| aangapen | 9,11 | staren | 8,53 |
| aangenaam | 8,65 | leuk | 4,34 |
| aangezien | 10,78 | omdat | 6,00 |
| aanhangwagen | 9,00 | auto | 4,22 |
| bundel | 9,00 | boek | 4,55 |
| bunker | 9,15 | oorlog | 7,40 |
| bureau | 6,56 | stoel | 3,95 |
| burgemeester | 7,42 | gemeente | 6,95 |
| bus | 4,99 | school | 3,61 |
| concreet | 10,78 | duidelijk | 7,44 |
| conditie | 9,94 | sport | 5,68 |
| confetti | 5,96 | carnaval | 5,15 |
| connectie | 11,44 | verbinding | 8,83 |
| consument | 11,89 | koper | 9,18 |
| eenheid | 9,28 | leger | 7,05 |
| eenhoorn | 8,26 | sprookje | 4,65 |
| gehaast | 8,39 | snel | 5,53 |
| gehaktbal | 6,78 | vlees | 4,61 |
| geheel | 8,78 | alles | 5,33 |
| gemeente | 6,95 | stad | 5,79 |
| hart | 5,56 | liefde | 7,09 |
| hinken | 7,33 | manken | 7,78 |
| hoorn | 8,24 | blazen | 5,03 |
| huid | 6,89 | vel | 5,84 |
| jaar | 4,45 | maand | 6,34 |
| kasteel | 5,41 | ridder | 6,28 |
| kerstboom | 4,39 | kerstmis | 5,03 |
| kleren | 4,39 | broek | 4,19 |
| kopie | 8,83 | dubbel | 6,71 |
| kurk | 7,15 | wijn | 7,31 |
| landkaart | 8,90 | reizen | 6,28 |
| leerlingen | 6,85 | school | 3,61 |
| likken | 5,90 | tong | 4,57 |
| majesteit | 7,94 | koning | 5,32 |
| matras | 6,03 | bed | 3,76 |
| meloen | 7,38 | fruit | 4,70 |
| middeleeuwen | 9,33 | ridder | 6,28 |
| missen | 6,40 | verdriet | 6,21 |
| moeras | 8,96 | drassig | 10,69 |
| nachtegaal | 8,33 | vogel | 4,24 |
| nek | 5,19 | hals | 6,71 |
| nier | 9,34 | orgaan | 10,09 |
| noordpool | 8,06 | koud | 4,09 |
| octopus | 7,85 | inktvis | 7,51 |
| oefenen | 7,00 | proberen | 6,78 |
| oerwoud | 7,60 | bos | 5,34 |
| ongeluk | 6,03 | auto | 4,22 |
| oud | 5,28 | jong | 6,28 |
| pan | 5,21 | koken | 5,71 |
| pannenkoek | 4,89 | lekker | 4,53 |
| paraplu | 5,61 | regen | 4,49 |
| persoonlijkheid | 9,72 | karakter | 9,28 |
| piloot | 6,88 | vliegtuig | 5,69 |
| politiek | 10,44 | saai | 6,78 |
| raam | 4,58 | venster | 5,32 |
| redder | 7,15 | zee | 4,78 |
| regen | 4,49 | nat | 4,59 |
| religie | 9,61 | godsdienst | 6,96 |
| rolstoel | 7,05 | gehandicapt | 7,84 |
| rook | 6,31 | vuur | 4,89 |
| samen | 6,17 | twee | 3,85 |
| schaken | 8,51 | spel | 4,31 |
| semester | 10,39 | school | 3,61 |
| snelweg | 7,83 | auto | 4,22 |
| spelden | 7,71 | naaien | 7,03 |
| surfen | 9,24 | zee | 4,78 |
| taart | 6,11 | lekker | 4,53 |
| tante | 5,04 | familie | 4,96 |
| tegel | 7,04 | vloer | 5,08 |
| termijn | 10,84 | periode | 9,06 |
| trommel | 6,65 | muziek | 4,84 |
| tulp | 7,13 | bloem | 4,27 |
| tweeling | 6,22 | dubbel | 6,71 |
| uitgang | 7,03 | deur | 4,44 |
| uitvinder | 8,25 | professor | 9,47 |
| uniek | 9,34 | enig | 8,71 |
| vallen | 4,46 | pijn | 3,79 |
| vangst | 8,00 | vis | 4,53 |
| verdeling | 8,56 | erfenis | 10,33 |
| vertegenwoordiger | 11,79 | verkoper | 7,15 |
| veter | 5,38 | schoen | 3,71 |
| vlam | 6,65 | vuur | 4,89 |
| walnoot | 8,10 | okkernoot | 9,12 |
| wandelaar | 6,83 | bos | 5,34 |
| werpen | 6,90 | gooien | 4,84 |
| winkel | 5,09 | kopen | 6,15 |
| yoghurt | 5,46 | wit | 4,54 |
| zaal | 6,18 | feest | 5,75 |
| zagen | 6,21 | hout | 5,41 |
| zebra | 6,15 | gestreept | 7,25 |
| zeef | 7,33 | bloem | 4,27 |
| zenuwen | 8,34 | stress | 10,40 |
| Gemiddeldes | **7,55** |  | **6,01** |
| STDV | **1,87** |  | **1,79** |

