Woordenschatstimulatie bij Meertalige Kinderen met een Voorschoolindicatie: Een Systematische Literatuurreview

Student: M.S. Wentel (S4064674)

Begeleider: Erica Kamphorst, MSc.

2e beoordelaar: Anoek Sluiter-Oerlemans

Rijksuniversiteit Groningen

Faculteit der Gedrags- en Maatschappijwetenschappen

Bachelorwerkstuk Pedagogische Wetenschappen (PABA-A412)

Aantal woorden: 9368

Juni, 2022

rijksuniversiteit groningen

Abstract

The best possible language development is of great importance for children. Due to globalization and migrant flows, among other things, there are more and more young children in the Netherlands who are at risk of language delay because not much Dutch is spoken at home. These children can receive *Voor- & Vroegschoolse Educatie* (VVE; translated as 'Pre- and Early-School Education') to prevent or catch up with delays. The quality of VVE must improve, partly as a result of the COVID-pandemic. One of the areas for improvement is the way vocabulary is taught. Vocabulary development is especially important for multilingual children. In this systematic literature review, an attempt is made to provide professionals within VVE with tools in the design of vocabulary instruction. The PRISMA guidelines served as the basis for the method behind this review. The data is sorted by internal and external factors. More external than internal factors were addressed in the included studies, but the contribution to vocabulary development appears to be about the same. Professionals are advised to differentiate their instruction based on child and word characteristics. For example, the level of vocabulary in L1 and L2, the imageability and concreteness of words and whether the word is a content or function word. Further implications, limitations and recommendations are discussed.

Keywords: Multilingualism, Preschool, Vocabulary, Language development, Instruction

Woordenschatstimulatie bij Meertalige VVE-Kinderen: Een Systematische Literatuurreview

In Nederland stijgt het aantal jonge kinderen met een migratieachtergrond dat thuis een andere taal dan Nederlands te horen krijgt. In de huidige review wordt een overzicht gecreëerd van recente studies naar manieren om woordenschatontwikkeling te stimuleren bij deze kinderen, aangezien woordenschatstimulatie een belangrijke rol kan spelen in het tegengaan van onderwijsachterstanden bij meertalige kinderen. Een optimale taalontwikkeling en het voorkomen van onderwijsachterstanden zijn van groot belang op korte en lange termijn. Het doel van deze review is om professionals zo handvatten te bieden bij woordenschatinstructie.

1.1 Vroege Taalontwikkeling

Taal is de basis van menselijke communicatie en is van belang om gedachten, gevoelens en kennis met elkaar te kunnen delen. Leeftijdsadequaat ontwikkelende kinderen zeggen hun eerste woordjes rond hun eerste levensjaar en rond hun tweede beginnen ze woorden te combineren. Deze vroege taalontwikkeling blijkt een voorspeller te zijn van latere spreek- en schrijfvaardigheid (Preston et al., 2010). Blootstelling aan taal kan volgens Yeung en Werker (2009) niet vroeg genoeg beginnen, het liefst nog voordat kinderen naar school gaan. Er zijn aanwijzingen dat er sprake is van een sensitieve periode in de eerste zeven levensjaren en dat wanneer een kind de taal tijdens deze periode niet goed ontwikkelt, er een groot risico is op problemen met taal in de rest van het leven (Yeung & Werker, 2009).

1.2 Gevolgen van Taalproblemen

Al voordat kinderen op de peuterspeelzaal komen, blijken er tussen hen grote verschillen te zijn in taalvaardigheden. Sommige kinderen halen deze achterstand op hun

leeftijdsgenoten in, maar bij andere kinderen wordt deze alleen nog maar groter tijdens de schoolloopbaan (Herrera et al., 2021). Een dergelijke achterstand kan problemen opleveren op korte en lange termijn. Uit onderzoek van Manning et al. (2019) is bijvoorbeeld gebleken dat dreumesen en peuters die minder woorden gebruiken en/of later beginnen met praten, last hebben van meer en heftigere driftbuien. Dit kan betekenen dat taalvaardigheden invloed kunnen hebben op de mentale gemoedstoestand (Manning et al., 2019).

Op de langere termijn kunnen problemen met taalvaardigheden leiden tot problemen met leren. Het taalniveau bepaalt hoe goed een kind een uitleg kan begrijpen om nieuwe kennis op te doen (Araújo & Costa, 2015). Ook kunnen opdrachten soms moeilijker begrepen worden. Bijvoorbeeld bij het vak rekenen: in Nederland wordt vaak gebruik gemaakt van verhalende rekenopgaven. Als de leerling door taalproblemen de inhoud van de rekensom niet begrijpt, kan hij de rekensom zelf ook niet uitvoeren en kunnen taalproblemen zo tot rekenproblemen leiden (Oortwijn et al., 2005). Woordenschat is hierbij van groot belang, omdat het een essentiële voorspeller is van lees- en luistervaardigheden en academisch succes (Coyne et al., 2022; Pan et al., 2018; Sun et al., 2021). Om specifiek te zijn gaat het hierbij met name om receptieve woordenschat (RW), het kunnen begrijpen van woorden (het tegenovergestelde van expressieve woordenschat (EW), het kunnen produceren van woorden). RW en EW kunnen schriftelijk of oraal zijn (Garcia, 2021).

Uit een longitudinaal onderzoek van Clegg et al. (2005) blijkt dat het hebben van taalproblemen op jonge leeftijd zelfs nog gevolgen kan hebben tijdens de volwassenleeftijd. In het onderzoek zijn 17 jongens met een *Specific Language Impairment* (SLI) vanaf hun kindertijd (4-9 jaar) gevolgd voor een periode van ongeveer 30 jaar. Een SLI is een hardnekkig taalprobleem dat bij 7-8% van voorschoolse kinderen. Het kan lees-, schrijf-, luister- en spreekvaardigheden beïnvloeden. Er is gecorrigeerd op IQ (>70) en auditieve problemen zijn uitgesloten. De gevolgen voor later van het hebben van een SLI op jonge

leeftijd bleken significante problemen met Theory of Mind, het verbaal kortetermijngeheugen en fonologische verwerking te zijn. Daarnaast is een verhoogd risico vastgesteld op moeilijkheden bij het sociaal adaptief vermogen en op psychiatrische stoornissen als volwassene (Clegg et al., 2005). Hoewel dit een hardnekkig taalprobleem is, geeft het wel een indicatie de gevolgen van taalproblemen in het algemeen.

1.3 VVE-Indicatie

Een zo goed mogelijke taalontwikkeling is dus van belang. Wanneer een kind tussen de 2,5 en 6 jaar in Nederland risico loopt op een achterstand kan het een zogenoemde VVE-indicatie krijgen, ook wel voorschoolindicatie genoemd (Gemeente Den Haag, 2021). De focus ligt op taalachterstanden, maar de achterstand kan ook binnen een ander ontwikkelingsdomein bevinden. VVE-geïndiceerde kinderen krijgen Voor- en Vroegschoolse Educatie (VVE) aangeboden en (deels) vergoed. Voorschoolse educatie is voor zogenoemde *doelgroeppeuters* binnen peuterspeelzalen en kinderdagverblijven en vroegschoolse educatie is voor *doelgroepkleuters* binnen groep 1 en 2. Het doel van VVE is om kinderen met zo min mogelijk achterstand in groep 3 te laten instromen en zo onderwijsachterstanden te voorkomen (Inspectie van het Onderwijs, z.d.; Ministerie van Algemene Zaken, 2021a; Ministerie van Algemene Zaken, 2021b). Beginnen met een achterstand zorgt namelijk voor een verhoogde kans op vroegtijdig schoolverlaten, een lagere opleiding, armoede en een minder goede gezondheid (Nederlands Jeugdinstituut, 2022). De VVE-indicatie kan worden verkregen via het consultatiebureau en kan op verschillende gronden worden gegeven. Deze gronden kunnen per gemeente verschillen (Ministerie van Algemene Zaken, 2021a). Voorbeelden van gronden voor een VVE-indicatie zijn wanneer het opleidingsniveau van één of beide ouders lager is dan mbo-2 niveau en/of wanneer er thuis geen of weinig Nederlands gesproken wordt (Gemeente Den Haag, 2021).

1.4 Meertaligheid in Nederland

Er zijn in Nederland steeds meer peuters en kleuters met een andere thuistaal dan Nederlands, die daarmee in aanmerking kunnen komen voor een VVE-indicatie. Momenteel spreekt ongeveer 10% van de gezinnen met kinderen thuis geen Nederlands (Schmeets, 2021). Velen hiervan hebben een migratieachtergrond en zijn afkomstig uit het buitenland. Op 1 februari 2022 had 25,3% van de bevolking een migratieachtergrond (eerste en tweede generatie meegeteld). De landen van herkomst en de duur van verblijf in Nederland variëren zeer – en daarmee de mate waarin de kinderen thuis in aanraking komen met de Nederlandse taal. De meeste migranten in Nederland komen uit Polen, de voormalige Sovjet-Unie, Syrië, Duitsland, Roemenië of Turkije (Centraal Bureau voor de Statistiek, 2022b). In 2021 werden er bijna 80% meer asielverzoeken ingediend dan het jaar ervoor, bijna 25,000 mensen waarvan ruim één op de drie uit Syrië afkomstig (Centraal Bureau voor de Statistiek, 2022a). Door de oorlog in Oekraïne is er daarnaast nog een nieuwe grote vluchtelingenstroom ontstaan, volgens de Volkskrant zelfs de grootste Europese vluchtelingenstroom sinds de Tweede Wereldoorlog (Van der Ploeg, 2022). Deze groep vluchtelingen bestaat vooral uit vrouwen en kinderen (Ministerie van Justitie en Veiligheid, 2022). In 2021 had 27,9% van de 0- tot 5-jarigen een migratieachtergrond, maar dat zal naar verwachting onder andere door deze twee groepen vluchtelingen nog verder stijgen (Centraal Bureau voor de Statistiek, 2021). Er valt een grotere multiculturele instroom in VVE te verwachten.

Door VVE van goede kwaliteit aan te bieden kan de ontwikkeling van kinderen die risico lopen op een achterstand extra gestimuleerd worden en is het doel om kansenongelijkheid tegen te gaan (Inspectie van het Onderwijs, 2020). Onderwijsachterstanden van peuters kunnen niet altijd helemaal worden weggewerkt, maar wel worden teruggedrongen. Een voorwaarde is wel dat de VVE daadwerkelijk van goede kwaliteit is (Leseman & Veen, 2016).

1.5 Kwaliteit van VVE

Om de kwaliteit van VVE op peil te houden controleert de Inspectie van het Onderwijs de VVE op kinderdagverblijven en basisscholen. Uit metingen van 2019 is gebleken dat VVE aan de vereiste kwaliteit voldoet, maar dat er wel ruimte is voor verbetering van het onderwijsleerproces binnen de VVE. Vooral het educatief handelen van medewerkers moet volgens de Inspectie worden verbeterd (Inspectie van het Onderwijs, 2020). Ditzelfde blijkt ook uit eerdere onderzoeken (Slot et al., 2017; Veen & Leseman, 2015). Deze verbetering is vanwege de coronacrisis nog harder nodig, stelt Minister Wiersma in een kamerbrief (Wiersma, 2021). Volgens de onderwijsinspectie valt er bij vroegschoolse educatie winst te behalen op het gebied van feedback leveren door de leerkracht, het bevorderen van de interactie tussen en met kinderen, en het stimuleren van de ontwikkeling van zelfstandig en probleemoplossend werken (Inspectie van het Onderwijs, 2020). Veen en Leseman geven de educatieve kwaliteit binnen Nederlandse voorschoolse educatie een lage tot gemiddelde score. Bij één vierde van de voorscholen kunnen vrijwel alle verdiepende aspecten van kwaliteit binnen het educatief handelen wel verbeterd worden. Eén van de verbeterpunten die specifiek genoemd wordt, is de stimulatie van taalontwikkeling, met name het leren van nieuwe woorden (Veen & Leseman, 2015).

Meertalige peuters lopen een grotere kans op het hebben van een kleinere woordenschat in de meerderheidstaal dan ééntaligen (Mancilla-Martinez et al., 2020). Een focus op het leren van nieuwe woorden kan een belangrijke rol spelen in het tegengaan van onderwijsachterstanden bij meertalige kinderen (Gibson, 2016). Een rijke taalomgeving is hierbij van groot belang en wanneer deze thuis niet geboden kan worden, is aanvullende woordenschatstimulatie extra belangrijk, bijvoorbeeld op de (voor)school (Sun et al., 2021; Wasik, 2010).

1.6 Interne en Externe Factoren

In factoren die woordenschatontwikkeling beïnvloeden, kan onderscheid gemaakt worden tussen interne en externe factoren, vergelijkbaar met *nature* en *nurture*. Interne factoren zijn hierbij eigenschappen van het kind, bijvoorbeeld intelligentie en cognitieve vaardigheden. Externe factoren zijn omgevingsfactoren, zoals de kwaliteit en kwantiteit van taalstimulatie thuis en op school. De relatieve invloed van interne en externe factoren op taalontwikkeling is nog niet op grote schaal expliciet verkend voor jonge tweetalige kinderen. Er is wel veel onderzoek gedaan naar interne en externe factoren, maar dit wordt niet altijd zo benoemd en in niet veel onderzoeken worden deze met elkaar vergeleken. Dit is wel van belang om de taalontwikkeling van deze populatie goed te kunnen begrijpen (Sun et al., 2021).

1.7 Dit Onderzoek

Door de coronacrisis en de toename van het aantal meertalige peuters en kleuters is het belang van het verbeteren van de kwaliteit van VVE nog groter geworden. Meer professionals zoals pedagogisch medewerkers, leraren en orthopedagogen zullen te maken krijgen met deze kinderen. Het doel van dit onderzoek is om hen te inspireren en handvatten te bieden voor hoe ze de woordenschatontwikkeling van deze doelgroep het beste kunnen stimuleren, en hiermee de kwaliteit van VVE te verbeteren. Er is de laatste jaren veel internationaal onderzoek gedaan naar woordenschatstimulatie. Deze literatuurreview biedt een overzicht van de meest recente studies (>2019) naar woordenschatstimulatie bij meertalige kinderen. Een dergelijk overzicht bestaat nog niet, maar wel van oudere studies. Dat eerdere onderzoek wordt in §4.2 vergeleken met dit huidige onderzoek.

De onderzoeksvraag die in dit onderzoek centraal staat, luidt: Met behulp van welke interne en externe factoren kan woordenschatontwikkeling van meertalige jonge kinderen het beste worden gestimuleerd? Een deelvraag hierbij is welke relatief meer invloed uitoefenen, interne of externe factoren.

De onderzoekspopulatie wordt nader gedefinieerd door (twee- of) meertalige kinderen van 2 tot 6 jaar waarbij de thuistaal (L1) anders is dan de meerderheidstaal (L2), niet zijnde kinderen waarbij vrijwillige vreemde taalverwerving het doel is. Dit laatste zou bijvoorbeeld het geval zijn als ouders willen dat hun kind een betere plek op de arbeidsmarkt kan krijgen en de tweede taal thuis niet wordt gesproken. Naar de laatste groep is op verzoek van de minister recentelijk al onderzoek gedaan in het kader van de vraag of meertalige kinderopvang wettelijk mogelijk gemaakt moet worden (Aalberse et al., 2021; Wiersma, 2021).

2. Methode

In deze systematische literatuurreview zijn verschillende studies gesynthetiseerd om een concreet antwoord op de onderzoeksvraag te formuleren. De PRISMA-richtlijnen[1] zijn hierbij als basis gebruikt om het onderzoek zo systematisch en transparant mogelijk te maken (Liberati et al., 2009).

2.1 Zoekstrategie

Op 21 april 2022 is er in de zoekmachines ERIC, PsychINFO en WebOfScience een zoekslag uitgevoerd. De inhoudelijke zoekwoorden waren *woordenschat, ontwikkeling/-stimulatie*, *peuters/kleuters* en *tweetalig/meertalig*, vertaald naar het Engels en aangevuld met synoniemen (zie Bijlage A voor de precieze zoekslag). De synoniemen kwamen van de Thesauri van ERIC en PsychINFO, van Google en van de beschrijvende steekwoorden die bij relevante artikelen waren gegeven in de zoekmachines. In de zoekslag zijn opzettelijk methode-kenmerken weggelaten, omdat zoekmachines niet altijd in de volledige tekst kunnen

[1] PRISMA staat voor Preferred Reporting Items for Systematic Reviews and Meta-Analysis

zoeken. Dit zou betekenen dat studies die geen methode in de titel of abstract benoemden onterecht niet uit de zoekslag naar voren konden komen.

