

Effect van een Robot op het aanleren van Productieve Woordenschat in een Tweede Taal bij
Kleuters

Master's thesis (201600204)

Studiejaar 2016-2017

Universiteit Utrecht

Masteropleiding Pedagogische Wetenschappen

Masterprogramma Clinical Child, Family and Education Studies

Lisa G. A. Limpens (4173554)

Supervisor: Sanne van der Ven

Tweede beoordelaar: Cathy van Tuijl

Datum: 23-6-2017

Voorwoord

Voor u ligt het afstudeeronderzoek waarin het effect van een robot op het aanleren van productieve woordenschat in een tweede taal bij kleuters is bestudeerd. Dit onderzoek is onderdeel van het Europese L2TOR project.

Ondanks dat de afgelopen maanden zeer intensief zijn geweest, heb ik dit onderzoek vooral als ontzettend leerzaam en leuk ervaren. Bij dezen bedank ik Sanne van der Ven voor haar ondersteuning en zeer prettige begeleiding in het gehele proces van mijn masterscriptie. Ook wil ik graag Rianne van den Berghe bedanken voor de ondersteuning en fijne samenwerking tijdens de testmomenten met de robot. Verder bedank ik mijn medestudenten die zich hebben ingezet voor onze dataverzameling. Ten slotte wil ik graag de kinderen die hebben deelgenomen aan het onderzoek bedanken voor hun grote inzet en de scholen voor hun gastvrijheid.

Ik wens u veel leesplezier.

Utrecht, 23 juni 2017

Lisa Limpens

Abstract

Aim: Current research studies the effect of a robot on teaching productive vocabulary [PV] in a second language to Dutch toddlers. For this purpose, toddlers played a game on a tablet with or without a peer or a robot. It's also studied whether there's a difference in learning outcomes and perception with respect to the robot between boys and girls.

Method: The sample involved 68 kids, of which 39 boys and 29 girls, aging between 52 and 78 months ($M = 67.60$, $SD = 6.92$). They were classified in child-, peer- or robot condition. On both short and long term (one week after vocabulary training) their English PV was measured. Their perception with respect to the robot was determined as well.

Results: The vocabulary training appears to improve the English PV of toddlers significantly. Condition has only a small effect on learning outcomes. Within the robot condition, sex has a medium effect on learning outcomes. However, these effects of condition and sex on learning outcomes are not significant, just as the effect of sex on perception.

Conclusion: English PV of toddlers doesn't appear to improve to a greater extent due to teaching with a tablet and a robot with respect to teaching with only a tablet or a tablet and a peer. Specifically, concerning learning with a robot, there's no difference between boys and girls concerning learning outcomes and perception with respect to the robot.

Keywords: Second language tutoring, social robot, productive vocabulary, preschool children

Samenvatting

Doel: In huidig onderzoek is het effect van een robot op het aanleren van productieve woordenschat [PW] in een tweede taal bij Nederlandstalige kleuters onderzocht. Hiertoe hebben kleuters een spel (vocabularyretraining) op een tablet gespeeld, al dan niet samen met een leeftijdgenoot of robot. Tevens is onderzocht of jongens en meisjes verschillen in leeropbrengst en in hun perceptie ten aanzien van de robot.

Methode: De steekproef betrof 68 kinderen, waarvan 39 jongens en 29 meisjes, in de leeftijd van 52 en 78 maanden ($M = 67.60$, $SD = 6.92$). Zij werden ingedeeld in de kind-, peer- of robotconditie. Op zowel korte, als lange termijn (één week na de vocabularyretraining) is hun Engelse PW gemeten. Ook werd hun perceptie ten aanzien van de robot vastgesteld.

Resultaten: De vocabularyretraining blijkt in elke conditie significant de Engelse PW van kleuters te bevorderen. Conditie blijkt slechts een klein effect te hebben op de leeropbrengst. Binnen de robotconditie heeft Sekse een medium effect op de leeropbrengst. Echter zijn deze effecten van Conditie en Sekse op de leeropbrengst niet significant, evenals het effect van Sekse op perceptie.

Conclusie: De Engelse PW van kleuters blijkt niet in meerdere mate bevorderd te worden door het leren met een tablet en robot dan door het leren met enkel een tablet of met een tablet en leeftijdgenoot. Specifiek betreffende het leren met de robot, verschillen jongens en meisjes niet in leeropbrengst en hun perceptie ten aanzien van de robot.