### A2. Frequentie waarden

|  |  |  |  |
| --- | --- | --- | --- |
| Stimulus | Frequentie waarde (in miljoenen) | Meest voorkomende Reactie | Frequentie waarde (in miljoenen) |
| aaien | 2,10 | strelen | 2,65 |
| aal | 0,80 | paling | 0,94 |
| aambeeld | 0,50 | hamer | 8,55 |
| aanbellen | 0,89 | deur | 247,48 |
| aanbieding | 4,14 | koopje | 3,38 |
| aandacht | 56,69 | concentratie | 6,24 |
| aandoening | 2,84 | ziekte | 37,14 |
| aanduiden | 0,39 | wijzen | 22,34 |
| aangapen | 0,16 | staren | 13,08 |
| aangenaam | 57,86 | leuk | 595,57 |
| aangezien | 25,54 | omdat | 869,39 |
| aanhangwagen | 0,34 | auto | 458,00 |
| bundel | 0,85 | boek | 31,05 |
| bunker | 5,10 | oorlog | 178,96 |
| bureau | 66,93 | stoel | 51,20 |
| burgemeester | 33,46 | gemeente | 6,91 |
| bus | 64,83 | school | 246,95 |
| concreet | 0,82 | duidelijk | 170,82 |
| conditie | 7,43 | sport | 20,12 |
| confetti | 0,71 | carnaval | 2,10 |
| connectie | 5,79 | verbinding | 16,88 |
| consument | 0,43 | koper | 8,96 |
| eenheid | 24,72 | leger | 107,98 |
| eenhoorn | 2,45 | sprookje | 6,11 |
| gehaast | 4,07 | snel | 463,60 |
| gehaktbal | 1,14 | vlees | 61,67 |
| geheel | 23,42 | alles | 1708,42 |
| gemeente | 6,91 | stad | 272,61 |
| hart | 196,37 | liefde | 208,90 |
| hinken | 0,25 | manken | 0,05 |
| hoorn | 8,00 | blazen | 17,65 |
| huid | 39,56 | vel | 6,20 |
| jaar | 762,67 | maand | 93,64 |
| kasteel | 27,60 | ridder | 13,58 |
| kerstboom | 3,82 | kerstmis | 29,07 |
| kleren | 104,85 | broek | 67,28 |
| kopie | 16,85 | dubbel | 17,08 |
| kurk | 1,62 | wijn | 60,44 |
| landkaart | 0,96 | reizen | 32,38 |
| leerlingen | 14,02 | school | 246,95 |
| likken | 7,50 | tong | 31,90 |
| majesteit | 42,74 | koning | 138,53 |
| matras | 5,15 | bed | 239,93 |
| meloen | 2,01 | fruit | 12,94 |
| middeleeuwen | 2,63 | ridder | 13,58 |
| missen | 85,39 | verdriet | 26,48 |
| moeras | 9,74 | drassig | 0,18 |
| nachtegaal | 1,07 | vogel | 32,27 |
| nek | 57,72 | hals | 7,50 |
| nier | 3,91 | orgaan | 2,45 |
| noordpool | 5,19 | koud | 95,95 |
| octopus | 2,20 | inktvis | 2,08 |
| oefenen | 24,99 | proberen | 224,54 |
| oerwoud | 3,00 | bos | 46,51 |
| ongeluk | 100,96 | auto | 458,00 |
| oud | 183,26 | jong | 91,61 |
| pan | 9,38 | koken | 33,80 |
| pannenkoek | 1,46 | lekker | 276,15 |
| paraplu | 3,43 | regen | 26,48 |
| persoonlijkheid | 8,90 | karakter | 17,22 |
| piloot | 30,12 | vliegtuig | 89,92 |
| politiek | 30,87 | saai | 31,47 |
| raam | 70,84 | venster | 3,45 |
| redder | 4,92 | zee | 67,80 |
| regen | 26,48 | nat | 30,57 |
| religie | 7,29 | godsdienst | 4,02 |
| rolstoel | 8,37 | gehandicapt | 2,58 |
| rook | 46,63 | vuur | 100,57 |
| samen | 370,64 | twee | 1007,51 |
| schaken | 4,00 | spel | 95,11 |
| semester | 6,08 | school | 246,95 |
| snelweg | 18,34 | auto | 458,00 |
| spelden | 0,96 | naaien | 18,02 |
| surfen | 5,10 | zee | 67,80 |
| taart | 31,35 | lekker | 276,15 |
| tante | 62,34 | familie | 304,12 |
| tegel | 0,57 | vloer | 31,72 |
| termijn | 6,29 | periode | 13,90 |
| trommel | 1,85 | muziek | 107,46 |
| tulp | 0,48 | bloem | 13,49 |
| tweeling | 15,80 | dubbel | 17,08 |
| uitgang | 14,68 | deur | 247,48 |
| uitvinder | 3,09 | professor | 67,21 |
| uniek | 12,74 | enig | 51,75 |
| vallen | 180,98 | pijn | 266,16 |
| vangst | 3,84 | vis | 50,08 |
| verdeling | 1,58 | erfenis | 8,05 |
| vertegenwoordiger | 5,28 | verkoper | 9,51 |
| veter | 1,33 | schoen | 13,45 |
| vlam | 7,94 | vuur | 100,57 |
| walnoot | 0,71 | lekker | 276,15 |
| wandelaar | 0,55 | bos | 46,51 |
| werpen | 7,82 | gooien | 49,62 |
| winkel | 65,13 | kopen | 130,64 |
| yoghurt | 2,15 | wit | 33,48 |
| zaal | 15,41 | feest | 98,26 |
| zagen | 52,66 | hout | 23,58 |
| zebra | 3,06 | gestreept | 0,50 |
| zeef | 0,57 | bloem | 13,49 |
| zenuwen | 16,37 | stress | 13,86 |
| Gemiddeldes | **32,84** |  | **126,78** |
| STDV | **89,65** |  | **234,79** |