Vier beoordelaars hebben de zoekslag gecontroleerd voordat de definitieve zoekslag is uitgevoerd in de zoekmachines. ERIC is gekozen, omdat deze zoekmachine gericht is op het onderwijsvakgebied. PsychINFO is gericht op psychologie en gekozen omdat taalontwikkeling onder de ontwikkelingspsychologie valt. WebOfScience is gebruikt, omdat dit een multidisciplinaire database is met zeer veel studies, wat breed zoeken mogelijk maakt. In deze zoekmachine mogen maximaal 50 tekens per zoekbalk worden ingegeven, daar is de zoekslag op aangepast (zie Bijlage A). Verder is de zoeksyntax voor de zoekmachines vergelijkbaar.

Tijdens het literatuuronderzoek voor het schrijven van de inleiding kwam nog een aantal artikelen naar voren dat bruikbaar leek voor de review (Gibson, 2016; Scheepers et al., 2021; Wasik, 2010; Yeung & Werker, 2009). Deze zijn los van de zoekslag meegenomen in hetzelfde screeningproces, zoals hierna beschreven.

2.2 Screening en Selectiecriteria

Nadat alle gevonden studies waren geëxporteerd naar de referentiemanager Zotero (Corporation for Digital Scholarship, 2006) en Rayyan, een programma speciaal gericht op systematische literatuurreviews (Ouzzani et al., 2016), zijn de duplicaten handmatig verwijderd. Eerst in Zotero, daarna ter controle in Rayyan, omdat dit programma op een andere wijze op duplicaten scant. Vervolgens zijn de studies gescreend. Hiervoor zijn eerst in- en exclusiecriteria opgesteld om te bepalen welke studies wel en niet in de review meegenomen zouden worden. Zo werden studies (1) geëxcludeerd wanneer ze niet (gedeeltelijk) keken naar interne en/of externe factoren die woordenschat konden bevorderen. Woordenschat moest hierbij een afhankelijke variabele zijn. De context, bijvoorbeeld op school of thuis, maakte in de eerste screeningsfase niet uit. Studies werden (2) alleen

geïncludeerd wanneer de gemiddelde leeftijd van participanten tussen de 2 en 6 jaar viel, omdat dit de leeftijd is waarop kinderen een VVE-indicatie kunnen hebben. Studies werden (3) alleen meegenomen wanneer kinderen die meertalig opgroeien de hoofdfocus of een losse subgroep in de studie waren. (4) Er mocht geen sprake zijn van een specifieke klinische populatie, bijvoorbeeld bestaande uit autistische kinderen, behalve wanneer er ook een niet-klinische controlegroep was. Kinderen met een taalachterstand werden niet gezien als klinische populatie, maar kinderen met een taalstoornis (bijvoorbeeld SLI) wel, omdat deze kinderen visuele en verbale informatie anders verwerken dan leeftijdsgenoten en dus niet goed vergelijkbaar zijn (Gladfelter et al., 2019). (5) Alleen wanneer (een deel van) de participanten thuis een andere taal dan de meerderheidstaal sprak(en), werd de studie geïncludeerd. Dus niet wanneer er op vrijwillige basis een vreemde taal werd aangeleerd. (6) De meerderheidstaal moest een Germaanse taal zijn, zoals Nederlands en Vlaams. Hieronder vallen ook (varianten van) het Engels, Duits (ook Jiddisch), Fries, Afrikaans en de Scandinavische talen (Deens, Noors, Zweeds, IJslands en Faeröers; Hawkins, 2018). (7) De studie zelf moest in het Nederlands, Vlaams of Engels geschreven zijn en (8) mocht niet ouder zijn dan 2019. Ook werden studies (9) geëxcludeerd die niet getoetst waren door vakgenoten (*peer-reviewed*). Dit betekent dat enkel wetenschappelijke literatuur is meegenomen, bijvoorbeeld geen rapporten. (10) Wanneer studies niet als volledige tekst beschikbaar waren, werden de auteurs gemaild en als er na twee weken geen antwoord was, werden ook deze studies geëxcludeerd. Wanneer er kosten aan verbonden waren voor het verkrijgen van de tekst, werd de studie om budgetredenen ook geëxcludeerd. (11) Alleen empirische studies zijn geïncludeerd, dus geen meta-analyses en reviews. (12) Studies met minder dan 15 participanten werden geëxcludeerd, omdat de taalstimulatie van een kleine groep niet goed te generaliseren is naar een grotere groep kinderen.

Op basis van bovenstaande selectiecriteria is een protocol opgesteld (zie Bijlage B) met bijbehorende labels die in Rayyan door de eerste en tweede beoordelaar gebruikt konden worden om aan te geven waarom een studie geëxcludeerd werd.

Het protocol is proefgedraaid op de titels en abstracts van 10 studies en naar aanleiding hiervan zijn verschillende criteria toegevoegd. Ten eerste (13) dat (een gedeelte van) de factoren die in de studie werden onderzocht ook daadwerkelijk beïnvloed moesten kunnen worden. Dit is vanwege het onderzoeksdoel om professionals handvatten te kunnen bieden voor woordenschatstimulatie; SES of het hebben van broers/zussen zijn bijvoorbeeld geen factoren die hulpverleners kunnen veranderen. Als laatste is in dit stadium toegevoegd (14) dat studies geëxcludeerd worden wanneer ze woordenschat niet specifiek meten, maar bijvoorbeeld indirect via een andere taalvaardigheid. In sommige studies werd bijvoorbeeld wel gesproken over woordenschat, maar werd bijvoorbeeld alleen de leesvaardigheid (waar woordenschat voor nodig is) gemeten.

Met dit aangepaste protocol werden verschillende rondes van screening uitgevoerd. Criteria 8, 9 en 10 zijn tijdens het uitvoeren van de zoekslag in de zoekmachines al zo veel mogelijk toegepast. In Rayyan is de optie *Blind ON* aangezet, waardoor de twee beoordelaars niet meer van elkaar konden zien welke teksten de ander in- of excludeerde. In de eerste screeningsronde is op basis van titel en abstract gescreend en in de tweede ronde op basis van de volledige tekst. Het was van belang om in de eerste ronde relatief mild te screenen; een studie werd alleen geëxcludeerd wanneer zeker was dat deze niet aan de selectiecriteria voldeed. Bij twijfel, zoals door ontbrekende informatie, werd de studie bewaard tot de screening van de volledige tekst, waarbij wel kritisch beoordeeld is. Het protocol is in deze ronde verder aangescherpt, omdat de hoeveelheid overgebleven studies te groot was. Studies moesten vanaf nu ook (15) gebruik maken van een voor- en nameting, zodat de kwaliteit werd verhoogd door middel van het longitudinale aspect. (16) Nu werden enkel onderzoeken

meegenomen die op een (voor)schoolse instantie (inclusief kinderdagverblijf, etc.) hebben plaatsgevonden, dus niet thuis. Deze afweging is gemaakt op basis van het feit dat het aantal kinderen dat op een (voor)schoolse instantie woordenschatinstructie krijgt groter is dan het aantal kinderen dat thuis hulp ontvangt (onder andere door de leerplicht). Deze review poogt zo veel mogelijk professionals te kunnen helpen.

Een voorbeeld van een studie die in deze screeningsfase afviel is die van Rolletschek (2020). In de abstract werd gesproken over de effecten van geur op woordenschatontwikkeling van jonge tweetalige kinderen. Pas in de volledige tekst bleek dat het om kinderen van gemiddeld 9 jaar ging, wat niet aansluit op de doelpopulatie van dit onderzoek (Rolletschek, 2020).

Na deze fase zijn een referentie- en citatiecheck uitgevoerd via de geïncludeerde artikelen, om meer gerelateerde artikelen te vinden. Deze zijn op dezelfde wijze gescreend als hierboven beschreven. In Figuur 1 is in een PRISMA flow diagram te zien hoeveel studies er in elke fase van het onderzoek zijn meegenomen.

2.3 Dataextractie Procedure

Om de geselecteerde studies te synthetiseren, zijn ze alle grondig gelezen en geanalyseerd. Tijdens het lezen zijn verschillende aspecten, waaronder steekproefgegevens, interventiegegevens en kernuitkomsten, van de studies genoteerd en in de schema's geplaatst (zie Tabel 1 en Bijlage D en E). Elke studie is gecodeerd in steekwoorden, waarmee uiteindelijk het verband tussen studies is vastgelegd in een boomdiagram (zie Figuur 3). Van een aantal andere gegevens zijn kwantitatieve maten zoals gemiddeldes en percentages berekend, bijvoorbeeld hoe groot de gemiddelde steekproefgrootte was en hoeveel procent van de studies een bepaalde meetmanier heeft gebruikt (zie §3.1).

2.4 Kwaliteitsbeoordeling

Om de kwaliteit van geïncludeerde studies te bepalen is een tool ontwikkeld (zie Bijlage C) met onder andere items uit de Newcastle-Ottawa schaal (Wells et al., 2011) – aanbevolen door Deeks et al. (2003) - en de tool uit het artikel van Beckers et al. (2021). Voorbeelden van items die toegevoegd zijn, zijn of studies de power en effectgrootte (EG) benoemen. Statistische power is een maat die aangeeft in hoeverre een studie onderzochte verbanden in de populatie kan waarnemen, dus in hoeverre wat een steekproef zegt over de populatie waarheidsgetrouw kan zijn. De EG geeft een indicatie van het nut voor de praktijk van een uitkomst van een studie. Een onderzoek dat bijvoorbeeld een significante uitkomst heeft maar een kleine EG (en dus weinig effect), voegt in de praktijk weinig toe. In dit onderzoek wordt gekeken naar EG in de vorm van Cohen's *d*, omdat deze maat veelgebruikt is binnen de sociale wetenschap (Fritz et al., 2012) en het studies beter vergelijkbaar zijn wanneer dezelfde maat gebruikt wordt. Bij vraag 8 en 9 is er voor een minimale betrouwbaarheid van .70 voor meetinstrumenten gekozen, op basis van regels van de COTAN. Volgens COTAN is dit de minimale betrouwbaarheid waar een test aan moet voldoen om uitspraken te kunnen doen op groepsniveau (Evers et al., 2010).

De initiële kwaliteitsbeoordeling is eerst op 10 teksten toegepast en verbeterd. Zo zijn bijvoorbeeld bij vraag 8 en 9, die de betrouwbaarheid van de gebruikte meetinstrumenten beoordelen, meer opties toegevoegd om goed onderscheid te kunnen maken in de kwaliteit van de meetmanier. In plaats van 4 opties met stappen van 15 (a: $r \geq .85$, b: $r \geq .70$, …), was dit veranderd naar 6 opties met stappen van 10 (a: $r \geq .90$, b: $r \geq .80$, …).

Per aspect kon een studie één tot drie punten verdienen. Bij meerdere antwoordmogelijkheden werd het laagste aantal punten toegekend. In totaal zijn er 21 punten te verdienen. Niet alle vragen waren toepasbaar op elke studie, in dat geval is *n.v.t.* ('niet van toepassing') genoteerd en het maximaal aantal punten daarop aangepast. Per studie is het

percentage van het maximum puntenaantal berekend (zie Bijlage E). Een kwaliteitsscore van <40% werd als zeer laag beoordeeld, 40-50% als laag, 50-60% als vrij laag, 60-70% als matig, 70-80% als prima en 80-90% als goed. Van de studies is het aantal punten per aspect in Bijlage D weergegeven.

2.5 Tweede Beoordeling

Bij elke fase van dit onderzoek heeft een tweede beoordelaar steekproefsgewijs een controle uitgevoerd (10% bij de titel en abstract screening en 20% bij de overige stappen). De interbeoordelaars-betrouwbaarheid (IBBH) was bij de titel en abstract screening κ=.39, bij de screening van de volledige studies κ=.44, bij de data extractie procedure κ=.47 en bij de kwaliteitsbeoordeling κ=.36. Deze kappa scores geven aan dat de overeenkomst tussen beoordelaars gematigd is (bij kappa .21 tot .40 is er sprake van een *fair agreement* en .41 tot .60 is *moderate agreement*[2]*;* Agresti, 2018). De studies waarbij de twee beoordelaars het niet eens waren, zijn bediscussieerd. De IBBH is berekend in Cohen's Kappa, omdat deze corrigeert voor de gok kans (Cohen, 1968).

3. Resultaten

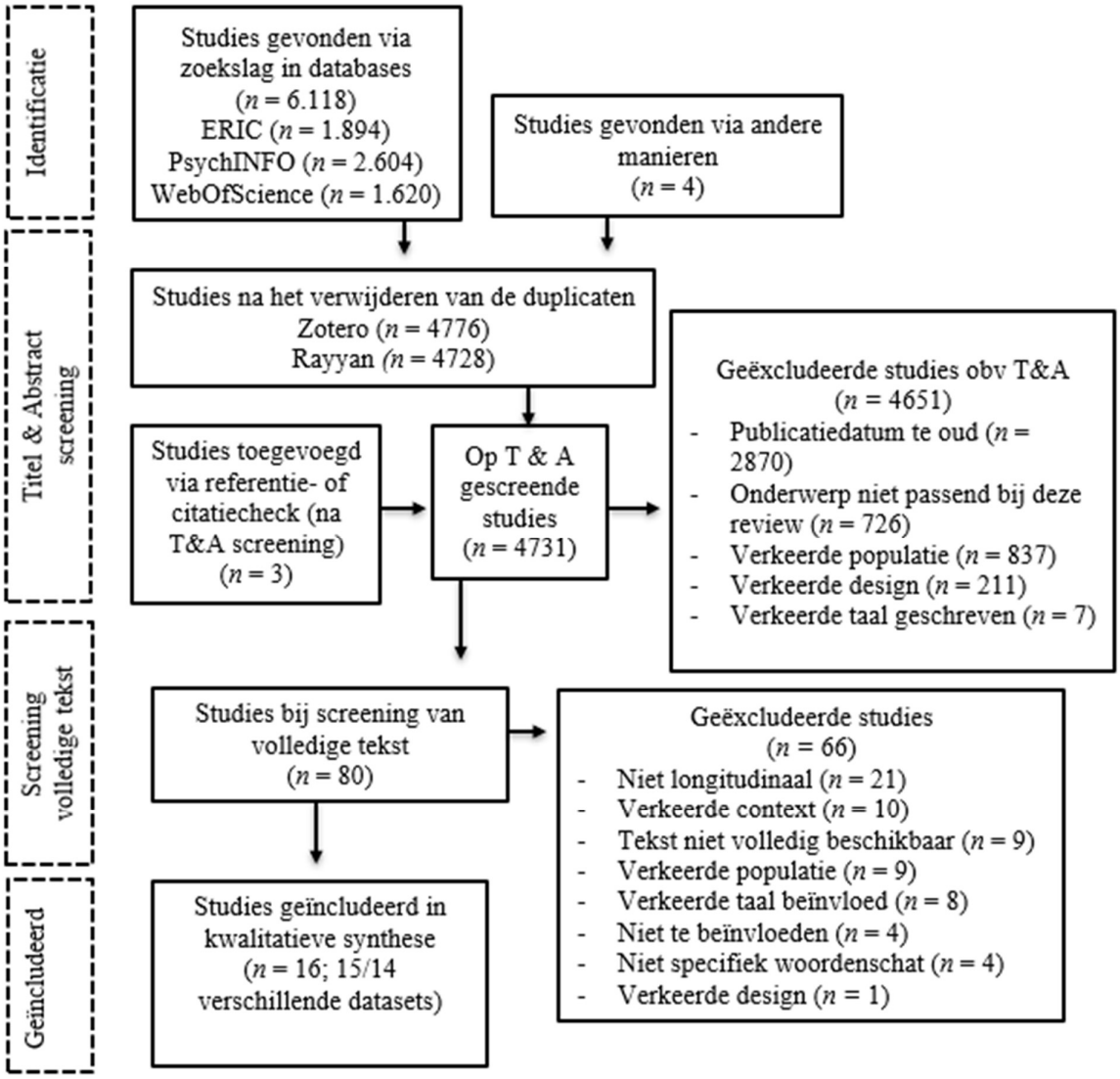
Uit de zoekslag in de drie zoekmachines kwamen in totaal 6.118 hits, waaraan 4 studies bij het schrijven van de inleiding met de hand zijn toegevoegd. Verwijderen van duplicaten leverde 4728 unieke studies op. Deze studies zijn op titels en abstracts gescreend, samen met 3 studies die na deze screening via referentie- of citatiecheck zijn toegevoegd. Er bleven 80 studies over voor screening van de volledige tekst. Van deze studies bleken er 16 geïncludeerd te kunnen worden (zie Figuur 1).