Keywords: Tweede taal leren, sociale robot, productieve woordenschat, kleuters

Het effect van een robot op het aanleren van productieve woordenschat in een tweede taal bij kleuters

Huidig onderzoek bestudeert het effect van een tablet, leeftijdgenoot en robot op het leren van Engels aan Nederlandstalige kleuters. In het Nederlandse onderwijs blijkt namelijk het promoten van tweetaligheid en de belangstelling voor het Engels toe te nemen. Steeds meer scholen bieden reeds in de onderbouw Engels aan, soms al vanaf groep 1 (Aarts & Ronde, 2006). Tevens zijn op scholen technologieën zoals computers en tablets niet meer weg te denken als hulpmiddel (Chinnery, 2006; Stockwell, 2007). Zulke systemen kunnen ondersteuning bieden tijdens educatieve activiteiten en onder andere worden ingezet bij het leren en/of bevorderen van een (tweede) taal (Baloian, Pino, & Vargas, 2013; Chanprasert & Han, 2013; Golonka, Bowles, Frank, Richardson, & Freynik, 2012; Haßler, Major, & Hennessy, 2016).

Wetenschappelijk onderzoek toont namelijk dat het gebruik van een tablet doorgaans significant bijdraagt aan het bevorderen van Engelse woordenschat als tweede taal bij vijf- en zesjarige kinderen (Baloian et al., 2013). Hetzelfde geldt voor interactie met een menselijke leeftijdgenoot, een proximaal proces, tijdens een spelactiviteit (Bronfenbrenner & Morris, 2006; Bulotsky-Shearer et al., 2012; Justice, Petscher, Schatschneider, & Mashburn, 2011; Lin, Justice, Paul, & Mashburn, 2016). Tevens blijkt leren middels samenwerking met een sociale robot, die zich als leeftijdgenoot kan gedragen, in combinatie met een tablet de Engelse woordenschat van kleuters als tweede taal significant te bevorderen (Mazzoni & Benvenuti, 2015; Tanaka & Matsuzoe, 2012; Westlund & Breazeal, 2015). Uit onderzoek volgt dat deze samenwerking met een robot effectiever is dan zowel leren met enkel een tablet, als samenwerken met een leeftijdgenoot (Kose-Bagci, Ferrari, Dautenhahn, Syrdal, & Nehaniv, 2009; Mazzoni & Benvenuti, 2015).

Kennedy, Baxter, en Belpaeme (2015) concluderen in hun onderzoek dat, hoewel niet significant, meisjes in meerdere mate vooruitgang tonen dan jongens na het leren met een tablet in combinatie met een robot. In de studie van Schermerhorn, Scheutz, en Crowell (2008) is dit verschil daarentegen wel significant. Verschillende auteurs suggereren dat dit genderverschil het gevolg is van uiteenlopende percepties die jongens en meisjes ten aanzien van robots blijken te hebben (Kennedy et al., 2015; Nomura, Kanda, Suzuki, & Kato, 2008; Schermerhorn et al., 2008). Meisjes zouden, vergeleken met jongens, robots bijvoorbeeld meer als machine zien en minder van de inhoud afgeleid worden door sociaal gedrag van de robot (Kennedy et al., 2015).

De belangstelling voor het toepassen van een robot (als ondersteuning) in het onderwijs lijkt toegenomen, ook binnen wetenschappelijk onderzoek (Chang, Lee, Chao, Wang, & Chen, 2010). Echter, naar aanleiding van reeds besproken literatuur kan geconcludeerd worden dat onderzoek gericht op tweede taalverwerving nog relatief beperkt is. Vooral ten aanzien van (Nederlandse) kleuters is dit zeer schaars (Chang et al., 2010; Ortega & Iberri-Shea, 2005). Onderzoek naar een eventueel genderverschil in het effect van het leren met een robot op leerprestaties, en mogelijke verklaringen, is bijvoorbeeld enkel uitgevoerd bij achtjarigen en studenten (Kennedy et al., 2015; Schermerhorn et al., 2008). Ten slotte ontbreken doorgaans effectgroottes en vergelijkingen met andere methoden zoals leren met een tablet en al dan niet een leeftijdgenoot, waardoor inzicht in het daadwerkelijke voordeel van een robot voor het leerproces van kinderen beperkt is.