## B. Informatiebrief en toestemmingsverklaring voor schoolleiding en ouders

### B1. Informatiebrief ouders

FETC-GW-referentienummer: zucke102-03-03-2019

(dit nummer pas invullen ná goedkeuring door de FETC-GW)

**Informatie over deelname aan**

**Onderzoek woordassociaties bij kinderen.**

Beste ouder/verzorger,

**1. Inleiding**

Door middel van deze brief zou ik uw toestemming willen vragen voor het deelnemen van uw kind aan een onderzoek over woordassociaties. Dit onderzoek zal plaatsvinden op de basisschool Puur Sang – Locatie Sint Lucia. Ik heb toestemming gekregen om dit onderzoek uit te voeren van Rob van den Crommenacker, directeur Kindcentrum Puur Sang.

**2. Wat is de achtergrond en het doel van het onderzoek?**

Ik ben Marieke Maas en ben student aan de universiteit in Utrecht. Ik zit momenteel in mijn 4e jaar van de opleiding Kunstmatige Intelligentie. Op dit moment ben ik bezig met mijn scriptie over woordassociaties. Voor mijn scriptie wil ik onderzoeken of er een verschil is in woordassociaties tussen kinderen en volwassenen en wat dit te maken heeft met het mentale lexicon (alle woorden iemand kent).

**3. Hoe wordt het onderzoek uitgevoerd?**

Tijdens het onderzoek zal ik een woord opnoemen en vervolgens uw kind vragen naar de eerste 3 associaties die uw kind bij dit woord heeft. Ik zal het onderzoek als een soort spel opstellen zodat het leuker is voor de kinderen en ze minder snel worden afgeleid.

*IB\_v22032019*

(Naar een voorbeeld van CCMO: www.ccmo.nl)

FETC-GW-referentienummer: zucke102-03-03-2019

(dit nummer pas invullen ná goedkeuring door de FETC-GW)

**4. Wat wordt er van uw kind verwacht?**

In totaal zal ik voor 36 woorden naar de associaties van de kinderen vragen. De antwoorden die uw kind geeft zal ik noteren. Het onderzoek zal plaats vinden op de basisschool in een ongebruikt klaslokaal. Het zal bestaan uit één sessie en zal ongeveer 10 tot 15 minuten duren. Een voorbeeld zou kunnen zijn dat ik het woord ‘appel’ noem en uw kind de volgende associaties geeft: ‘fruit’, ‘peer’ en ‘lekker’.

**5. Wat zijn mogelijke voor- en nadelen van deelname aan dit onderzoek?**

Uw kind heeft zelf geen voordeel van deelname aan dit onderzoek. De resultaten van het onderzoek kunnen echter wel nuttig zijn. Zo kan dit onderzoek meer informatie opleveren over de taalontwikkeling van kinderen.

Mogelijke nadelen zijn:

* Onderbreking van de normale les.
* Langere tijd stilzitten.
* Het onderzoek vergt een lange concentratie van uw kind.

**6. Wat gebeurt er als uw kind toch niet mee wil doen tijdens het onderzoek?**

Het kan zijn dat uw kind tijdens het onderzoek niet wil meewerken. Wanneer dit gebeurt zal ik eerst vragen waarom uw kind niet mee wil werken en of hij/zij het misschien nog een keer wil proberen. Wanneer dit niet zal blijken te werken zal het onderzoek gestopt worden.