---

[2] De Engelse termen zijn in dit geval gegeven, omdat het onderscheid ertussen zich niet goed laten vertalen naar het Nederlands

Figuur 1

PRISMA Flow Diagram



3.1 Studie Kenmerken

Kenmerken van de gesynthetiseerde studies zijn gepresenteerd in Tabel 1 en Bijlage E. In Tabel 1 staan de steekproefgegevens beschreven. Steekproefgroottes varieerden van 43 tot 1151 personen en gemiddeld hadden studies 278 participanten. Deze participanten vielen, voor zover bekend, binnen een range van 3 tot 7;2 jaar oud en waren gemiddeld 4;8. De meeste onderzoeken (68.8%) zijn uitgevoerd binnen de VS, met Engels als meerderheidstaal

en Spaans als meest voorkomende thuistaal. De overige onderzoeken vonden plaats in Nederland (12.5%), Duitsland (6.3%), Noorwegen (6.3%) en Singapore (6.3%).

Tabel 1
Kenmerken Participanten (Studies Alfabetisch Geordend)

Studie	*n*	*Leeftijd*	*Plek van afname, meerderheidstaal*	*Thuistaal*
Baker et al., 2021	351	"kindergarten"[a]	VS, Engels	Spaans
Garcia, 2021	645	*M* = 3;10	VS, Engels	Spaans
Grøver et al. (2020)	429	*M* = 4;5 Range: 3-4	Noorwegen, Noors	Gevarieerd (>11 talen)
Hadley et al., 2021	118	*M* = 4;10	VS, Engels	Gevarieerd (o.a. Spaans, Arabisch, Nepals)
Janssen et al., 2019	85	*M* = 4;11 Range: 4-6;3	Nederland, Nederlands	Gevarieerd (> 10 talen)
Leeuwestein et al., 2021	67	*M* = 4;9 Range: 4-6	Nederland, Nederlands	Turks
Miles et al., 2019	81	*M* = 5;9 Range: 5;5 – 7;2	VS, Engels	Spaans, Chinees, Russisch
Pollard-Durodola et al., 2021	334	*M* = 4;8	VS, Engels	Spaans
Reilly et al., 2020	1151	*M* = 3;11 Range: 3-5	VS, Engels	Spaans
Rivera Pérez et al., 2019	43	*M* = 4;2 Range: 3;1 – 5;0	VS, Engels	Spaans
Silverman et al., 2019	196	"kindergarten"[a]	VS, Engels	Gevarieerd
Sun et al., 2021	736	*M* = 4;9 Range: 4;1 – 5;5	Singapore, Engels	Mandarijn
Wong et al., 2021	50	*M* = 4;5 Range: 4 – 6	VS, Engels	Gevarieerd (meeste Spaans)
Wong & Neuman, 2021	51	*M* = 4;7	VS, Engels	Spaans & Haïti-Creools
Wong & Samudra, 2019	43	*M* = 4;7	VS, Engels	Spaans, Haïti-Creools, Fula

Noot. Op basis van ongeveer hetzelfde aantal participanten (*n* = 50 en *n* = 51) met ongeveer dezelfde gemiddelde leeftijd (4;5 en 4;7), dezelfde locatie (het noordoosten van de VS),

hetzelfde publicatiejaar (2021) en het feit dat Wong en Neuman aan beide onderzoeken meewerkten, wordt verondersteld dat de onderzoeken van Wong en Neuman (2021) en Wong et al. (2021) zijn gebaseerd op (ongeveer) dezelfde steekproef. Ook is er een kans dat er overlap is met het onderzoek van Wong en Samudra (2019), aangezien dit ook door één van dezelfde onderzoekers is uitgevoerd in dezelfde regio en de gemiddelde leeftijd ook ongeveer gelijk is (*M* = 4;7). Echter, het publicatiejaar is twee jaar eerder en de steekproef is kleiner (43 i.p.v. 50/51). Dit betekent dat er van de 16 onderzoeken waarschijnlijk 15 of 14 op opzichzelfstaande datasets zijn gebaseerd.

[a]Niet gespecificeerd. In de VS is "kindergarten" voor 4- tot 6-jarigen.

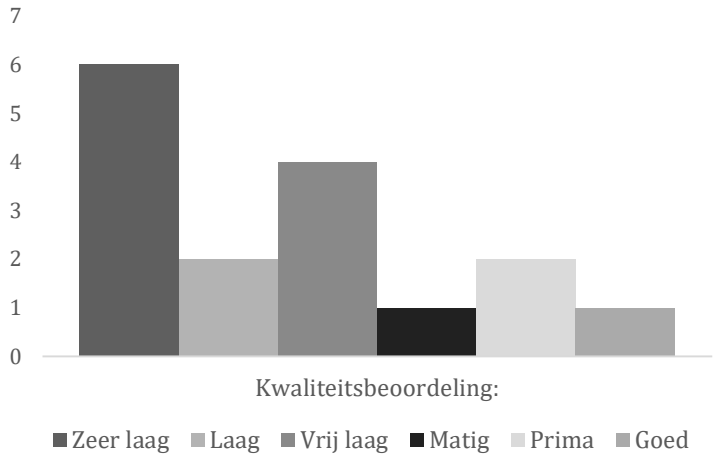
In Bijlage E zijn de inhoudelijke kenmerken van de gesynthetiseerde studies weergegeven. Hieruit blijkt dat de meeste studies zich richtten op receptieve woordenschat (RW) (50%, *n* = 8), daarnaast richtte 12.5% (*n* = 2) zich op expressieve woordenschat (EW) en was 37.5% (*n* = 6) gericht op beiden. Om RW te meten is het vaakst de 4[e] editie van de Peabody Picture Vocabulary Test (PPVT-4) gebruikt (in 62.5% van de studies, *n* = 10) en/of een test die door de onderzoekers zelf is ontworpen (in 75% van de studies, *n* = 12). Soms werd gebruik gemaakt van meerdere tests, bijvoorbeeld de PPVT als voormeting en een zelfontworpen test om targetwoordenschat (woordenschat van woorden waar de focus op ligt tijdens instructie) te meten als nameting. De betrouwbaarheid van alle tests lag tussen de .60 en de .99 en is in verschillende vormen gemeten (Interbeoordelaars-betrouwbaarheid (IBBH), test-hertest betrouwbaarheid (T-H) en split-half betrouwbaarheid (S-H)). De duur tussen voor-, nameting en follow-up varieerde van één dag tot 3 jaar. Bij 5 studies (31.3%) vonden voor- en nameting op dezelfde dag plaats. Effectgroottes in Cohen's *d* waren maar in 2 studies gegeven (zie Grøver et al. (2020) en Silverman et al. (2019) in Tabel 1).

Aspecten van de effectgroottes, betrouwbaarheid en duur zijn ook meegenomen in het kwaliteitsassessment van de studies (zie Bijlage E voor de totale kwaliteitsscore, Bijlage C

voor het gebruikte beoordelingsformulier en Bijlage D voor de uitwerking per studie). De verdeling van de kwaliteitsbeoordeling die de geïncludeerde studies kregen is te zien in Figuur 2. Gemiddeld scoorden studies 50.7% van het totaal aantal punten.

Figuur 2

Verdeling van de Kwaliteit van Geïncludeerde Studies



3.2 Manieren van Woordenschatstimulatie: Interne en Externe Factoren

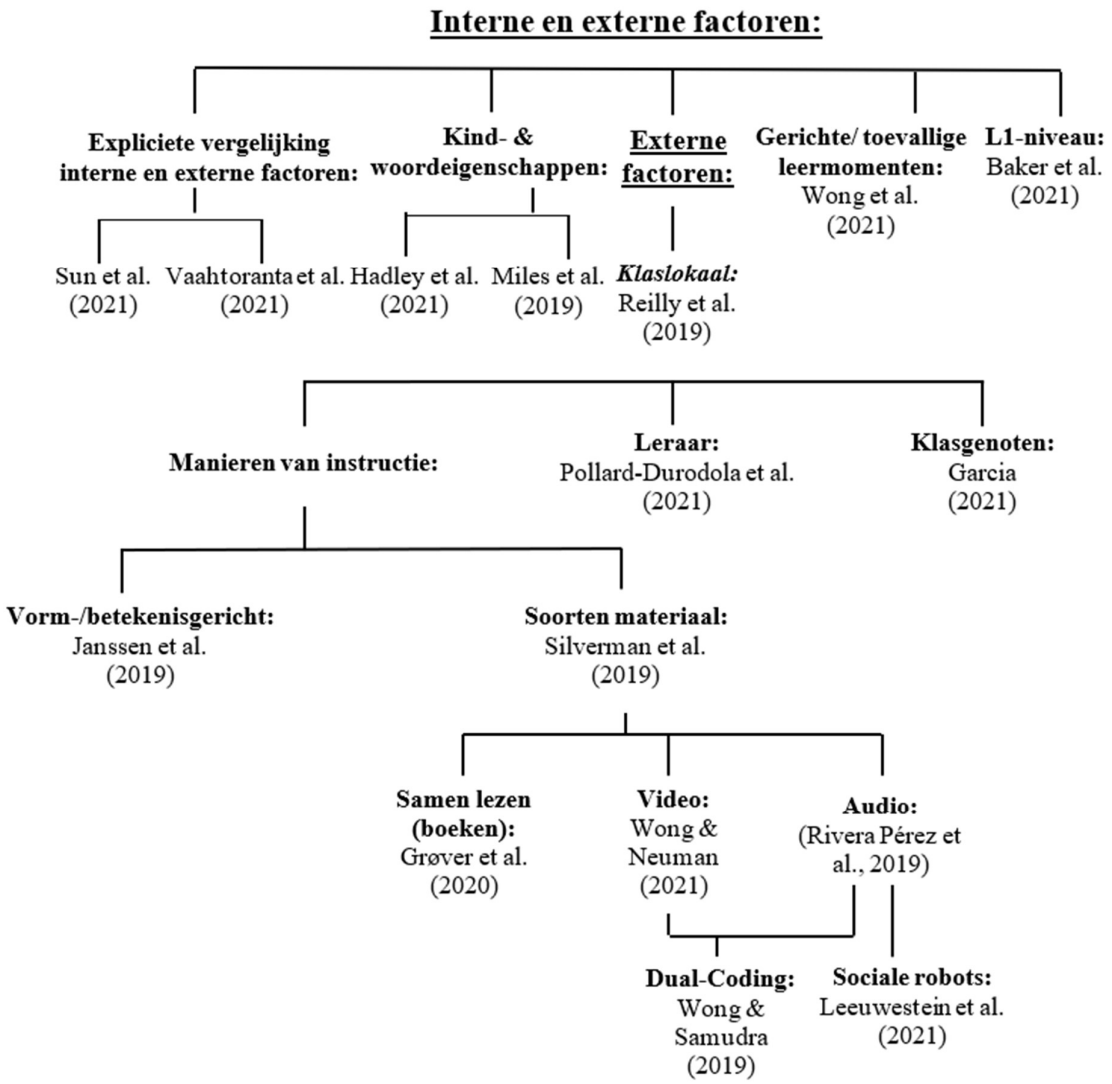
Het verband tussen de verschillende artikelen is te zien in Figuur 3 en wordt in onderstaande bespreking als leidraad voor de opbouw gebruikt en verder verduidelijkt.

3.2.1 Interne en Externe Factoren

Van de 16 geïncludeerde studies zijn er 6 studies die interne en externe factoren bevatten. Twee hiervan vergelijken expliciet interne en externe factoren met elkaar. De eerste studie die interne en externe factoren vergelijkt is die van Sun et al. (2021), waarin als externe factoren het aantal boeken thuis, de leeshoeveelheid thuis, de SES (in de vorm van opleidingsniveau van ouders) en het opleidingsniveau en de ervaring van de leraar zijn vergeleken met de interne factoren werkgeheugen, non-verbale intelligentie, leer-gerelateerde sociale vaardigheden, leeftijd en gender. Uit deze studie bleek dat interne en externe factoren beide een significante bijdrage leveren aan RW-ontwikkeling, en interne factoren net iets meer bijdroegen (53%). Met name non-verbale intelligentie, sociale vaardigheden, leesinput

Figuur 3

Samenhang tussen geanalyseerde studies



thuis in de meerderheidstaal en het opleidingsniveau van de moeder voorspelden RW. Daarna volgden leeftijd en het aantal maanden dat kinderen op de opvang hadden doorgebracht. De overige factoren hadden beduidend minder invloed en waren niet significant (Sun et al., 2021).

De tweede studie die externe en interne factoren vergelijkt is die van Vaahtoranta et al. (2021). Ook in deze studie wordt gekeken naar taalblootstelling thuis (frequentie en duur). De rol van deze externe factor wordt vergeleken met die van het fonologisch

kortetermijngeheugen als interne factor bij RW- en EW-ontwikkeling. Taalvaardigheid van ouders is hierbij toegevoegd als co-variaat. Uitkomsten gaven aan dat RW en EW werden voorspeld door L2-vaardigheid van de moeder, fonologisch kortetermijngeheugen en de duur van de blootstelling aan L2 (op lange termijn). De correlaties waren ongeveer even sterk, wat betekent dat volgens deze studie externe en interne factoren beide ongeveer even belangrijk lijken. Vaahtoranta et al. (2021) vonden ook dat kwaliteit en kwantiteit van de actuele dagelijkse blootstelling aan de meerderheidstaal en thuistaal buiten school geen significante voorspeller was. Gebruik van L1 heeft dus ook geen significante negatieve invloed op L2. Deze resultaten spreken die van Sun et al. (2021) tegen, omdat die wel stellen dat blootstelling aan L2 thuis significant effect heeft op woordenschatontwikkeling. Wel moet gezegd worden dat de invloed in deze studie bij EW in de buurt kwam van significant (Vaahtoranta et al., 2021).

In studies van Hadley et al. (2021) en Miles et al. (2019) is gekeken naar de invloed van verschillende woord- en kind-factoren op woordenschatsontwikkeling. In het onderzoek van Hadley et al. (2021) zijn als kind-factoren (interne factoren) de EW-voormeting van de targetwoorden, de leeftijd in maanden, het executief functioneren, de algemene RW en SES meegenomen. De woordkenmerken (externe factoren) in deze studie betroffen beeldbaarheid (het gemak waarmee iemand een beeld kan vormen bij het woord), concreetheid (in hoeverre het woord iets beschrijft dat waargenomen kan worden; dit is niet hetzelfde als beeldbaarheid/abstractheid) en vorm (in hoeverre het voorwerp omlijnd kan worden of de actie te volgen is). Uit het onderzoek bleek woordenschatontwikkeling te variëren op basis van de beeldbaarheid van woorden en het executief functioneren (met name werkgeheugen en zelfbeheersing) en de algemene woordenschat van een kind. Buiten het verband met EW-ontwikkeling is onderzocht of bepaalde woorden makkelijker te leren zijn voor bepaalde

kinderen. Effecten van beeldbaarheid waren groter voor oudere kinderen en kinderen met hogere RW-scores. Echter, de interactie-effecten zijn erg klein (Hadley et al., 2021).

In de studie van Miles et al. (2019) is gekeken naar welke taalvaardigheden het leren van inhouds- en functiewoorden voorspelden. Dit is gedaan via een interventie waarbij enkele woorden zonder context via flashcards werden aangeleerd. Fonetisch bewustzijn schijnt het leren van inhoudswoorden te voorspellen en de voormeting op RW scheen de groei in functiewoorden te voorspellen. De laatstgenoemde relatie wordt gemodereerd door alfabetkennis (Miles et al., 2019).