Huidige, longitudinale en toetsende, pilotstudie is onderdeel van het L2TOR project (<http://www.l2tor.eu>). Er wordt gebruik gemaakt van een NAO-robot, genaamd Robyn, welke sociaal gedrag vertoont door bijvoorbeeld te kunnen praten en bewegen. Uitkomsten van huidig onderzoek kunnen eerder benoemde kennisgebreken verkleinen. Dit kan bijdragen aan optimalisering van Engelse taallessen aan Nederlandstalige kleuters en de toepassing van sociale robots als leermaatje. Hiertoe wordt antwoord gezocht op de vraag: “In hoeverre heeft een robot effect op het aanleren van productieve woordenschat in een tweede taal bij kleuters?”. Ten eerste wordt het effect van (1) leren met enkel een tablet en van (2) leren met een tablet en leeftijdgenoot vergeleken met het effect van (3) leren met een tablet en een robot. Daarnaast worden, wat betreft het effect van een robot, eventuele verschillen tussen jongens en meisjes bestudeerd, evenals de mate waarin hun perceptie ten aanzien van de robot hier mogelijk aan bijdraagt.

Op basis van reeds besproken literatuur is de verwachting dat, met betrekking tot de eerste onderzoeksvraag, de groep kinderen die leert met de tablet en robot meer productieve woordenschat in een tweede taal aanleert dan de groep kinderen die enkel met de tablet of met de tablet en een leeftijdgenoot leert. De tweede hypothese is dat, betreffende kinderen die leren in samenwerking met de robot, meisjes een grotere leeropbrengst behalen dan jongens. Hoewel de onderzoeken van Kennedy en collega's (2015) en Schermerhorn en collega's (2008) betrekking hebben op andere taken (planeten en wiskunde), blijken gevonden sekseverschillen namelijk niet veroorzaakt te zijn door de (inhoud van de) vragen die gesteld werden. Indien de tweede hypothese wordt aangenomen, luidt de derde hypothese: ‘De perceptie van jongens en meisjes ten aanzien van de robot draagt bij aan het sekseverschil in effect van de robot op hun productieve, Engelse woordenschat’. Hiertoe dient tevens een

verband te zijn tussen sekse en de perceptie van de kinderen enerzijds en perceptie en de leeropbrengst anderzijds, dus worden ook deze effecten binnen de derde hypothese verwacht en onderzocht.

Methodie

Participanten

De steekproef bestond uit 68 kinderen, namelijk 39 jongens (57%) en 29 meisjes (43%). Zij waren in de leeftijd van 52 tot en met 78 maanden ($M = 67.60$, $SD = 6.92$). De kinderen zaten in groep 1 of groep 2 van een reguliere basisschool en hadden Nederlands als moedertaal. Tweektaligheid gold als exclusie criterium.

De kinderen in de steekproef werden over drie onderzoeksgroepen verdeeld (zie Tabel 1). Zij gingen namelijk leren met enkel een tablet (kindconditie), met een tablet en leeftijdgenoot (peerconditie) of met een tablet en de robot (robotconditie). Uit een Levene's test volgt geen gelijkheid in variantie tussen deze groepen betreffende leeftijd, $F(2, 65) = 5.15$, $p = .008$. Tevens toonde een ANOVA analyse significante verschillen in gemiddelde leeftijd tussen de drie groepen, $F(2, 65) = 3.31$, $p = .043$, $\eta_p^2 = .092$. Uit de Games-Howell post-hoc toets bleek de gemiddelde leeftijd in de kindconditie namelijk significant hoger te zijn dan de gemiddelde leeftijd in de robotconditie ($M_{\text{verschil}} = 4.77$, $p = .037$).

Tabel 1

Beschrijvende Statistieken per Conditie

Conditie	N	Geslacht ^a		Leeftijd ^b	
		Jongen	Meisje	M_{leeftijd}	SD_{leeftijd}
Kind	23 (34%)	12 (52%)	11 (48%)	70.48	4.25
Peer	22 (32%)	15 (68%)	7 (32%)	66.57	7.51
Robot	23 (34%)	12 (52%)	11 (48%)	65.70	6.92

Noot. ^a = Aantal jongens en meisjes per conditie, ^b = Leeftijd in maanden.

Procedure

Op basis van een gemakssteekproef werd telefonisch contact opgenomen met directies van enkele reguliere basisscholen. Gedurende het telefoongesprek, en indien gewenst ook per mail, werd het onderzoek toegelicht. Directies en leerkrachten kregen enkele dagen tijd om deelname te overwegen en eventuele vragen en suggesties met betrekking tot het onderzoek kenbaar te maken. Vijf scholen hebben aan het onderzoek deelgenomen: twee in Gelderland,

één in Noord-Holland, één in Brabant en één in Limburg. Deelnemende kinderen hadden actieve toestemming van hun ouders.