**7. Vrijwillige deelname**

Deelname aan het onderzoek is vrijwillig. Wanneer u besluit dat uw kind toch niet mag meedoen, hoeft uw verder niets te doen. U hoeft hier ook geen reden voor op te geven. Als uw kind wel mag deelnemen aan het onderzoek, kan u uw toestemming op ieder moment intrekken. Bovendien kunt u binnen 5 dagen na het onderzoek nog uw toestemming intrekken. De gegevens van uw kind zullen dan niet worden meegenomen in de analyse.

*IB\_v22032019*

*(Naar een voorbeeld van CCMO:* [*www.ccmo.nl*](http://www.ccmo.nl)*)*

FETC-GW-referentienummer: zucke102-03-03-2019

(dit nummer pas invullen ná goedkeuring door de FETC-GW)

**8. Wat gebeurt er met de verzamelde gegevens?**

Het onderzoek zal anoniem zijn. Het enige dat ik noteer is de leeftijd en het geslacht van uw kind. De data worden dus opgeslagen zonder de namen van de kinderen en er is geen manier om daarna terug te halen welke reactie bij welk kind hoort. Wij zijn verplicht de geanonimiseerde onderzoeksgegevens 10 jaar te bewaren. De anonieme data worden bewaard en beheerd door mij onder begeleiding van Dr. S. Zuckerman, afdeling Taalwetenschap, Universiteit Utrecht, en kunnen gebruikt worden voor onderzoeksdoeleinden, en kunnen ook met andere onderzoekers gedeeld worden.

**9. Is er een vergoeding wanneer U besluit aan dit onderzoek mee te doen?**

Elk kind krijgt een kleine beloning (sticker) na het onderzoek.

**10. Goedkeuring van dit onderzoek**

De Linguïstiek kamer van de Facultaire Ethische Toetsingscommissie (FETC-GW), heeft dit onderzoek goedgekeurd. Wanneer u een klacht wilt indienen over de procedure omtrent dit onderzoek, dan kunt u contact opnemen met de secretaris van de FETC-GW, e-mail: [fetc-gw@uu.nl](mailto:fetc-gw@uu.nl) .

Als uw kind mag mee doen, zou u dan s.v.p. bijgevoegde toestemmingsverklaring willen invullen, ondertekenen, en aan uw kind meegeven? Bij voorbaat hartelijk dank.

Voor verdere informatie over het onderzoek of andere vragen kunt u terecht bij:

Marieke Maas – 0629893705 .

Met vriendelijke groet,

Marieke Maas

*IB\_v22032019*

*(Naar een voorbeeld van CCMO:* [*www.ccmo.nl*](http://www.ccmo.nl)*)*

### B2. Toestemmingsverklaring ouders



FETC-GW-referentienummer: zucke102-03-03-2019

(dit nummer pas invullen ná goedkeuring door de FETC-GW)

**TOESTEMMINGSVERKLARING voor deelname aan:**

**Onderzoek woordassociaties bij kinderen.**

Ik, ouder of voogd van het hieronder genoemde kind, bevestig:

* dat ik via de informatiebrief naar tevredenheid over het onderzoek ben ingelicht;
* dat ik in de gelegenheid ben gesteld om vragen over het onderzoek te stellen en dat mijn eventuele vragen naar tevredenheid zijn beantwoord;
* dat ik gelegenheid heb gehad om grondig over deelname aan het onderzoek na te denken;
* dat ik uit vrije wil toestemming geef dat mijn kind deelneemt aan het onderzoek.

Ik stem er mee in dat:

* de verzamelde gegevens voor wetenschappelijke doelen worden verkregen en bewaard zoals in de informatiebrief vermeld staat;
* de verzamelde, geheel anonieme, onderzoeksgegevens door wetenschappers kunnen worden gedeeld en/of worden hergebruikt om eventueel andere onderzoeksvragen mee te beantwoorden;

Ik begrijp dat:

* ik het recht heb om mijn toestemming voor het gebruik van mijn data in te trekken zoals vermeld staat in de informatiebrief.

Naam ouder of voogd: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Geboortedatum: \_\_\_ / \_\_\_ / \_\_\_\_ (dd/mm/jjjj)

Handtekening ouder/vgd.: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Datum, plaats: \_\_\_ / \_\_\_ / \_\_\_\_, \_\_\_\_\_\_\_\_\_\_

Naam kind: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Geboortedatum: \_\_\_ / \_\_\_ / \_\_\_\_ (dd/mm/jjjj)

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

Datum: \_\_\_ / \_\_\_ / \_\_\_\_\_ (dd/mm/jjjj)

Handtekening:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**In te vullen door de uitvoerend onderzoeker:**

Ik verklaar dat ik bovengenoemde deelnemer heb uitgelegd

Wat deelname inhoudt en dat ik borg sta voor

de privacy van de gegevens.