In hoeverre het beter werkt om gericht en intentioneel iets aangeleerd te krijgen of wanneer leerervaringen zich meer toevallig voordoen (een externe factor) en of dit verschilt voor kinderen met een hoge of lage begin L2-RW (een interne factor), is waar Wong et al. (2021) naar kijken. Een narratieve context, waarbij leren het meest 'toevallig' was, was voor beide subgroepen minder effectief wat betreft ontwikkeling van simpele woordenschat dan een verklarende en participatieve context met meer gerichte leermomenten. Voorbeelden hiervan zijn visuele en geluidseffecten en herhalingen bij het verklarende type en een karakter in de video dat een soort gesprek met de kijker voert in het participatieve type. Wat betreft ontwikkeling van dieper woordbegrip was alleen het verklarende type instructie bij de groep met hogere begin L2-RW een significante voorspeller (Wong et al., 2021).

Een studie die naar een interne factor kijkt door middel van een externe factor is die van Baker et al. (2021). Zij onderzochten door middel van een interventie of het niveau van L1 invloed heeft op RW en EW van L2. Boven op een basisinterventie werd aan de experimentele groep aanvullende instructie aangeboden, waarbij woorden uit de basisinterventie werden verdiept. Alle kinderen hadden baat bij de interventie, met name kinderen met een grotere L1-RW. Die kinderen hadden gemiddeld significant meer groei in

Engelse RW en EW. Effecten waren 1 jaar later nog steeds aanwezig in de EW (Baker et al., 2021).

3.2.2 Externe Factoren

10 studies in deze review keken naar de invloed van één of meerdere externe factoren op woordenschatontwikkeling. Al deze externe factoren bevinden zich binnen het klaslokaal. Reilly et al. (2020) vergelijken de invloed van verschillende klaslokaalkenmerken op de EW- en RW-scores. Kinderen uit klaslokalen met genoeg instructie, geen over-structurering en ruimte voor het kind om activiteiten te initiëren hadden hogere EW- en RW-scores, ook na de follow-up. Meer emotionele steun ging samen met meer groei in RW (Reilly et al., 2020).

Binnen het klaslokaal zijn klasgenoten een duidelijke externe factor. Garcia (2021) onderzocht wat het effect is van het woordenschatniveau van klasgenootjes op RW. De studie stelde dat er meer groei was wanneer klasgenootjes een hoger niveau van RW hadden. Variabiliteit tussen klasgenoten hing niet significant samen met groei in woordenschat (Garcia, 2021).

Misschien wel de meest onderzochte externe factor is de leraar en zijn eigenschappen en didactische vaardigheden. In aanvullend op Sun et al. (2021), die keken naar het opleidingsniveau en de ervaring van de leraar, onderzochten Pollard-Durodola et al. (2021) de sociale validiteit van leraren bij een verdiepende samen-lees-interventie (*content enriched shared book reading*) en de invloed daarvan op woordenschat van tweetalige kinderen. Met sociale validiteit wordt de tevredenheid met een bepaalde interventie bedoeld van degene die deze implementeert, bijvoorbeeld of deze als noodzakelijk en effectief gezien wordt. Een positieve mening over de interventie hing significant samen met meer groei in target-RW en EW op 3/4 testen (op gebied van definitie- en conceptkennis; Pollard-Durodola et al., 2021).

Een leraar kan binnen het klaslokaal verschillende manieren van instructie toepassen. Naast het aanpassen van instructie op basis van woord- en kind-eigenschappen, zoals bij de

studies van Hadley et al. (2021) en Miles et al. (2019) is gedaan, en of leermomenten gericht of toevallig plaats moeten vinden, zoals onderzocht in de studie van Wong et al. (2021), kan ook worden gekeken of bij instructie van verschillende woorden beter op de vorm of betekenis gefocust kan worden. Met deze vraag hebben Janssen et al. (2019) zich beziggehouden. In hun onderzoek hebben ze een leesinterventie onderzocht, waarbij het weekprogramma voor de twee vergeleken groepen gelijk was, maar bij één groep werd meer op fonologische aspecten (de vorm) gelet en bij de andere meer op semantische aspecten (de betekenis), al werden vorm en betekenis bij beide instructies wel benoemd. Er werd een gelijke toename in target-RW gemeten. Kinderen die vormgerichte instructie kregen, hadden een significant grotere toename in semantische kennis van targetwoordenschat. Deze toename werd indirect voorspeld door het niveau waarop een kind in woordparen die qua klank erg op elkaar lijken, kon leren. Dit effect werd vervolgens weer volledig gemedieerd door algemene RW (Janssen et al., 2019).

In de studie van Silverman et al. (2019) is onderzocht welke soorten materiaal het beste gebruikt kunnen worden. Het gebruik van video lijkt effectiever te zijn dan digitale tekst en tekst op papier voor de RW en EW. Wel is de sterkte van het effect afhankelijk van verschillende factoren zoals leeftijd (wel een interne factor) en de soort woordenschat, namelijk RW of EW (Silverman et al., 2019).

Hoewel dus mogelijk minder effectief dan het gebruik van video's, onderzochten Grøver et al. (2020) in hun onderzoek een interventie die gebruik maakt van een specifieke vorm van tekst op papier, namelijk het collectief lezen van boeken met een bijbehorend vrij te volgen script. Deze interventie bleek volgens Grøver et al. significant effect te hebben op target-RW en -EW, maar niet op algemene woordenschat (Grøver et al., 2020).

Welk onderzoek wel gebruik maakt van video's is dat van Wong en Neuman (2021). Mediaclips van Sesamstraat waarbij verschillende technieken werden gebruikt, werden

bekeken (expliciete definities, herhalen en visuele en geluidseffecten). De RW nam significant toe. Kinderen met meer L2-blootstelling thuis boekten over het algemeen de meeste vooruitgang, met name bij video's met expliciete definities en geluids- en visuele effecten. Dit sluit aan op Sun et al. (2021) die stelden dat meer L2-stimulatie thuis samenging met hogere RW-scores en spreekt het onderzoek van Vaahtoranta et al. (2021) tegen, die dit als niet-significant bevonden. Echter, kinderen die thuis minder werden blootgesteld aan L2 scoorden significant hoger wanneer er gebruik werd gemaakt van herhaling dan kinderen met meer L2-blootstelling (Wong & Neuman, 2021).

Buiten dat de studies deels door dezelfde onderzoekers zijn uitgevoerd, maakt de studie van Wong en Samudra (2019) ook gebruik van video's. Deze studie draait echter rondom het begrip *dual-coding.* Volgens deze theorie kan woordenschatontwikkeling extra gestimuleerd worden wanneer visuele en auditieve informatie tegelijk worden aangeboden, omdat zo gebruik gemaakt wordt van twee subsystemen in het brein, waardoor een completer beeld van een woord opgeslagen kan worden. Woorden kunnen hierdoor aan mentale plaatjes gelinkt worden. Uit het onderzoek bleek dat met name kinderen met ouders met lage L2-taalvaardigheid of uit een taalarme omgeving inderdaad woorden beter leerden via een scherm wanneer deze *dual-coded* werden aangeboden. De taalkwaliteit thuis is dus van belang, wat in lijn is met de onderzoeken van Sun et al. (2021) en Wong en Neuman (2021), maar dat van Vaahtoranta et al. (2021) tegenspreekt. Visuele aandacht bleek ook een significante voorspeller voor de woordenschat nameting, maar alleen wanneer er aan *dual-coding* werd gedaan (Wong & Samudra, 2019).

Iets dat lijkt op *dual-coding* is *audio prompting.* Hierbij worden van tevoren opgenomen woorden of zinnen toegevoegd aan het woordleerproces. Rivera Pérez et al. (2019) stelden vast dat dit significant effect had op het woord benoemen van tweetalige

kinderen. Wanneer instructie tweetalig werd aangeboden had dit ook significant effect op het kunnen definiëren van woorden, ook na de follow-up (Rivera Pérez et al., 2019).

Ook Leeuwestein et al. (2021) maken gebruik van *audio prompting,* al benoemen ze dit zelf niet expliciet. Ze maakten gebruik van sociale robots, waarvan één enkel Nederlandse woorden aanleerde en één Nederlandse woorden ook kon vertalen naar het Turks. Beide robots zorgden voor significante groei in woordenschat, maar er is geen significant verschil in effect gevonden tussen de twee. Kinderen vonden het wel fijner om met de tweetalige robot te werken (Leeuwestein et al., 2021). Wat opvallend is aan de resultaten van deze studie, is dat deze tegenstrijdig zijn met die van de studie van Rivera Pérez et al. (2019), waarin de groep die tweetalige instructie kreeg aangeboden wel hogere scores had dan de groep die ééntalige instructie kreeg.

Er zijn geen studies geïncludeerd die enkel interne factoren onderzochten.

4. Discussie

In het huidige onderzoek is met behulp van een systematische literatuurreview toegewerkt naar het beantwoorden van de hoofdvraag met behulp van welke interne en externe factoren de woordenschatontwikkeling van meertalige VVE-geïndiceerde kinderen het beste kan worden gestimuleerd. Een deelvraag hierbij is welke relatief meer invloed uitoefenen, interne of externe factoren.

4.1 Beantwoording Onderzoeksvraag en Deelvragen

Op basis van de 16 geïncludeerde studies in deze review zou geconcludeerd kunnen worden dat woordenschatontwikkeling een complex en multidimensionaal proces is en dat er verscheidene interne en externe factoren van invloed zijn op de RW- en EW-ontwikkeling. De interne factoren die op basis van deze review een bijdrage lijken te leveren aan woordenschatontwikkeling zijn: leeftijd, non-verbale intelligentie, sociale vaardigheden, het aantal maanden dat het kind naar de opvang gaat, fonologisch kortetermijngeheugen,

fonetisch bewustzijn, werkgeheugen, zelfbeheersing, algemene begin-RW in L1 en L2 en alfabetkennis. Externe factoren die uit de geïncludeerde studies naar voren kwamen als waarschijnlijk relevant voor woordenschatontwikkeling, zijn onder andere factoren in het klaslokaal, zoals veel instructie, geen over-structurering, veel ruimte voor door het kind geïnitieerde activiteiten, hoge emotionele steun, een hoog niveau van klasgenootjes en een hoge sociale validiteit van interventies bij leraren. Bij instructie kunnen waarschijnlijk het beste gerichte leermomenten worden gecreëerd met tweetalige vorm- en betekenisgerichte instructie (voor meer semantische kennis het liefste vormgericht), het leren in woordparen die qua klank op elkaar lijken, rekening houdend met woordkenmerken (beeldbaarheid, concreetheid, vorm, inhouds- en functiewoorden) en collectief lezen met behulp van een vrij te volgen script. Woorden lijken het beste in de vorm van video's aangeboden te kunnen worden, anders als tekst op papier en minder graag als digitale tekst. Video-aspecten die hierbij effectief lijken, zijn visuele en geluidseffecten (met name wanneer visuele en auditieve informatie tegelijk worden aangeboden (*dual-coding*)), een karakter in de video dat een soort gesprek met de kijker voert, expliciete definities van woorden benoemen en het herhalen van woorden (dit lijkt voornamelijk voor kinderen die thuis weinig L2-blootstelling ervaren te werken). Ook *audio prompting*, van tevoren opgenomen audio laten horen bij het bijbehorende woord, bijvoorbeeld door middel van een sociale robot, lijkt te kunnen bijdragen aan woordenschatontwikkeling.

De sociale robot die gebruik maakte van tweetalige *audio prompting* was niet effectiever dan de ééntalige robot in de studie van Leeuwestein et al. (2021), hoewel uit onderzoek van Rivera Pérez et al. (2019) wel bleek dat meertalige instructie effectiever wel zou zijn. Óf en hoe L1 gebruikt kan worden bij woordenschatstimulatie van L2 kan dus op basis van deze review niet worden vastgesteld. Een tweede tegenstrijdigheid binnen de review is de invloed van kwaliteit en kwantiteit van blootstelling aan de meerderheidstaal

thuis; in het onderzoek van Vaahtoranta et al. (2021) bleek dit geen significant effect te hebben op woordenschat, terwijl in de onderzoeken van Sun et al. (2021), Wong en Neuman (2021) en van Wong en Samudra (2019) thuisfactoren wel significante voorspellers waren. Deze significant bevonden thuisfactoren waren het taal- en opleidingsniveau van de moeder, leesinput thuis in de meerderheidstaal en de duur van de blootstelling aan L2 in jaren. Een mogelijke verklaring is dat de kinderen uit de studie van Vaahtoranta et al. (2021) gemiddeld ouder zijn dan die in de studies die tegenspreken (5;4 in vergelijking tot 4;9 en 4;7), waardoor de kinderen mogelijk meer tijd op de voorschool hebben doorgebracht. Hier zijn ze de hele dag aan de meerderheidstaal blootgesteld, waardoor de totale blootstelling aan de meerderheidstaal waarschijnlijk hoger ligt en het daardoor mogelijk is dat de invloed van de blootstelling aan de meerderheidstaal thuis is afgenomen.

Een overige factor die niet significant bleek te zijn, was variabiliteit in het taalniveau van klasgenoten.

Kortom, het beste is om instructie aan te passen op kind- en woordkenmerken, zoals het niveau van het kind in L1 en L2, de beeldbaarheid en concreetheid van woorden en of het woord een inhouds- of functiewoord is (zie §4.7 voor implicaties in de praktijk). Daarnaast kunnen het bieden van veel emotionele support, het vaak herhalen van woorden en het gebruik maken van d*ual-coding* een positieve bijdrage leveren aan woordenschatontwikkeling.

De deelvraag - welke relatief meer invloed uitoefenen, interne of externe factoren - kan ook worden beantwoord. Het grootste deel van de behandelde mogelijkheden voor woordenschatstimulatie bestaat uit externe factoren, wat echter niet per definitie betekent dat deze meer invloed uitoefenen dan interne factoren. In de studies die interne en externe factoren vergeleken, hadden interne en externe factoren namelijk ongeveer evenveel invloed op woordenschatontwikkeling. Hoewel er nog meer bewijs aangeleverd zou moeten worden

om de hypothese volledig te bewijzen, is dit resultaat is in lijn met de c*ritical mass hypothesis*, die stelt dat interne factoren zich pas kunnen ontwikkelen en pas hun invloed uit kunnen oefenen wanneer er een kantelpunt door genoeg input, 'massa', van externe factoren is bereikt (Marchman & Bates, 1994). Een kind heeft dan dus eerst genoeg stimulatie en instructie nodig om interne mechanismen die bijdragen aan woordenschatontwikkeling, zoals bijvoorbeeld het fonologisch kortetermijngeheugen, goed te kunnen laten werken.

4.3 Vergelijkbaar Onderzoek

In een enigszins vergelijkbare literatuurreview van Banse (2019) is gekeken hoe de ontwikkeling van tweetalige kinderen op vier domeinen, waaronder taal en geletterdheid, gestimuleerd kan worden. Een overeenkomst tussen beide onderzoeken is dat beide een studie bevatten waarin meertalige instructie effectiever was dan ééntalige L2-instructie. Een verschil hierbij is dat de huidige review een studie bevat die tegenstrijdige resultaten laat zien. Mogelijk is dit te verklaren doordat de studie in kwestie, het onderzoek van Leeuwestein et al. (2021), een vrij specifieke manier van één- en meertalige woordenschatinstructie onderzoekt, namelijk instructie door middel van sociale robots. Dat meertalige instructie bij dit onderzoek niet effectiever was, hoeft niet te betekenen dat dit resultaat te generaliseren is naar alle één- en meertalige instructie. Andere manieren van meertalige instructie kunnen wel effectiever blijken dan deze specifieke manier.

Een tweede verschil tussen de twee reviews is dat Banse (2019) in zijn review een studie behandelt die een negatieve correlatie tussen woordenschat van L1 en L2 heeft gevonden, wat de positieve correlatie uit de studie van Baker et al. (2021) in de huidige review tegenspreekt. Een mogelijke verklaring voor dit verschil is het verschil in leeftijd. De participanten in de studie van Baker et al. gingen namelijk al naar *kindergarten* (de Amerikaanse vroegschoolse educatie) en in de behandelde studie in het artikel van Banse nog

niet. Dit verschil in leeftijd kan zorgen voor een verschil in woordenschat, wat het verband tussen de woordenschat in L1 en L2 mogelijk anders maakt.