*TV\_Kinderen\_0-11\_NL\_v11042019* bestemd voor onderzoek met kinderen van 0 t/m 11 jaar

### B3. Informatiebrief schoolleiding

FETC-GW-referentienummer: zucke102-03-03-2019

(dit nummer pas invullen ná goedkeuring door de FETC-GW)

**Informatie over**

**Onderzoek woordassociaties bij kinderen.**

Beste schoolleiding,

**1. Inleiding**

Door middel van deze brief wil ik graag informatie verschaffen over het onderzoek over woordassociaties bij kinderen dat binnenkort zal plaatsvinden op uw basisschool.

**2. Wat is de achtergrond en het doel van het onderzoek?**

Ik ben Marieke Maas en ik ben student aan de universiteit in Utrecht. Ik zit momenteel in mijn 4e jaar van de opleiding Kunstmatige Intelligentie. Op dit moment ben ik bezig met mijn scriptie over woordassociaties. Voor mijn scriptie wil ik onderzoeken of er een verschil is in woordassociaties tussen kinderen en volwassenen en wat dit te maken heeft met het mentale lexicon (alle woorden iemand kent).

**3. Hoe wordt het onderzoek uitgevoerd?**

Tijdens het onderzoek zal ik een woord opnoemen en vervolgens het kind vragen naar de eerste 3 associaties die het kind bij dit woord heeft. Ik zal het onderzoek als een soort spel opstellen zodat het leuker is voor de kinderen en ze minder snel worden afgeleid. Ook krijgen ze een kleine beloning naderhand (sticker).

**4. Wat wordt er van u verwacht?**

In totaal zal ik voor 36 woorden naar de associaties van de kinderen vragen. De antwoorden die het kind geeft zal ik noteren. Het onderzoek zal plaats vinden op de basisschool in een ongebruikt klaslokaal. Het zal bestaan uit één sessie en zal ongeveer 10 tot 15 minuten duren. Een voorbeeld zou kunnen zijn dat ik het woord ‘appel’ noem en uw kind de volgende associaties geeft: ‘fruit’, ‘peer’ en ‘lekker’.

FETC-GW-referentienummer: zucke102-03-03-2019

(dit nummer pas invullen ná goedkeuring door de FETC-GW)

**5. Wat zijn mogelijke voor- en nadelen van deelname aan dit onderzoek?**

Het kind heeft zelf geen voordeel van deelname aan dit onderzoek. De resultaten van het onderzoek kunnen echter wel nuttig zijn. Zo kan dit onderzoek meer informatie opleveren over de taalontwikkeling van kinderen.

Mogelijke nadelen zijn:

* Onderbreking van de normale les.
* Langere tijd stilzitten.
* Het onderzoek vergt een lange concentratie van uw kind.

**6. Wat gebeurt er als uw kind toch niet mee wil doen tijdens het onderzoek?**

Het kan zijn dat het kind tijdens het onderzoek niet wil meewerken. Wanneer dit gebeurt zal ik eerst vragen waarom het kind niet mee wil werken en of hij/zij het misschien nog een keer wil proberen. Wanneer dit niet zal blijken te werken zal het onderzoek gestopt worden.

**7. Vrijwillige deelname**

Deelname aan het onderzoek is vrijwillig. Wanneer een ouder besluit dat zijn/haar kind toch niet mag meedoen, hoeft er verder niets te gebeuren. De ouder hoeft hier ook geen reden voor op te geven. Als het kind wel toestemming krijgt om deel te nemen aan het onderzoek, dan kan deze toestemming op ieder moment ingetrokken worden. Er kan ook tijdens het onderzoek besloten worden om te stoppen. Bovendien kan binnen 5 dagen na het onderzoek de toestemming alsnog ingetrokken worden. De gegevens van het kind zullen dan niet worden meegenomen in de analyse.

**8. Wat gebeurt er met de verzamelde gegevens?**

Het onderzoek zal anoniem zijn. Het enige dat ik noteer is de leeftijd en het geslacht van het kind. De data worden dus opgeslagen zonder de namen van de kinderen en er is geen manier om daarna terug te halen welke reactie bij welk kind hoort. Wij zijn verplicht de geanonimiseerde onderzoeksgegevens 10 jaar te bewaren. De anonieme data worden bewaard en beheerd door mij onder begeleiding van Dr. S. Zuckerman, afdeling Taalwetenschap, Universiteit Utrecht, en kunnen gebruikt worden voor onderzoeksdoeleinden, en kunnen ook met andere onderzoekers gedeeld worden.

FETC-GW-referentienummer: zucke102-03-03-2019

(dit nummer pas invullen ná goedkeuring door de FETC-GW)

**9. Is er een vergoeding wanneer U besluit aan dit onderzoek mee te doen?**

Elk kind krijgt een kleine beloning (sticker) na het onderzoek.