In het algemeen kunnen de verschillen tussen de reviews verklaard worden doordat er geen overlap in de studies zit, omdat het huidige onderzoek naar studies vanaf 2019 kijkt, wat het publicatiejaar van het artikel van Banse (2019) is. Andere verschillen zijn het gebruik van andere zoekmachines, geen focus op Germaanse talen en een bredere blik; Banse focust met de vier domeinen niet alleen op woordenschat en heeft een groter aantal studies meegenomen (vanaf 2000; Banse, 2019).

4.4 Limitaties Geïncludeerde Studies

Aan de geïncludeerde studies in het huidige onderzoek zitten haken en ogen. Ten eerste was de kwaliteit beduidend laag; ongeveer de helft van de studies scoorde zeer laag en de gemiddelde score was rond de 50%. Dit betekent dat de meeste resultaten voorzichtig geïnterpreteerd dienen te worden. Wat betreft de kwaliteit onder andere opviel, is dat de power en effectgroottes in weinig onderzoeken zijn benoemd. Dit sluit aan op een eerdere review van Fritz et al. (2012), waarin minder dan de helft van de studies de effectgrootte was benoemd. In de geïncludeerde studies die een kwaliteitspunt kregen voor de power, werd deze enkel genoemd en niet berekend, zoals dat er benoemd is dat een *within-subjects* design is gebruikt om de power te verhogen.

Een ander kwaliteitsaspect is de mate waarin onderzoeken longitudinaal te noemen zijn. Eén derde van de onderzoeken had een voor- en nameting op dezelfde dag. Hiermee wordt onderzoek naar de effectiviteit van interventies tekortgedaan en kan geen lange termijneffect worden gemeten.

Een andere opvallende factor is dat het grootste deel van de geïncludeerde onderzoeken is afgenomen in de VS met Spaans als thuistaal. Dit kan gevolgen hebben voor de mate waarin de uitkomsten van het huidige onderzoek te generaliseren zijn naar populaties

in andere landen, omdat de omstandigheden daar anders zijn (taal, cultuur, schoolsysteem, etc.).

Een laatste limitatie van de geïncludeerde studies is dat niet alle benoemde factoren even makkelijk te beïnvloeden zijn door professionals. Hoewel een inclusiecriterium was dat een deel van de factoren in een studie te beïnvloeden moest zijn, is er geen rekening gehouden met het gemak om deze factor te kunnen beïnvloeden. Met name interne factoren, zoals fonologisch kortetermijngeheugen en werkgeheugen zijn vaak moeilijker te beïnvloeden en vereisen eerst de juiste scholing van professionals en mogelijk meer onderzoek. Echter, niet alle factoren hoeven beïnvloedbaar te zijn om nuttig te zijn voor deze review; leeftijd is bijvoorbeeld een factor die niet te beïnvloeden is, maar waar wel rekening mee gehouden kan worden in instructie.

4.5 Sterke en Zwakke Aspecten Huidig Onderzoek

Ook aan de huidige review zitten sterke en zwakke kanten. Hoewel volgens Grant en Booth (2009) het liefst minimaal 3 onderzoekers en een jaar nodig zijn om een kwalitatief goede systematische literatuurreview uit te voeren en alle literatuur rondom een bepaald onderwerp te behandelen, was dat niet het doel van en haalbaar voor het huidige onderzoek. Dit onderzoek is geschreven in vier maanden door één onderzoeker met de hulp van twee tweede beoordelaars. Het bevat een vrij uitgebreide kwaliteitsbeoordeling en door de gebruikte figuren en tabellen is het onderzoek overzichtelijk en goed te volgen. De kracht van dit onderzoek zit in de systematiek, waardoor elke stap te herleiden en repliceren is.

In geval van replicatie zijn er wel een aantal verbeterpunten waar rekening mee gehouden kan worden. Zo moeten er andere studies geïncludeerd worden, zoals meer studies die ingaan op hoe L1 gebruikt kan worden in L2 instructie. Om andere studies te vinden, moet de zoekslag aangepast worden. Er is namelijk een grote kans dat er studies zijn die gebruik maken van andere terminologie dan de huidige zoekslag, maar wel antwoord kunnen

geven op de onderzoeksvraag. Suggesties voor Engelse zoektermen die toegevoegd kunnen worden, zijn *'language minority children', 'nonnative speak*', migrant, refugee, boost*, 'majority language', lexicon, semantic knowledge* en *'word learning'*. Bij replicatie van dit onderzoek zou men naast de meerderheidstaal ook rekening kunnen houden met de thuistaal. In dit onderzoek was dit veelal Spaans, maar Turks, Marokkaans, Surinaams, Syrisch of Oekraïens zou voor dit onderzoek relevanter zijn (zie §1.4).

Verder zijn er ook suggesties voor verbetering van het instrument voor kwaliteitsassessment. Effectgroottes hadden inhoudelijk kunnen worden beoordeeld (dus niet alleen óf de EG is genoemd, maar ook de grootte ervan beoordelen) en ook met andere maten dan Cohen's *d*. Hoewel dit de meest gebruikte maat is, wordt deze niet altijd op de juiste manier gebruikt en zijn er betere alternatieven (Dankel et al., 2017). Als de power vaker benoemd zou zijn, zou ook hier inhoudelijk over geoordeeld kunnen worden. Een alternatief is om de steekproefgrootte te beoordelen, aangezien er een positief verband is tussen steekproefgrootte en power (Agresti, 2018).

Verder kunnen kwaliteitsitems worden toegevoegd die reflecteren op de hoeveelheid gebruikte instrumenten (meer instrumenten kunnen samen een concept beter in kaart brengen) en missing data. Ook had gekeken kunnen worden of er op de oorspronkelijke groeicurve is gereflecteerd in de studies, aangezien kinderen ook zonder interventie vooruitgaan op woordenschat en je het effect boven op deze ontwikkeling wilt meten. Als laatste had onderscheid gemaakt kunnen worden in verschillende soorten betrouwbaarheid. Guttman's lambda-2 zit bijvoorbeeld standaard dichter bij de daadwerkelijke betrouwbaarheid dan Crohnbach's alfa (Agresti, 2018).

Een overige limitatie van de huidige review is dat niet alle gemeenten weinig blootstelling aan Nederlands thuis als grond voor een VVE-indicatie zien, de gemeente Amsterdam bijvoorbeeld niet (Kaspers, 2022). Bovendien was de interbeoordelaars-

betrouwbaarheid maar matig. Dit is te verklaren door de grote hoeveelheid studies die in korte tijd gescreend en geanalyseerd moesten worden, waardoor dit niet altijd even precies is gebeurd.

4.6 Aanbevelingen

Naast aanbevelingen voor verbetering van een soortgelijk onderzoek als de huidige review, zijn er meer aanbevelingen voor toekomstig onderzoek. Ten eerste is het van belang dat empirisch onderzoek vaker de power en effectgrootte specifiek benoemt en berekent. Deze maten kunnen namelijk aangeven in hoeverre een uitkomst daadwerkelijk toepasbaar is op de praktijk, zoals beschreven in §2.4. Volgens Fox en Mathers (1997) is het benoemen van de power essentieel voor elk onderzoek waarbij een hypothese wordt getoetst. Daarom was de bevinding dat dit maar heel weinig wordt gedaan ook zeer opvallend en moet dit in toekomstig onderzoek worden verbeterd. Daarnaast moet er naar meer contexten onderzoek gedaan worden, zoals verschillende landen, thuistalen, meerderheidstalen en schoolsystemen. Dit is belangrijk om een breder en beter te generaliseren beeld te kunnen vormen van woordenschatontwikkeling bij meertalige kinderen. Op basis van de geïncludeerde onderzoeken lijkt er ook meer onderzoek gedaan te moeten worden naar óf en hoe L1 ingezet moet worden bij L2-instructie. Een *scoping* review wordt aangeraden om eerst in kaart te brengen hoeveel onderzoek er ook van voor 2019 naar dit onderwerp is gedaan (Grant & Booth, 2009). Op basis daarvan kan worden vastgesteld of er meer empirisch onderzoek naar dit onderwerp gedaan moet worden. Ook lijken er veel methoden te zijn die enkel zijn onderzocht bij kinderen die maar één taal of vrijwillig een vreemde taal erbij leren (bijvoorbeeld het gebruik van muziek, geuren en koken). Deze methoden zouden onderzocht kunnen worden bij kinderen met een andere thuistaal dan de meerderheidstaal, om te kijken of het ook toepasbaar is op de populatie van het huidige onderzoek. De laatste aanbeveling is het reviewen van literatuur over woordenschatstimulatie van meertalige kinderen uit een

klinische populatie. Bij de screening bleek hier veel onderzoek naar te zijn, bijvoorbeeld naar meertalige kinderen met Autisme Spectrum Stoornis en ADHD. Een overzicht en synthese van de beschikbare kennis hierover kan nuttig zijn voor professionals binnen het onderwijs.

4.7 Implicaties voor de Praktijk

Ondanks de limitaties kunnen de resultaten uit de huidige review wel adviezen opleveren voor professionals en hen handvatten bieden omtrent de vormgeving van woordenschatinstructie. Het overkoepelende advies is dat er zo veel mogelijk gedifferentieerd moet worden op basis van kind- en woordkenmerken. Intentionele leermomenten en een inputrijke omgeving moeten worden nagestreefd, waar door genoeg stimulatie van externe factoren ook interne factoren zich kunnen ontwikkelen. Het liefst worden verschillende stimulatiemethoden gebruikt en meerdere zintuigen tegelijk gestimuleerd (met name gehoor en visie). Dit kan bijvoorbeeld via *audio prompting* door middel van sociale robots, wat leraren kan helpen om een link te leggen tussen L1 en L2 wanneer ze zelf de L1 van het kind niet spreken. Ook kan *dual-coding* worden toegepast met behulp van video's of door afbeeldingen op het bord te laten zien. Met name kinderen die thuis weinig L2-stimulatie ondervinden, zoals de doelpopulatie, hebben hier baat bij. Voor hen werken video's die woorden herhalen het beste.

Kinderen die zelf een relatief laag L2-woordenschatniveau hebben, kunnen het beste beginnen met het leren van beeldbare en concrete inhoudswoorden via plaatjes en definities. Abstractere woorden, zoals functiewoorden, vereisen eerst een basiswoordenschat en kunnen vervolgens het beste in verschillende contexten worden aangeleerd. Kinderen die hierbij wel een grote L1-woordenschat hebben, zullen waarschijnlijk sneller groeien dan kinderen met een lager niveau. Wissel kinderen met verschillende niveaus af, zodat ze elkaar positief kunnen beïnvloeden.

Focus verder bij instructie het liefst op vorm in plaats van betekenis, geef synoniemen en bied het kind ruimte om zelf activiteiten te initiëren. Collectief boeklezen kan nuttig zijn, met name door verdiepende vragen te stellen. Zorg daarnaast dat het klaslokaal niet over-gestructureerd is, er veel emotionele support geboden wordt en dat sociale validiteit van interventies bij leraren hoog is, dus dat leraren ook geloven in de effectiviteit van wat ze doen. Professionals moeten zich daar dus in verdiepen voordat ze methodes toepassen (ook om de kwaliteit van implementatie te bevorderen).

4.8 Conclusie

Deze literatuurreview biedt een overzicht van recente studies naar interne en externe factoren die woordenschatontwikkeling beïnvloeden bij jonge kinderen die met een andere thuistaal dan de meerderheidstaal opgroeien en hierdoor een VVE-indicatie hebben gekregen. Wanneer ondanks de lage kwaliteit van de studies een advies aan professionals geformuleerd zou moeten worden, zou dit zijn dat zo veel mogelijk gedifferentieerd moet worden op basis van kind- en woordkenmerken, zoals het niveau van woordenschat in L1 en L2, de beeldbaarheid en concreetheid van woorden en of het woord een inhouds- of functiewoord is. Daarnaast is het bieden van veel emotionele support, het vaak herhalen van woorden en het gebruik maken van *dual-coding* aan te raden. Interne en externe factoren lijken ongeveer evenveel invloed te hebben op woordenschatontwikkeling.

Referenties

Aalberse, S. P., Andringa, S. J., Hanekamp, K., Keydeniers, D. J., Kuiken, F., Thieme, A. M. M., Verhagen, J., & Vos, E. M. J. (2021). *Meertaligheid in dagopvang: Een onderzoek naar de effecten van het gebruik van twee talen op het kinderdagverblijf (MIND: multilingualism in daycare)*. https://www.eerstekamer.nl/bijlage/20220202/meertaligheid_in_dagopvang_een/document3/f=/vlq4betzigk5.pdf

Agresti, A. (2018). *Statistical Methods for the Social Sciences* (5e ed.). Pearson.

Araújo, L., & Costa, P. (2015). Home Book Reading and Reading Achievement in EU Countries: The Progress in International Reading Literacy Study 2011 (PIRLS). *Educational Research and Evaluation*, *21*(5–6), 422–438. https://doi.org/10.1080/13803611.2015.1111803

Banse, H. W. (2019). Dual language learners and four areas of early childhood learning and development: What do we know and what do we need to learn? *Early Child Development and Care*, *191*(9), 1347–1360. https://doi.org/10.1080/03004430.2019.1658086

Baker, D., McCoach, B. D., Ware, S., Coyne, M. D., & Rattan, S. M. (2021). Effects of Spanish vocabulary knowledge on the English word knowledge and listening comprehension of bilingual students. *International Journal of Bilingual Education and Bilingualism*, *0* 1–15. https://doi.org/10.1080/13670050.2021.1908219

Beckers, D., Karssen, L. T., Vink, J. M., Burk, W. J., & Larsen, J. K. (2021). Food parenting practices and children's weight outcomes: A systematic review of prospective studies. *Appetite, 158*, 105010. https://doi.org/10.1016/j.appet.2020.105010

Centraal Bureau voor de Statistiek. (2021, 10 september). *Bevolking; migratieachtergrond, generatie, leeftijd, regio, 1 januari.* Geraadpleegd op 29 maart 2022, van https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84910NED/table?ts=1648408547027

Centraal Bureau voor de Statistiek. (2022a, 31 januari). *Asielinstroom 2021 vooral tweede helft van het jaar toegenomen*. Geraadpleegd op 29 maart 2022, van https://www.cbs.nl/nl-nl/nieuws/2022/05/asielinstroom-2021-vooral-tweede-helft-van-het-jaar-toegenomen

Centraal Bureau voor de Statistiek. (2022b, 28 februari). *Hoeveel immigranten komen naar Nederland?* Geraadpleegd op 29 maart 2022, van https://www.cbs.nl/nl-nl/dossier/dossier-asiel-migratie-en-integratie/hoeveel-immigranten-komen-naar-nederland-

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental Language Disorders—A Follow-Up in Later Adult Life. Cognitive, Language and Psychosocial Outcomes. *Journal of Child Psychology and Psychiatry*, *46*(2), 128–149. https://doi.org/10.1111/j.1469-7610.2004.00342.x

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213–220. https://doi.org/10.1037/h0026256

Corporation for Digital Scholarship. (2006). *Zotero* (6.0.8). Roy Rosenzweig Center for History and New Media at George Mason University. https://www.zotero.org/

Coyne, M. D., McCoach, D. B., Ware, S. M., Loftus-Rattan, S. M., Baker, D. L., Santoro, L. E., & Oldham, A. C. (2022). Supporting vocabulary development within a multitiered system of support: Evaluating the efficacy of supplementary kindergarten vocabulary intervention. *Journal of Educational Psychology, 0*. https://doi.org/10.1037/edu0000724

Dankel, S. J., Mouser, J. G., Mattocks, K. T., Counts, B. R., Jessee, M. B., Buckner, S. L., Loprinzi, P. D., & Loenneke, J. P. (2017). The widespread misuse of effect sizes. *Journal of Science and Medicine in Sport*, *20*(5), 446–450. https://doi.org/10.1016/j.jsams.2016.10.003

Deeks, J. J., Dinnes, J., D'Amico, R., Sowden, A. J., Sakarovitch, C., Song, F., Petticrew, M., Altman, D. G., International Stroke Trial Collaborative Group, & European Carotid Surgery Trial Collaborative Group (2003). Evaluating non-randomised intervention studies. *Health technology assessment (Winchester, England), 7*(27), 1–173. https://doi.org/10.3310/hta7270