**10. Goedkeuring van dit onderzoek**

De Linguïstiek kamer van de Facultaire Ethische Toetsingscommissie (FETC-GW), heeft dit onderzoek goedgekeurd. Wanneer u een klacht wilt indienen over de procedure omtrent dit onderzoek, dan kunt u contact opnemen met de secretaris van de FETC-GW, e-mail: [fetc-gw@uu.nl](mailto:fetc-gw@uu.nl) .

Voor verdere informatie over het onderzoek of andere vragen kunt u terecht bij:

Marieke Maas – 0629893705 .

Met vriendelijke groet,

Marieke Maas

## C. Reacties van kinderen en volwassenen

### C1. Top 3 reacties van de kinderen.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Stimulus | Woordsoort | Top 3 Reacties Kinderen | Woordsoort | Aantal | Percentage |
| Bouwen | O ww | Blokken | T znw | 15/46 | 33% |
|  |  | Huis | T znw | 4/46 | 9% |
|  |  | Maken | O ww | 3/46 | 7% |
| Computer | T znw | Scherm | T znw | 6/46 | 13% |
|  |  | Spelen | O ww | 6/46 | 13% |
|  |  | Spelletjes | T znw | 5/46 | 11% |
| Drinken | O ww | Water | O znw | 9/46 | 20% |
|  |  | Melk | O znw | 7/46 | 15% |
|  |  | Beker | T znw | 4/46 | 9% |
| Gitaar | T znw | Muziek | O znw | 18/46 | 39% |
|  |  | Spelen | O ww | 5/46 | 11% |
|  |  | Instrument | T znw | 4/46 | 9% |
| Groeien | On ww | Groot | Bnw | 9/46 | 20% |
|  |  | Groter | Bwn | 6/46 | 13% |
|  |  | Krimpen | On ww | 4/46 | 9% |
| Huilen | On ww | Verdrietig | Bnw | 9/46 | 20% |
|  |  | Tranen | T znw | 5/46 | 11% |
|  |  | Pijn | O znw | 4/46 | 9% |
| Kip | T znw | Eieren | T znw | 16/46 | 35% |
|  |  | Haan | T znw | 8/46 | 17% |
|  |  | Boerderij | T znw | 2/46 | 4% |
| Klein | Bnw | Groot | Bnw | 11/46 | 24% |
|  |  | Spelen | O ww | 5/46 | 11% |
|  |  | Kind | T znw | 3/46 | 7% |
| Lezen | O ww | Boek | T znw | 12/46 | 26% |
|  |  | Letters | T znw | 6/46 | 13% |
|  |  | Schrijven | O ww | 4/46 | 9% |
| Liefde | O znw | Verliefd | Bnw | 9/46 | 20% |
|  |  | Lief | Bnw | 5/46 | 11% |
|  |  | Trouwen | On ww | 4/46 | 9% |
| Melk | O znw | Drinken | O ww | 14/46 | 30% |
|  |  | Koeien | T znw | 10/46 | 22% |
|  |  | Ranja | O znw | 3/46 | 7% |
| Moeilijk | Bnw | Makkelijk | Bnw | 14/46 | 30% |
|  |  | - |  | 4/46 | 9% |
|  |  | Leren | On ww | 2/46 | 4% |
| Rood | Bnw | Kleur | T znw | 10/46 | 22% |
|  |  | Groen | Bnw | 6/46 | 13% |
|  |  | Oranje | Bnw | 4/46 | 9% |
| Rugzak | T znw | School | T znw | 13/46 | 28% |
|  |  | Spullen | O znw | 8/46 | 17% |
|  |  | Tas | T znw | 6/46 | 13% |
| Snurken | On ww | Slapen | O ww | 27/46 | 27% |
|  |  | Bed | T znw | 4/46 | 4% |
|  |  | Geluid | T znw | 3/46 | 3% |
| Spelen | O ww | Leuk | Bnw | 3/46 | 7% |
|  |  | Speelgoed | O znw | 3/46 | 7% |
|  |  | Buiten | Bw | 2/46 | 4% |
| Springen | On ww | Trampoline | T znw | 12/46 | 26% |
|  |  | Hoog | Bnw | 6/46 | 13% |
|  |  | Lucht | O znw | 4/46 | 9% |
| Suiker | O znw | Snoep | O znw | 7/46 | 15% |
|  |  | Zoet | Bnw | 3/46 | 7% |
|  |  | Klontjes | T znw | 3/46 | 7% |
| Zacht | Bnw | Hard | Bnw | 12/46 | 26% |
|  |  | Knuffel | T znw | 8/46 | 17% |
|  |  | Kussen | T znw | 4/46 | 9% |
| Zand | O znw | Zandbak | T znw | 10/46 | 22% |
|  |  | Schep | T znw | 7/46 | 15% |
|  |  | Spelen | O ww | 5/46 | 11% |

### C2. Top 3 reacties van de volwassenen (gebaseerd op het Small World of Words project)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Stimulus | Woordsoort | Top 3 Reacties Volwassenen | Woordsoort | Aantal | Percentage |
| Bouwen | O ww | Huis | T znw | 71/100 | 71% |
|  |  | Steen | T znw | 6/100 | 6% |
|  |  | Baksteen | T znw | 3/100 | 3% |
| Computer | T znw | Internet | O znw | 13/100 | 13% |
|  |  | Scherm | T znw | 12/100 | 12% |
|  |  | Pc | T znw | 9/100 | 9% |
| Drinken | O ww | Bier | O znw | 23/100 | 23% |
|  |  | Water | O znw | 19/100 | 19% |
|  |  | Dorst | O znw | 15/100 | 15% |
| Gitaar | T znw | Muziek | O znw | 49/100 | 49% |
|  |  | Snaar | T znw | 11/100 | 11% |
|  |  | Snaren | T znw | 9/100 | 9% |
| Groeien | On ww | Kind | T znw | 15/100 | 15% |
|  |  | Kinderen | T znw | 10/100 | 10% |
|  |  | Groot | Bnw | 8/100 | 8% |
| Huilen | On ww | Wenen | O ww | 26/100 | 26% |
|  |  | Tranen | T znw | 22/100 | 22% |
|  |  | Verdriet | O znw | 16/100 | 16% |
| Kip | T znw | Ei | T znw | 30/100 | 30% |
|  |  | Eten | O znw | 6/100 | 6% |
|  |  | Haan | T znw | 5/100 | 5% |
| Klein | Bnw | Groot | Bnw | 36/100 | 36% |
|  |  | Kabouter | T znw | 9/100 | 9% |
|  |  | Kind | T znw | 7/100 | 7% |
| Lezen | O ww | Boek | T znw | 73/100 | 73% |
|  |  | Schrijven | O ww | 4/100 | 4% |
|  |  | Letters | T znw | 3/100 | 3% |
| Liefde | O znw | Rood | Bnw | 14/100 | 14% |
|  |  | Hart | T znw | 13/100 | 13% |
|  |  | Geluk | O znw | 5/100 | 5% |
| Melk | O znw | Koe | T znw | 36/100 | 36% |
|  |  | Wit | Bnw | 30/100 | 30% |
|  |  | Koffie | O znw | 3/100 | 3% |
| Moeilijk | Bnw | Examen | T znw | 14/100 | 14% |
|  |  | Gemakkelijk | Bnw | 13/100 | 13% |
|  |  | Makkelijk | Bnw | 9/100 | 9% |
| Rood | Bnw | Bloed | O znw | 19/100 | 19% |
|  |  | Kleur | Bnw | 12/100 | 12% |
|  |  | Liefde | O znw | 12/100 | 12% |
| Rugzak | T znw | Vakantie | T znw | 13/100 | 13% |
|  |  | Trekken | O ww | 9/100 | 9% |
|  |  | Reis | T znw | 8/100 | 8% |
| Snurken | On ww | Slapen | O ww | 29/100 | 29% |
|  |  | Lawaai | O znw | 7/100 | 7% |
|  |  | Ronken | On ww | 6/100 | 6% |
| Spelen | O ww | Kinderen | T znw | 20/100 | 20% |
|  |  | Bal | T znw | 11/100 | 11% |
|  |  | Plezier | O znw | 7/100 | 7% |
| Springen | On ww | Trampoline | T znw | 16/100 | 16% |
|  |  | Touw | T znw | 9/100 | 9% |
|  |  | Springtouw | T znw | 6/100 | 6% |
| Suiker | O znw | Zoet | Bnw | 44/100 | 44% |
|  |  | Snoep | O znw | 6/100 | 6% |
|  |  | Wit | Bnw | 6/100 | 6% |
| Zacht | Bnw | Kussen | T znw | 15/100 | 15% |
|  |  | Dons | O znw | 7/100 | 7% |
|  |  | Hard | Bnw | 5/100 | 5% |
| Zand | O znw | Zee | O znw | 37/100 | 37% |
|  |  | Strand | O znw | 35/100 | 35% |
|  |  | Kasteel | T znw | 4/100 | 4% |