Evers, A. V. A. M., Sijtsma, K., Lucassen, W., & Meijer, R. R. (2010). Het COTAN beoordelingssysteem voor de kwaliteit van tests herzien: What's new? *De Psycholoog*, *45*(9), 48-55. https://research.rug.nl/en/publications/het-cotan-beoordelingssysteem-voor-de-kwaliteit-van-tests-herzien

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect Size Estimates: Current Use, Calculations, and Interpretation. *Journal of Experimental Psychology-General*, *141*(1), 2–18. https://doi.org/10.1037/a0024338

Fox, N., Mathers, N. (1997). Empowering research: statistical power in general practice research. *Family Practice*, *14(*4), 324–329. https://doi.org/10.1093/fampra/14.4.324

Garcia, E. B. (2021). Peer effects on dual language learners' English and Spanish receptive vocabulary development. *Early Childhood Research Quarterly*, *56*, 180–189. https://doi.org/10.1016/j.ecresq.2021.03.015

Gemeente Den Haag. (2021, 1 mei*). Subsidieregeling voorschoolse educatie Den Haag 2021*. Overheid.nl. Geraadpleegd op 1 april 2022, van https://lokaleregelgeving.overheid.nl/CVDR658194?&show-wti=true

Gibson, C. (2016). Bridging English Language Learner Achievement Gaps through Effective Vocabulary Development Strategies. *English Language Teaching*, *9*(9), 134–138. https://eric.ed.gov/?id=EJ1110015

Gladfelter, A., Barron, K. L., & Johnson, E. (2019). Visual and verbal semantic productions in children with ASD, DLD, and typical language. *Journal of Communication Disorders*, *82*. https://doi.org/10.1016/j.jcomdis.2019.105921

Grant, M. J. & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*(2), 91-108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Grøver, V., Rydland, V., Gustafsson, J.-E., & Snow, C. E. (2020). Shared Book Reading in Preschool Supports Bilingual Children's Second-Language Learning: A Cluster-Randomized Trial. *Child Development*, *91*(6), 2192–2210. https://doi.org/10.1111/cdev.13348

Hadley, E. B., Dedrick, R. F., Dickinson, D. K., Kim, E., Hirsh-Pasek, K., & Golinkoff, R. M. (2021). Exploring the relations between child and word characteristics and preschoolers' word-learning. *Journal of Applied Developmental Psychology*, *77*. https://doi.org/10.1016/j.appdev.2021.101332

Hawkins, J. A. (2018). Germanic languages. In *The world's major languages* (pp. 51-59). https://www.taylorfrancis.com/chapters/edit/10.4324/9781315644936-3/germanic-languages-john-hawkins

Herrera, S., Phillips, B. M., Newton, Y. C., Dombek, J. L., & Hernandez, J. A. (2021). *Effectiveness of Early Literacy Instruction: Summary of 20 Years of Research* https://eric.ed.gov/?id=ED613947

Inspectie van het Onderwijs. (2020). *Kwaliteit van de Voor- en Vroegschoolse Educatie in Nederland (Rapport meting VVE 2019)*.

https://www.onderwijsinspectie.nl/onderwijssectoren/voor-en-vroegschoolse-educatie/documenten/rapporten/2020/05/19/meting-vve-2019

Inspectie van het Onderwijs. (z.d.). *Wat is vve?* Voor- en Vroegschoolse educatie | Inspectie van het Onderwijs. Geraadpleegd op 26 februari 2022, van https://www.onderwijsinspectie.nl/onderwijssectoren/voor-en-vroegschoolse-educatie/wat-is-voor-en-vroegschoolse-educatie

Janssen, C., Segers, E., McQueen, J. M., & Verhoeven, L. (2019). Comparing Effects of Instruction on Word Meaning and Word Form on Early Literacy Abilities in Kindergarten. *Early Education and Development*, *30*(3), 375–399. https://doi.org/10.1080/10409289.2018.1547563

Kaspers, F. (2022). *Hoe werkt de VVE-indicatie en welke criteria zijn er?* Onderwijsconsument. Geraadpleegd op 23-05-2022 van https://www.onderwijsconsument.nl/hoe-werkt-de-vve-indicatie/

Leeuwestein, H., Barking, M., Sodaci, H., Oudgenoeg-Paz, O., Verhagen, J., Vogt, P., Aarts, R., Spit, S., de Haas, M., de Wit, J., & Leseman, P. (2021). Teaching Turkish-Dutch Kindergartners Dutch Vocabulary with a Social Robot: Does the Robot's Use of Turkish Translations Benefit Children's Dutch Vocabulary Learning? *Journal of Computer Assisted Learning*, *37*(3), 603–620. https://doi.org/10.1111/jcal.12510

Leseman, P. & Veen, A. (2016). *Ontwikkeling van kinderen en relatie met kwaliteit van voorschoolse instellingen. Resultaten uit het pre-COOL cohortonderzoek* (Nr. 947). https://kohnstamminstituut.nl/rapport/ontwikkeling-van-kinderen-en-relatie-met-kwaliteit-van-voorschoolse-instellingen-resultaten-uit-het-pre-cool-cohortonderzoek/

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That

Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Medicine*, *6*(7), e1000100. https://doi.org/10.1371/journal.pmed.1000100

Mancilla-Martinez, J., Hwang, J. K., Oh, M. H., & Pokowitz, E. L. (2020). Patterns of Development in Spanish-English Conceptually Scored Vocabulary among Elementary-Age Dual Language Learners. *Journal of Speech, Language, and Hearing Research*, *63*(9), 3084–3099. https://doi.org/10.1044/2020_JSLHR-20-00056

Manning, B. L., Roberts, M. Y., Estabrook, R., Petitclerc, A., Burns, J. L., Briggs-Gowan, M., Wakschlag, L. S., & Norton, E. S. (2019). Relations between toddler expressive language and temper tantrums in a community sample. *Journal of Applied Developmental Psychology*, *65*, 101070. https://doi.org/10.1016/j.appdev.2019.101070

Marchman, V.A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*, 339–366. https://doi.org/10.1017/S0305000900009302

Miles, K. P., McFadden, K. E., & Ehri, L. C. (2019). Associations between Language and Literacy Skills and Sight Word Learning for Native and Nonnative English-Speaking Kindergarteners. *Reading and Writing: An Interdisciplinary Journal*, *32*(7), 1681–1704. https://doi.org/10.1007/s11145-018-9919-5

Ministerie van Algemene Zaken. (2021a, 18 november). *Voorschoolse of vroegschoolse educatie krijgen voor peuter of kleuter*. Wetten en regelingen | Rijksoverheid.nl. Geraadpleegd op 8 maart 2022, van https://www.rijksoverheid.nl/wetten-en-regelingen/productbeschrijvingen/voorschoolse-of-vroegschoolse-educatie-krijgen-voor-peuter-of-kleuter

Ministerie van Algemene Zaken. (2021b, 24 november). *Financiering aanpak onderwijsachterstanden*. Voorschoolse en vroegschoolse educatie | Rijksoverheid.nl.

Geraadpleegd op 8 maart 2022, van https://www.rijksoverheid.nl/onderwerpen/voorschoolse-en-vroegschoolse-educatie/financiering-onderwijsachterstanden

Ministerie van Justitie en Veiligheid. (2022, 23 maart). *Kamerbrief over aanpak opvang vluchtelingen Oekraïne*. Kamerstuk | Rijksoverheid.nl. https://www.rijksoverheid.nl/documenten/kamerstukken/2022/03/17/tk-brief-oekraine

Nederlands Jeugdinstituut (NJI). (2022). *Voor- en vroegschoolse educatie (vve)*. NJI. Geraadpleegd op 8 maart 2022, van https://www.nji.nl/voor-en-vroegschoolse-educatie-vve

Oortwijn, M. B., Boekaerts, M., & Vedder, P. (2005). Is coöperatief rekenen op multiculturele basisscholen effectiever dan klassikaal rekenen? *Panama-Post*, *24*, 3-11. https://www.volgens-bartjens.nl/documenten/archief/bartjens/pp_24_4_16_-_24_oortwijn_is_cooperatief_rekenen_op_multiculturele_basisscholen_effectiever_dan_klassikaal_rekenen.pdf

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1). https://doi.org/10.1186/s13643-016-0384-4

Pan, Y. C., Tsai, T. H., Huang, Y. K., & Liu, D. (2018). Effects of expanded vocabulary support on L2 listening comprehension. *Language Teaching Research*, *22*(2), 189–207. https://doi.org/10.1177/1362168816668895

Pollard-Durodola, S. D., Gonzalez, J. E., Saenz, L., Soares, D., Davis, H. S., Resendez, N., & Zhu, L. (2021). The Social Validity of Content Enriched Shared Book Reading Vocabulary Instruction and Preschool DLLs' Language Outcomes. *Early Education and Development*, *0*(0), 1–23. https://doi.org/10.1080/10409289.2021.1946761

Preston, J. L., Frost, S. J., Mencl, W. E., Fulbright, R. K., Landi, N., Grigorenko, E., Jacobsen, L., & Pugh, K. R. (2010). Early and late talkers: School-age language, literacy and neurolinguistic differences. *Brain*, *133*(8), 2185–2195. https://doi.org/10.1093/brain/awq163

Reilly, S. E., Johnson, A. D., Luk, G., & Partika, A. (2020). Head Start Classroom Features and Language and Literacy Growth among Children with Diverse Language Backgrounds. *Early Education and Development*, *31*(3), 354–375. https://doi.org/10.1080/10409289.2019.1661935

Rivera Pérez, J. F., Creaghead, N. A., Washington, K., Guo, Y., Raisor-Becker, L., & Combs, S. (2019). Using Audio Prompting to Assist Monolingual Speech-Language Pathologists to Teach English-Spanish Vocabulary to English Learners. *Communication Disorders Quarterly*, *41*(1), 3–11. https://doi.org/10.1177/1525740118819659

Rolletschek, H. (2020). The Effects of Odor on Vocabulary Learning. *Language Teaching Research Quarterly*, *18*, 21–39. https://eric.ed.gov/?id=EJ1270078

Scheepers, M., Geertsema, S., le Roux, M., & Graham, M. (2021). Phonological Awareness and Learning to Read in Afrikaans: The Role of Working Memory. *South African Journal of Childhood Education*, *11*(1). http://eric.ed.gov/?id=EJ1316859

Schmeets, H. L. C. (2021, 16 juli). *Talen en dialecten in Nederland*. Centraal Bureau voor de Statistiek. https://www.cbs.nl/nl-nl/longread/statistische-trends/2021/talen-en-dialecten-in-nederland?onepage=true#c-3--Conclusie-en-discussie

Silverman, R., Artzi, L., McNeish, D., Hartranft, A., Martin-Beltran, M., & Peercy, M. (2019). The relationship between media type and vocabulary learning in a cross age peer-learning program for linguistically diverse elementary school students.

Contemporary Educational Psychology, *56*, 106–116. https://doi.org/10.1016/j.cedpsych.2018.12.004

Slot, P. L., Jepma, I., Muller, P., Romijn, B. R., & Leseman, P. P. M. (2017). *Kwaliteit van de Nederlandse kinderdagopvang, peuteropvang, buitenschoolse opvang en gastouderopvang: Landelijke Kwaliteitsmonitor Kinderopvang Meting 2017*. https://dspace.library.uu.nl/bitstream/handle/1874/360591/kinderopvang.pdf?sequence=1

Sun, H., Bornstein, M. H., & Esposito, G. (2021). The Specificity Principle in Young Dual Language Learners' English Development. *Child Development*, *92*(5), 1752–1768. https://doi.org/10.1111/cdev.13558

Vaahtoranta, E., Suggate, S., Lenhart, J., & Lenhard, W. (2021). Language exposure and phonological short-term memory as predictors of majority language vocabulary and phonological awareness in dual language learning. *Bilingualism: Language and Cognition*, *24*(2), 319–332. https://doi.org/10.1017/S1366728920000541

Van der Ploeg, J. (2022, 11 maart). Oekraïense vluchtelingen zijn welkom, maar hoelang en hoeveel? *de Volkskrant*. Geraadpleegd op 29 maart 2022, van https://www.volkskrant.nl/nieuws-achtergrond/oekraiense-vluchtelingen-zijn-welkom-maar-hoelang-en-hoeveel~b5429baf/?referrer=https%3A%2F%2Fwww.ecosia.org%2F

Veen, A., & Leseman, P. (2015). *Pre-COOL cohortonderzoek: Resultaten van de voorschoolse periode* (Nr. 932). Kohnstamm Instituut. https://pre-cool.kohnstamminstituut.nl/resultaten-en-publicaties.html

Wasik, B. A. (2010). What Teachers Can Do to Promote Preschoolers' Vocabulary Development: Strategies From an Effective Language and Literacy Professional

Development Coaching Model. *The Reading Teacher, 63*(8), 621–633. https://doi.org/10.1598/RT.63.8.1

Wells, G. A., Shea, B., & O'Connell, D. (2011). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses.* www.ohri.ca/programs/clinical_epidemiology/oxford.asp

Wiersma, A. (2021, 16 december). Verzamelbrief Kinderopvang december 2021 [Kamerbrief]. Geraadpleegd op 27 februari, van https://www.tweedekamer.nl/kamerstukken/brieven_regering/detail?id=2021Z23735&did=2021D50340

Wong, K. M., Flynn, R. M., & Neuman, S. B. (2021). L2 vocabulary learning from educational media: The influence of screen-based scaffolds on the incidental–intentional continuum. *TESOL Journal*, *12*(4), e641. https://doi.org/10.1002/tesj.641

Wong, K. M., & Neuman, S. B. (2021). Learning L2 vocabulary on screen: The role of screen-based pedagogical supports on dual language learners. *Computer Assisted Language Learning*, *0*(0), 1–24. https://doi.org/10.1080/09588221.2021.1999983

Wong, K. M., & Samudra, P. G. (2019). L2 Vocabulary Learning from Educational Media: Extending Dual-Coding Theory to Dual-Language Learners. *Computer Assisted Language Learning*, *34*(8), 1182-1204. https://doi.org/10.1080/09588221.2019.1666150

Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-Month-olds use distinct objects as cues to categorize speech information. *Cognition*, *113*(2), 234–243. https://doi.org/10.1016/j.cognition.2009.08.010

Bijlage A

Zoekslag

ERIC & PsychINFO:

(((Vocabulary OR "Word knowledge") AND (Develop* OR Acquisit* OR Educat* OR Stimul* OR Enrich* OR Guid* OR Instruct* OR Strateg* OR Encourag* OR Activate OR Promote OR Strength* OR Motivat*)) OR "Effective Vocabulary Development Strategies") AND (((Second OR Foreign OR Novel OR Third OR Multi OR Dual) AND (Language OR Linguistic)) OR Bilingual* OR Multilingual* OR Trilingual* OR NT2 OR Polyglot OR Diglot OR English Language Learners OR L1 OR L2) AND (Toddler* OR Early childhood OR Preschool OR 2 year old OR 3 year old OR 4 year old OR 5 year old OR 6 year old OR Young OR Early learning OR Kindergarten OR Daycare)

WebOfScience:

Eerste zoekbalk:

((Vocabulary OR "Word knowledge") AND (Develop* OR Acquisit* OR Educat* OR Stimul* OR Enrich* OR Guid* OR Instruct* OR Strateg* OR Encourag* OR Activate OR Promote OR Strength* OR Motivat*)) OR "Effective Vocabulary Development Strategies"

AND (Tweede zoekbalk):

((Second OR Foreign OR Novel OR Third OR Multi OR Dual) AND (Language OR Linguistic)) OR Bilingual* OR Multilingual* OR Trilingual* OR NT2 OR Polyglot OR Diglot OR English Language Learners OR L1 OR L2

AND (Derde zoekbalk):

Toddler* OR Early childhood OR Preschool OR 2 year old OR 3 year old OR 4 year old OR 5 year old OR 6 year old OR Young OR Early learning OR Kindergarten OR Daycare