## D. Reacties van het kunstmatige model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Stimulus | Part-of-Speech | Response | Part-of-Speech | Hmg.? |
| table | C.N. | tables | C.N. | Ja |
|  |  | tray | C.N. | Ja |
|  |  | dining\_room | C.N. | Ja |
| house | C.N. | houses | C.N. | Ja |
|  |  | bungalow | C.N. | Ja |
|  |  | apartment | C.N. | Ja |
| foot | C.N. | feet | C.N. | Ja |
|  |  | pound | C.N. | Ja |
|  |  | inch | C.N. | Ja |
| needle | C.N. | needle | C.N. | Ja |
|  |  | syringe | C.N. | Ja |
|  |  | hypodermic\_needles | C.N. | Ja |
| apple | C.N. | apples | C.N. | Ja |
|  |  | pear | C.N. | Ja |
|  |  | fruit | M.N. | Nee |
| doctor | C.N. | physician | C.N. | Ja |
|  |  | doctors | C.N. | Ja |
|  |  | gynecologist | C.N. | Ja |
| milk | M.N. | dairy | Adj. | Nee |
|  |  | cow\_milk | M.N. | Ja |
|  |  | milk\_powder | M.N. | Ja |
| water | M.N. | potable\_water | M.N. | Ja |
|  |  | sewage | M.N. | Ja |
|  |  | groundwater | M.N. | Ja |
| sand | M.N. | sands | M.N. | Ja |
|  |  | sandy | Adj. | Nee |
|  |  | dirt | M.N. | Ja |
| sugar | M.N. | refined\_sugar | M.N. | Ja |
|  |  | cane\_sugar | M.N. | Ja |
|  |  | cocoa | M.N. | Ja |
| air | M.N. | skies | M.N. | Ja |
|  |  | airspace | M.N. | Ja |
|  |  | airborne | Adj. | Nee |
| cheese | M.N. | cheeses | M.N. | Ja |
|  |  | cheddar | M.N. | Ja |
|  |  | goat\_cheese | M.N. | Ja |
| to send | T.V. | sending | T.V. | Ja |
|  |  | sent | T.V. | Ja |
|  |  | sends | T.V. | Ja |
| to bring | T.V. | bringing | T.V. | Ja |
|  |  | brought | T.V. | Ja |
|  |  | come | I.V. | Ja |
| to find | T.V. | discover | T.V. | Ja |
|  |  | finding | T.V. | Ja |
|  |  | found | T.V. | Ja |
| to take | T.V. | taking | T.V. | Ja |
|  |  | took | T.V. | Ja |
|  |  | takes | T.V. | Ja |
| to hit | T.V. | hitting | T.V. | Ja |
|  |  | hits | T.V. | Ja |
|  |  | struck | T.V. | Ja |
| to invite | T.V. | invites | T.V. | Ja |
|  |  | inviting | T.V. | Ja |
|  |  | invited | T.V. | Ja |
| to skate | I.V. | skating | I.V. | Ja |
|  |  | skates | I.V. | Ja |
|  |  | skated | I.V. | Ja |
| to come | I.V. | coming | I.V. | Ja |
|  |  | go | I.V. | Ja |
|  |  | bring | T.V. | Nee |
| to live | I.V. | living | I.V. | Ja |
|  |  | reside | I.V. | Ja |
|  |  | lived | I.V. | Ja |
| to laugh | I.V. | chuckle | I.V. | Ja |
|  |  | laughing | I.V. | Ja |
|  |  | giggle | I.V. | Ja |
| to stand | I.V. | stands | I.V. | Ja |
|  |  | stood | I.V. | Ja |
|  |  | standing | I.V. | Ja |
| to walk | I.V. | walking | I.V. | Ja |
|  |  | walks | I.V. | Ja |
|  |  | walked | I.V. | Ja |
| dark | Adj. | darkened | T.V. | Nee |
|  |  | darker | Adj. | Ja |
|  |  | gray | Adj. | Ja |
| soft | Adj. | softer | Adj. | Ja |
|  |  | softest | Adj. | Ja |
|  |  | squishy | Adj. | Ja |
| cold | Adj. | chilly | Adj. | Ja |
|  |  | frigid | Adj. | Ja |
|  |  | chill | Adj. | Ja |
| white | Adj. | black | Adj. | Ja |
|  |  | blue | Adj. | Ja |
|  |  | brown | Adj. | Ja |
| sweet | Adj. | sweetest | Adj. | Ja |
|  |  | yummy | Adj. | Ja |
|  |  | buttery | Adj. | Ja |
| hard | Adj. | harder | Adj. | Ja |
|  |  | tough | Adj. | Ja |
|  |  | difficult | Adj. | Ja |
| quickly | Adv. | swiftly | Adv. | Ja |
|  |  | easily | Adv. | Ja |
|  |  | rapidly | Adv. | Ja |
| slowly | Adv. | gradually | Adv. | Ja |
|  |  | steadily | Adv. | Ja |
|  |  | rapidly | Adv. | Ja |
| sadly | Adv. | unfortunately | Adv. | Ja |
|  |  | alas | Interjection | Nee |
|  |  | thankfully | Adv. | Ja |
| now | Adv. | still | Adv. | Ja |
|  |  | already | Adv. | Ja |
|  |  | currently | Adv. | Ja |
| softly | Adv. | calmly | Adv. | Ja |
|  |  | gently | Adv. | Ja |
|  |  | shyly | Adv. | Ja |
| gently | Adv. | tenderly | Adv. | Ja |
|  |  | delicately | Adv. | Ja |
|  |  | softly | Adv. | Ja |