Bijlage B

Protocol

	Een studie wordt geëxcludeerd wanneer:	*Label:*
1.	Er niet (gedeeltelijk) gekeken wordt naar één of meerdere interne/externe factoren die woordenschat kunnen stimuleren. Bijvoorbeeld wanneer er alleen in kaart gebracht wordt hoe de woordenschat ontwikkeling verloopt, maar de ontwikkeling zelf niet wordt beïnvloed (bijv. dmv een interventie)	*Geen woordenschat stimulatie*
2.	De gemiddelde leeftijd van de participanten bij de baseline/voormeting niet in de leeftijdscategorie 2-6 jaar valt	*Verkeerde populatie*
3.	Twee-/ meertalige kinderen niet de hoofdfocus of een aparte subgroep zijn	*Niet hoofd-/subgroep*
4.	Er enkel sprake is van een klinische populatie (bijv. alleen autistische kinderen zonder niet-klinische controlegroep)	*Verkeerde populatie*
5.	Niet alle of geen deel van de participanten thuis een andere taal dan de meerderheidstaal spreken (bijv. wanneer de studie alleen op vrijwillige meertaligheid is gericht) of wanneer L1 wordt gestimuleerd i.p.v. L2	*Verkeerde populatie*
6.	De verkeerde taal wordt gestimuleerd. Bijvoorbeeld niet de meerderheidstaal of wanneer deze geen Germaanse taal is ((varianten op) het Nederlands, Vlaams, Afrikaans, Engels, Duits, Fries, Deens, Noors, Zweeds, IJslands en Faeröers)	*Verkeerde taal*
7.	De studie niet in het Nederlands, Vlaams of Engels is geschreven	*Verkeerde taal geschreven*
8.	De publicatiedatum eerder is dan 2019 of later dan april 2022	*Publicatiedatum te oud*
9.	De studie niet is getoetst door vakgenoten	*Niet peerreviewed*
10.	De volledige tekst niet beschikbaar is (of wanneer ervoor betaald moet worden)	*Niet fulltext*

(zie volgende pagina voor vervolg)

(Vervolg van vorige pagina)

	Een studie wordt geëxcludeerd wanneer:	*Label:*
11.	De studie geen empirische studie is (bijv. meta-analyses of reviews)	*Niet juiste design*
12.	De studie minder dan 15 participanten heeft	*Niet juiste design*
13.	Geen deel van de onderzochte factoren te beïnvloeden of zelf toe te voegen is	*Niet beïnvloedbaar*
14.	Woordenschat niet specifiek gemeten wordt (bijvoorbeeld indirect via een andere taalvaardigheid)	*Niet specifiek woordenschat*
15.	Wanneer er geen sprake is van een voor- en nameting	*Niet juiste design*
16.	Wanneer de stimulatie niet op een (voor)schoolse instantie (inclusief kinderdagverblijf, etc.) plaatsvindt, maar bijvoorbeeld thuis	*Verkeerde context*

Bijlage C

Kwaliteitsbeoordelingsformulier

Item	*Kwaliteitsaspect*	*Kwaliteitsscore (*)*
1.	Waren de onderzoeksvraag en het onderzoeksdoel duidelijk?	a) Ja (*) b) Nee
2.	Kwaliteit van de onderzoeksmethode	a) De onderzoeksmethode is van goede kwaliteit (bijvoorbeeld door aspecten als dubbel-blind en een randomized control trial) (**) b) De onderzoeksmethode kan beter, maar is niet slecht (*) c) De onderzoeksmethode is slecht d) De onderzoeksmethode past niet goed bij de onderzoeksvraag
3.	Werving van de deelnemers	a) De deelnemers zijn volledig random geselecteerd (**) b) De deelnemers zijn deels random geselecteerd (*) c) Er is sprake van een gelegenheidssteekproef of er wordt onvoldoende informatie gegeven
4.	Vergelijkbaarheid van de twee groepen, derde variabelen	a) Er wordt gecontroleerd op drie variabelen of meer die mogelijk verband houden met de voorspeller en uitkomstmaat (bijv. SES, sekse of leeftijd) (**) b) Er wordt op één of twee variabelen gecontroleerd (*) c) Er wordt niet op derde variabelen gecontroleerd

(Zie volgende pagina voor vervolg)

(Vervolg van vorige pagina)

Item	*Kwaliteitsaspect*	*Kwaliteitsscore (*)*
5.	Power	a) Power wordt genoemd (*) b) Power wordt niet genoemd
6.	Effectgrootte	a) Effectgrootte wordt expliciet benoemd in de vorm van Cohen's *d* (*) b) Effectgrootte wordt niet expliciet benoemd in de vorm van Cohen's *d*
7.	Representeerbaarheid steekproef t.o.v. populatie	a) De steekproef is een goede afspiegeling van de populatie (**) b) Er is twijfel over of de steekproef een goede afspiegeling is van de populatie. Bijvoorbeeld door de manier van werving (*) c) De steekproef is geen goede afspiegeling van de populatie
8.	Meting onafhankelijke variabele	a) De meting is met een zeer betrouwbaar instrument gedaan (≥ .90) (***) b) De meting is met een vrij betrouwbaar instrument gedaan (≥ .80) (**) c) De meting is met een betrouwbaar instrument gedaan (≥ .70) (*) d) De meting is met een twijfelachtig instrument gedaan (≥.60) e) De meting is niet betrouwbaar (betrouwbaarheid <.60) f) De betrouwbaarheid is niet gegeven
9.	Meting afhankelijke variabele (woordenschat)	a) De meting is met een zeer betrouwbaar instrument gedaan (≥ .90) (***) b) De meting is met een vrij betrouwbaar instrument gedaan (≥ .80)(**)

(Zie volgende pagina voor vervolg)

(Vervolg van vorige pagina)

Item	*Kwaliteitsaspect*	*Kwaliteitsscore (*)*
9.	Meting afhankelijke variabele (woordenschat)	c) De meting is met een betrouwbaar instrument gedaan (≥ .70) (*) d) De meting is met een twijfelachtig instrument gedaan (≥.60) e) De meting is niet betrouwbaar (betrouwbaarheid <.60) f) De betrouwbaarheid is niet gegeven
10.	Periode	a) De periode tussen de voor- en nameting is gelijk aan of meer dan een jaar (**) g) De periode tussen de voor- en nameting is tussen de 6 en 12 maanden (*) h) De periode tussen de voor- en nameting is korter dan 6 maanden
11.	Overige limitaties	a) Er is vrijwel geen sprake van overige limitaties (**) b) Er is wel sprake van overige limitaties, maar die verminderen de kwaliteit van de studie niet aanzienlijk (*) c) Er is sprake van één of meerdere overige limitatie(s) die de kwaliteit aanzienlijk verminderen

Bijlage D

Kwaliteitsbeoordeling

(Zie ook Bijlage C voor de bijbehorende vragen)

Tabel D1

Vraag 1 t/m 6 van het Kwaliteitsassessment

Studie	*1. Onderzoeksvraag & - doel*	*2. kwaliteit methode*	*3. werving deelnemers*	*4. Vergelijkbare groepen*	*5. power*	*6. effectwaarde*	*Evt. opmerking over oordeel*
(Baker et al., 2021)	✓ *	A **	B *	B *	X	X	
(Garcia, 2021)	✓ *	B *	B *	B *	X	X	2: aanbeveling: experimenteel design/interventie

(Zie volgende pagina voor vervolg)

(Vervolg van vorige pagina)

Studie	*1. Onderzoeksvraag & - doel*	*2. kwaliteit methode*	*3. werving deelnemers*	*4. Vergelijkbare groepen*	*5. power*	*6. effectwaarde*	*Evt. opmerking over oordeel*
(Grøver et al.,	✓	B	A	A	✓	✓	
2020)	*	*	**	**	*	*	
(Hadley et al.,	✓	B	C	B	✓	X	2: Aanbeveling: experimenteel design/ interventie
2021)	*	*		*	*		
(Janssen et al.,	✓	B	C	A	✓	X	2: Aanbeveling: experimenteel design/ interventie
2019)	*	*		**	*		
(Leeuwestein et	✓	C	C	B	X	X	2: Aanbeveling: interventie langer laten duren
al., 2021)	*			*			
(Miles et al.,	✓	B	C	B	X	X	
2019)	*	*		*			

(Zie volgende pagina voor vervolg)

(Vervolg van vorige pagina)

Studie	*1. Onderzoeksvraag & - doel*	*2. kwaliteit methode*	*3. werving deelnemers*	*4. Vergelijkbare groepen*	*5. power*	*6. effectwaarde*	*Evt. opmerking over oordeel*
(Pollard-Durodola et al., 2021)	✓ *	A **	B *	B *	X	X	
(Reilly et al., 2020)	✓ *	A **	B *	A **	X	✓ *	
(Rivera Pérez et al., 2019)	✓ *	A **	B*	B *	X	X	2: randomized control trial
(Silverman et al., 2019)	✓ *	B *	C	C	X	✓ *	
(Sun et al., 2021)	✓ *	B *	C	C	X	X	2: weinig informatie over de methode (bijvoorbeeld selectie van participanten)

(Zie volgende pagina voor vervolg)

(Vervolg van vorige pagina)

Studie	*1. Onderzoeksvraag & - doel*	*2. kwaliteit methode*	*3. werving deelnemers*	*4. Vergelijkbare groepen*	*5. power*	*6. effectwaarde*	*Evt. opmerking over oordeel*
(Vaahtoranta et al., 2021)	✓ *	B *	C	C	✓ *	X	2: aanbeveling: meer interne/ externe factoren
(Wong & Samudra, 2019)	X	B *	C	N.v.t	✓ *	X	1: het nut van de vraag hoe visuele aandacht en Dual-Coding in verband staan, is niet helemaal duidelijk
(Wong et al., 2021)	✓ *	B *	C	N.v.t	✓ *	X	
(Wong & Neuman, 2021)	✓ *	A **	C	N.v.t	✓ *	X	2: within-subjects design was zeer passend

Tabel D2

Vraag 7 t/m 11 van het Kwaliteitsassessment

Studie	*7. Steekproef - populatie*	*8. meting onafhankelijke var.*	*9. meting afhankelijke var.*	*10. periode*	*11. Overige limitaties*	*Evt. opmerking over oordeel*
(Baker et al., 2021)	A	A	B	A	A	
	**	***	**	**	**	
(Garcia, 2021)	A	A/B	A/B	B	B	
	**	**	**	*	*	
(Grøver et al., 2020)	A	N.v.t.	A/B/C	B	B	
	**		*	*	*	
(Hadley et al., 2021)	B	A/B/C	A	B	B	9: interbeoordelaars-betrouwbaarheid: >90%, maar geen andere
	*	*	*	*	*	betrouwbaarheid gegeven (daarom 1 ster)

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*7. Steekproef - populatie*	*8. meting onafhankelijke var.*	*9. meting afhankelijke var.*	*10. periode*	*11. Overige limitaties*	*Evt. opmerking over oordeel*
(Janssen et al., 2019)	B *	A ***	C *	C	A **	
(Leeuwestein et al., 2021)	B *	N.v.t.	F	C	B *	
(Miles et al., 2019)	B *	A/B/C/D	N.v.t	C	B *	
(Pollard-Durodola et al., 2021)	A **	N.v.t.	A/B **	C	B *	
(Reilly et al., 2020)	A **	A/B **	A/B **	A **	A **	Opvallend: grote *n*

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*7. Steekproef - populatie*	*8. meting onafhankelijke var.*	*9. meting afhankelijke var.*	*10. periode*	*11. Overige limitaties*	*Evt. opmerking over oordeel*
(Rivera Pérez et al., 2019)	A	N.v.t.	B	C	A	
	**		**		**	
(Silverman et al., 2019)	A	N.v.t.	A	B	B	
	**		***	*	*	
(Sun et al., 2021)	B	A/B	C	A	B	7: onvoldoende informatie
	*	**	*	**	*	
(Vaahtoranta et al., 2021)	B	B/C	A/B	C	B	
	*	*	**		*	
(Wong & Samudra, 2019)	B	A	F	C	B	
	*	***			*	

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*7. Steekproef - populatie*	*8. meting onafhankelijke var.*	*9. meting afhankelijke var.*	*10. periode*	*11. Overige limitaties*	*Evt. opmerking over oordeel*
(Wong et al., 2021)	B *	N.v.t.	D	C	B *	7: verdeling van afkomst niet representatief (76% Spaans, 18% Haïti-Creools)
(Wong & Neuman, 2021)	C	N.v.t.	D	C	A **	7: verdeling van afkomst niet representatief (ongeveer 75% Spaans, 25% Haïti-Creools)

Bijlage E

Studie-eigenschappen en Resultaten (Studies Alfabetisch Geordend)

Studie	*Kernbevindingen*	*Effect-grootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting (& follow-up)*	*Kwaliteit*
Baker et al. (2021)	Grotere Spaanse RW bleek een significante voorspeller voor grotere Engelse RW en EW	X	OV: Spaanse RW**:** Test de Vocabulario en Imágenes Peabody AV: (1) EW: Expressive measure of target words (ETW) (door de onderzoekers ontworpen) (2) RW: Receptive measure of target words (RTW) (door de onderzoekers ontworpen)	OV: $r = .93$[a] AV: (1) $\alpha = .88$ (2) $\alpha = .82$ (3) IBBH: r>90%:	Herfst (vm) → Eind van het jaar (nm) ≈ 9/10 maanden Follow-up: einde van *first grade* = 2 jaar sinds vm	16/21 (76.2%) Prima

Noot. IBBH = Interbeoordelaars-betrouwbaarheid, EW = Expressieve Woordenschat, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

[a] Betrouwbaarheidsmaat niet gespecificeerd

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effectgrootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting (& follow-up)*	*Kwaliteit*
Garcia (2021)	Het hebben van klasgenoten met grote RW bleek een significante positieve invloed op RW-groei te hebben. Variabiliteit had geen significant effect.	X	<u>OV:</u> Woordenschat van klasgenoten (*peers*): berekend als gemiddelde per kind, o.b.v. de scores op de PPVT van minimaal 4 (één/meertalige) klasgenoten <u>AV:</u> RW**:** PPVT- 4e editie	<u>OV:</u> - <u>AV:</u> IC: *r* =.96/ .97 Andere vormen van betrouwbaarheid: 87 < *r* < .93	Herfst (vm) → Lente (nm) ≈ Half jaar	12/21 (57.1%) Vrij laag

Noot. IC = Interne Consistentie, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting, PPVT = Peabody Picture Vocabulary Test

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effectgrootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaar-heid (r) instrumenten*	*Duur tussen vm & nm*	*Kwaliteit*
Grøver et al. (2020)	De onderzochte samen-lees-interventie bleek een significant, positief effect te hebben op RW- en EW-groei, maar enkel voor targetwoorden, niet voor algemene woordenschat.	✓ EW: *d* = .85 Target-RW: *d* = 1.15 Algemene RW: *d* = .68	<u>OV:</u> Interventie 'Extend Intervention': i.h.k.v. 4 thema's lezen ouders/professionals boeken m.b.v. een vrij te volgen script voor verdiepende gesprekken over targetwoorden. <u>AV:</u> (1) RW: 'Targeted receptive vocabulary' (door de onderzoekers ontworpen) (2) EW: Targeted expressive vocabulary' (door de onderzoekers ontworpen) (3) Voor de woorden die niet in de interventie voorkwamen (RW): British Picture Vocabulary Scale – 2	<u>OV:</u> - <u>AV:</u> (1) α=.74 (2) α= .88 (3) IBBH: *r* =.86 (4) α = .93	7,4 maanden	13/18 (72,2%) Prima

Noot. IBBH = Interbeoordelaars-betrouwbaarheid, EW = Expressieve Woordenschat, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effect-grootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting & follow-up*	*Kwaliteit*
Hadley et al. (2021)	EW-ontwikkeling bleek significant afhankelijk van beeldbaarheid (abstractheid) van woorden, executief functioneren en algemene woordenschat. Ook concreetheid en vorm van woorden hingen positief significant samen met EW-scores. Effecten waren sterker voor oudere kinderen en kinderen met een grotere algemene woordenschat.	X	<u>OV:</u> *Kind-kenmerken:* (1) Algemene RW (voormeting): PPVT-4 (2) Executief functioneren: Peg Tapping task *Woordkenmerken:* (3) Measure of Words' Perceptual Accessibility <u>AV:</u> EW (van targetwoorden): New Word Definition Test-Modified (NWDT-M) (door de onderzoekers ontworpen)	<u>OV:</u> (1) T-H: $.92 < r < .96$ (1) IC: $.94 < r < .95$ (2) T-H: $r = .80$ (3) IBBH: $r = 80\%$ <u>AV:</u> IBBH: $r > 90\%$	7 maanden Follow-up: 4,5 maand na nm	9/21 (42.9%) Laag

Noot. PPVT = Peabody Picture Vocabulary Test, IBBH = Interbeoordelaars-betrouwbaarheid, T-H = Test-hertest betrouwbaarheid, IC = Interne Consistentie, EW = Expressieve Woordenschat, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effect-grootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouw-baarheid (r) instrumenten*	*Duur tussen voor- en nameting & follow-up*	*Kwaliteit*
Janssen et al. (2019)	Vormgericht en betekenisgericht woordleren zorgden voor een gelijke toename in oppervlakkige kennis van target-RW, maar vormgerichte instructie zorgde voor significant meer semantische kennis. Het kunnen leren van woordparen die op elkaar lijken qua klank was een significante voorspeller hiervan.	X	<u>OV:</u> (1) Interventie: interactief boeklezen met vormgerichte of betekenisgerichte instructie (2) Algemene woordenschat: Taaltoets Alle Kinderen (3) Vaardigheid om woorden die fonetisch minimaal verschillen te leren: woordleertaak <u>AV:</u> (4) Targetwoorden RW: Door onderzoekers ontworpen test (5) Targetwoorden semantische kennis: Door onderzoekers ontworpen test	<u>OV:</u> (1) - (2) α = .97 (3) α = .96 <u>AV:</u> (4) α =.72 (5) α = .80	September (vm) – > november (nm) = 3 maanden	12/21 (57.1%) Vrij laag

Noot. RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effectgrootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting (& follow-up)*	*Kwaliteit*
Leeuwestein et al. (2021)	Sociale robots bleken significant effect te hebben bij het aanleren van RW. Een robot die Turkse vertalingen gaf bij de Nederlandse woorden had geen significant verschil in effect ten opzichte van een robot die enkel de Nederlandse woorden gaf.	X	<u>OV:</u> Interventie: Sociale robots die Nederlandse woorden aanleren, waarvan 1 de Turkse vertaling erbij geeft. <u>AV:</u> (1) Target-RW: plaatjes-selectie-test (door de onderzoekers ontworpen) (2) RW (NL en Turks): Diagnostic Test of Bilingualism	<u>OV:</u> - <u>AV:</u> (1) X (2) .89 <α<93	Vm & nm op dezelfde dag Follow-up: 1 week na nm	4/18 (22.2%) Zeer laag

Noot. RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effect-grootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen vm & nm*	*Kwaliteit*
Miles et al. (2019)	Alfabetkennis modereerde de significante, positieve relatie tussen taalvaardigheden en het leren van losse woorden zonder context via flashcards. Het leren van inhoudswoorden werd significant voorspeld door fonetisch bewustzijn, dat van functiewoorden door RW.	X	<u>OV:</u> (1) Interventie: Flashcards met woorden in isolatie (zonder context). Drie inhoudswoorden en drie functiewoorden per training van 15 min. (2) Alfabetkennis: door onderzoekers ontworpen test: of kinderen wel of niet pseudo-woorden kunnen oplezen (3) RW: PPVT-4 (Formulier A (functiewoorden) en formulier B (inhoudswoorden)) (4) Spelling kennis: door onderzoekers ontworpen test (5) Woordlezen: Word Identification subtest of the Woodcock Reading Mastery Test-III (6) Fonetisch bewustzijn: Sound Matching subsection of the Comprehensive Test of Phonological Processing <u>AV:</u> Door onderzoekers ontworpen flashcardtest om target-woorden (RW) te toetsen	<u>OV:</u> (1) N.v.t. (2) X (3) S-H: A: $r = .94$ B: $r = .95$ (4) $\alpha = .66$ (5) $.86 < \alpha < .99$ (6) T-H: $r = .70 < r < .92$ S-H: $r = .75$ <u>AV:</u> (7) X	Vm & nm op dezelfde dag	7/18 (38.9%) Zeer laag

Noot. PPVT = Peabody Picture Vocabulary Test, S-H = Split-half betrouwbaarheid, T-H = Test-hertest betrouwbaarheid, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effectgrootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting*	*Kwaliteit*
Pollard-Durodola et al. (2021)	De visie van leraren op een lees-interventie, bijvoorbeeld op gebied bruikbaarheid en effectiviteit (= sociale validiteit), bleek een significant effect te hebben op target-RW en EW van kinderen.	X	<u>OV:</u> Vragenlijst naar sociale validiteit van de interventie <u>AV:</u> RW: (1) PPVT – 4 (2) Researcher-Developed Receptive Picture Vocabulary Test (RDRPVT) EW: (3) Expressive Vocabulary Test-Second Edition (EVT-2) (4) Researcher-Developed Expressive Picture Vocabulary Test (RDEPVT)	<u>OV:</u> X (gevalideerd in twee eerdere onderzoeken) <u>AV:</u> (1) $r^a > .90$ (2) $r^a = .86$ (3) X (4) $r^a = .88$	2 weken voor en na interventie van 5 tot 14 weken afgenomen = 9 tot 18 weken ertussen	10/18 (55.6%) Vrij laag

Noot. PPVT = Peabody Picture Vocabulary Test, EW = Expressieve Woordenschat, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

[a] Betrouwbaarheidsmaat niet gespecificeerd

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effectgrootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting & follow-up*	*Kwaliteit*
Reilly et al. (2020)	Klaslokalen met veel instructie, geen over-structurering en meer ruimte voor door het kind geïnitieerde activiteiten zorgden voor significant hogere EW- en RW-scores (ook in first grade nog). Meer emotionele steun hing significant samen met meer groei in RW.	X	OV: Classroom Assessment Scoring System AV: (1) EW: Expressive One-Word Picture Vocabulary Test (2) RW: PPVT-4	OV: S-H: $.88 < r < .91$ AV: (1 & 2): IC: $.95 < \alpha < .97$ T-H: $.85 < \alpha < .96$	Vm: herfst 2009 Nm: lente 2010 ≈ 8 maanden Follow-ups: lente 2011 en 2012 = 1 & 2 jaar	17/21 (80.1%) Goed

Noot. PPVT = Peabody Picture Vocabulary Test, S-H = Split-half betrouwbaarheid, IC = Interne consistentie, T-H = Test-hertest betrouwbaarheid, EW = Expressieve Woordenschat, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effectgrootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting & follow-up*	*Kwaliteit*
Rivera Pérez et al. (2019)	Gebruik van audio opnames bij woordenschatinstructie bleek significant effect te hebben op het benoemen van woorden. Expressief definiëren bleek significant voor kinderen die instructie in het Spaans en Engels gemengd kregen, niet voor hen die alleen Engels kregen.	X	<u>OV:</u> Interventie (Spaans-Engelse instructie, enkel Engelse-instructie of controlegroep) <u>AV:</u> Test door onderzoekers: 24 target woorden getoetst d.m.v. bijpassende foto's. Benoemen en definiëren gevraagd	<u>OV:</u> Steekproef werd 100& volgens protocol uitgevoerd <u>AV:</u> IC: Engels: α = .83 Spaans: α = .71 IBBH (obv 10%-steekproef): benoemen: 98%, definiëren: 92%	Vm→ nm: 6 weken Follow-up: 6 weken na nm	11/18 (61.1%) Matig

Noot. IC = Interne consistentie, IBBH = Interbeoordelaars-betrouwbaarheid, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effectgrootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting*	*Kwaliteit*
Silverman et al. (2019) (1/2)	Van verschillende soorten input (tekst op papier, video en digitale tekst) was video-instructie effectiever dan tekst op papier bij de uitkomsten op EW. Video was ook effectiever dan digitale tekst voor EW- en RW-ontwikkeling. Digitale tekst was hierbij het minst effectief.	✓ Tekst op papier: *d*=.25 Video: *d*=.64 Digitale tekst: *d*= .14	<u>OV:</u> Interventie: Martha's True Stories (MTS) Buddies Program (peuters kregen een lees-buddy uit grade 4[a]). Experimentele groep kreeg extra les met de verschillende soorten input, waarin de targetwoorden uit het buddy-programma werden besproken. <u>AV:</u> RW: (1) PPVT-4	<u>OV: -</u> <u>AV:</u> (1) S-H: *r* = .94 T-H: *r* = .95 (2) Vm: H[b]= .87 Nm: H = .93 (1 & 2): Correlatie RAV-K & PPVT: vm: .66 & nm: .72	Oktober/ November (vm) ➔ April/Mei (nm) = 6 maanden	10/18 (55.6%) Vrij laag

Noot. PPVT = Peabody Picture Vocabulary Test, S-H = Split-half betrouwbaarheid, T-H = Test-hertest betrouwbaarheid, EW = Expressieve Woordenschat, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

[a] Kinderen uit grade 4 in Amerika zijn rond de 9/10 jaar oud

[b] Alternatief voor Crohnbach's alpha

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effect-grootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting*	*Kwaliteit*
Silverman et al. (2019) (2/2)			(2) Receptive Assessment of Vocabulary - Kindergarten (RAV-K) (door onderzoekers gemaakt) EW: (3) Expressive Vocabulary Assessment-Kindergarten (EVA-K) (door onderzoekers gemaakt)	(1) IBBH: *r* = .90 vm: H = .97 nm: H = .95 (1 & 3): Correlatie EVA-K en PPVT: vm: .66 & nm: .68		
Sun et al. (2021) (1/2)	De invloed van externe factoren (input kwantiteit thuis en op school, aantal boeken en leeshoeveelheid thuis, opleiding en ervaring van de leraar, SES) en interne,	X	<u>OV:</u> (1) Externe factoren: vragenlijsten aan ouders en leraren Interne factoren: (2) *Werkgeheugen:* Backward Digit Recall (door onderzoekers zelf aangepast)	<u>OV:</u> (1) X (2) α = .83 (3) α = .82 (4) α = .92	1 jaar	9/21 (42.9%) Laag

Noot. IBBH = Interbeoordelaars-betrouwbaarheid, EW = Expressieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effect-grootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting*	*Kwaliteit*
Sun et al. (2021) (2/2)	kind-gerelateerde factoren (werkgeheugen, nonverbale intelligentie, leer-gerelateerde sociale vaardigheden, leeftijd, gender) op RW waren ongeveer gelijk. Met name non-verbale intelligentie sociale vaardigheden en het opleidingsniveau van de moeder voorspelden RW.		(3) *Nonverbale intelligentie:* Raven's Colored Progressive Matrices (4) *Leergerelateerde sociale vaardigheden:* Child Behavior Rating Scale <u>AV:</u> RW: Bilingual Language Assessment Battery	<u>AV:</u> α = .77		
Vaahtoranta et al. (2021) (1/2)	RW en EW werden voorspeld door L2-taalvaardigheid van de moeder, fonologisch kortetermijngeheugen en	X	<u>OV:</u> Fonologisch kortetermijngeheugen: (1) NWR-G: 'Phonological Working Memory for Non-Words' van de	<u>OV:</u> (1) α = .83 (2) α = .72 (3) α = .88	5 maanden	8/21 (38.1%) Zeer laag

Noot. EW = Expressieve Woordenschat, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*
Vaahtoranta et al. (2021) (2/2)	de duur op lange termijn van de blootstelling aan L2. Huidige dagelijkse blootstelling buiten school aan L1 en L2 voorspelde geen woordenschatontwikkeling.	testbatterij 'Development Test for Three- to Five-Year-Old Children' (Test taal-specifieke (Duits-achtige) pseudo-woorden) (2) NWR-U: 'Quasi-Universal Nonword Repetition Task' (test taal-algemene pseudowoorden) Taalblootstelling: (3) Ouders vulden een dagboek-vragenlijst in (subschaal over taalervaringen) <u>AV:</u> (1) RW: Duitse versie van de PPVT-4 (2) EW: door onderzoekers ontworpen versie van de 'Expressive Vocabulary Test-2' (EVT-2)	<u>AV:</u> (1) S-H: $r = .96$ (2) S-H: Vm: $r = .84$ Nm: $r = .97$

Noot. PPVT = Peabody Picture Vocabulary Test, EW = Expressieve Woordenschat, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, Vm = Voormeting, Nm = Nameting, S-H = Split-half betrouwbaarheid

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effect-grootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting*	*Kwaliteit*
Wong et al. (2021)	Gerichte leermomenten, om precies te zijn d.m.v. video's gebaseerd op verklaren en participatie, waren effectiever op het gebied van simpele RW-groei dan toevallige leermomenten (d.m.v. verhalende video's). Voor dieper woordbegrip was alleen het verklarende type instructie bij de groep met hogere L2-RW op de voormeting verklarend.	X	<u>OV:</u> Interventie: 9 korte video's met targetwoorden. Drie typen: (A) kijker werd aangesproken en aangemoedigd om te participeren (meest gericht leren), (B) expliciete visuele-auditieve voorbeelden (tussen toevallig en gericht leren in) en (C) een verhaal (minst gericht leren). <u>AV:</u> (1) Vm RW: PPVT-IV Nm RW: twee door onderzoekers gecreëerde test: (2) *Simpele woordenschat:* 18 items (3) *Dieper woordbegrip:* 27 items	<u>OV:</u> - <u>AV:</u> (1) $.91 < \alpha < .94$ (2) $\alpha = .61$ (3) $\alpha = .69$	Vm en nm vonden op dezelfde dag plaats	6/16 (37.5%) Zeer laag

Noot. RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effect-grootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*Betrouwbaarheid (r) instrumenten*	*Duur tussen voor- en nameting*	*Kwaliteit*
Wong & Neuman (2021)	Kinderen die thuis niet veel worden blootgesteld aan de meerderheidstaal bleken met name baat te hebben bij video's die eerder geleerde woorden herhalen. Kinderen die thuis meer blootstelling hebben, bleken baat te hebben bij andere typen video's (met expliciete definities, visuele en geluidseffecten).	X	<u>OV:</u> Interventie: verschillende soorten digitale technieken in video's <u>AV:</u> RW: door onderzoekers ontworpen tests: (1) Woordenschat in-context test target-woorden aanwijzen op één van de drie plaatjes, wat screenshots waren van Sesamstraat (de context waarin het woord is gepresenteerd aan de kinderen) (2) Woordenschat in-nieuwe-context test: target-woorden aanwijzen op één van de drie plaatjes, maar hier werden ze in een nieuwe context gepresenteerd	<u>OV:</u> - <u>AV:</u> (1) α =.61 (2) α =.60	Vm en nm vonden op dezelfde dag plaats	6/16 (37.5%) Zeer laag

Noot. RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting

(Vervolg op volgende pagina)

(Vervolg van vorige pagina)

Studie	*Kernbevindingen*	*Effect-grootte*	*Meetmanier onafhankelijke en afhankelijke variabele*	*r* van *instrumenten*	*Duur tussen vm & nm*	*Kwaliteit*
(Wong & Samudra, 2019)	Video's die gebruik maken van *Dual-Coding* (auditieve en visuele informatie van het woord tegelijk presenteren) waren effectiever in stimulatie van RW, met name voor kinderen die thuis een minder stimulerende L2-taalomgeving hebben. Visuele aandacht blijkt alleen een voorspeller voor RW wanneer aan DC wordt gedaan.	X	OV: (1) Interventie: 3 video's waarbij de visuele representatie van het woord en de audio van het woord tegelijk werden afgespeeld en 3 video's waarbij dit na elkaar was. Alle participanten kregen alle video's te zien (*Within-subjects* design) (2) Aandachtspan: eye-tracking methode (met een Tobii Technology T120 eye-tracker) (3) Taalvaardigheid ouders (blootstelling aan Engels thuis): vragenlijst (4) Baseline Engelse RW: PPVT- IV AV: 12-item woorden-identificatietest (door onderzoekers ontworpen). 6 woorden met en 6 woorden zonder context	OV: (1) – (2) X (3) X (4) .91 < α < .94) AV: X	Vm en nm vonden op dezelfde dag plaats	7/19 (36.8%) Zeer laag

Noot. PPVT = Peabody Picture Vocabulary Test, RW = Receptieve Woordenschat, OV = Onafhankelijke Variabele, AV = Afhankelijke Variabele, vm = Voormeting, nm = Nameting, *r* = Betrouwbaarheid