Onderzoeksinstrumenten

Productieve, Engelse woordenschat. Onderzoekers hebben twee vertaaltaken samengesteld (zie Bijlage 1 en 2). Scores op deze taken representeerden de mate van productieve, Engelse woordenschat van de kinderen. Beiden bestonden uit zes items, betreffende de zes targetwoorden die in de vocabulairetraining werden aangeboden: *heavy*, *light*, *full*, *empty*, *in front of* en *behind*. Eén taak deed een beroep op het vertalen van deze woorden van het Engels naar het Nederlands (EN-NL), bijvoorbeeld middels het item ‘Wat is *heavy*?’. De andere taak vroeg om de Engelse vertaling van het Nederlandse woord (NL-EN), bijvoorbeeld ‘Wat is zwaar?’.

Vertaaltaak EN-NL werd op drie momenten bij de kinderen afgenomen: zowel vóór de leerfase (pretest) om te controleren of zij de Engelse woorden reeds kenden, als direct en één week na de leerfase (post- en late test) om te onderzoeken hoeveel zij op korte en lange termijn hadden geleerd. Vertaaltaak NL-EN is tijdens de post- en late test afgenomen. Wanneer de kinderen tijdens de post- en/of late test het antwoord op een item niet wisten of aanvankelijk fout hadden, ontvingen zij een prompt (bijvoorbeeld ‘Is *heavy* zwaar of licht?’). Zij scoorden voor een goed antwoord één punt, een goed antwoord na een prompt 0,5 punt en een fout antwoord 0 punten. Dit resulteerde in de scores EN-NL Post, NL-EN Post, EN-NL Late en NL-EN Late. Hoewel de pretest door alle (68) kinderen werd doorlopen, was dit aantal wegens ziekte bij de posttest 67 en bij de late test 66.

Perceptie. De kinderen in de robotconditie werden tijdens de posttest naar hun perceptie ten aanzien van de robot gevraagd: ‘Vind je Robyn meer een meester/juf of meer een vriendje/vriendinnetje?’, ‘Vind je Robyn meer een vriendje/vriendinnetje of meer een ding?’ en ‘Vind je Robyn meer een ding of meer een meester/juf?’. De optie die een kind het vaakst antwoordde, representeerde zijn/haar perceptie. Wanneer een kind ‘meester/juf’, ‘vriendje/vriendinnetje’ en ‘ding’ elk één keer koos, werd dit gescoord als zijnde een ‘Onduidelijke perceptie’. Dit resulteerde in vier categorieën.

Betrouwbaarheid en validiteit. De onderzoeksinstrumenten betroffen taken waarvan de betrouwbaarheid en validiteit nog onbekend zijn. Test-hertest betrouwbaarheid is bijvoorbeeld aannemelijk, maar vanwege leereffecten (op korte en lange termijn) niet concreet vast te stellen. Middels afstemming van de afname en scoring tussen verschillende onderzoeksters werd interbeoordelaarsbetrouwbaarheid nagestreefd. Daarnaast waren instructies gestandaardiseerd, wat eveneens de afname goed reproduceerbaar maakte. De

validiteit was daarentegen waarschijnlijk beperkt, op basis van de specifieke en het kleine aantal items van de vertaaltaken. Tevens werden middels de items in de perceptietaak mogelijke categorieën onderling niet geheel uitgesloten.

Data-verzameling

Het onderzoek vond plaats van februari tot en met mei 2017. Na afname van de pretest volgde de vocabulairetraining, oftewel de leerfase. Hierin leerden de kinderen middels een spel op een tablet Engelse woorden. Kinderen behorende de kindconditie voerden het spel geheel zelfstandig uit. Kinderen behorende tot de peer- en robotconditie werden daarentegen ondersteund door respectievelijk een peer of de robot. Zij volbrachten namelijk om beurten een spelscène, wat zowel interactie met de tablet, als samenwerking met de peer of robot mogelijk maakte. De pretest en leerfase duurden samen ongeveer 60 minuten en vonden plaats in een rustige kamer in de school waarin enkel de participant(en) en onderzoekers aanwezig waren.

Bij aanvang van het spel instrueerde de onderzoekster de kinderen dat de dierentuinmedewerker ziek was geworden en zij voor de dieren in de dierentuin mochten gaan zorgen. Tijdens het spel verzocht de onderzoekster het kind, de leeftijdgenoot of de robot verschillende taken uit te voeren door objecten en dieren, weergegeven op de tablet, te manipuleren (bijvoorbeeld ‘Plaats de ijsbeer *in front of* de boot’). De (ruimtelijke) targetwoorden, vier adjectieven en twee voorzetsels, werden door de onderzoekster zowel impliciet (‘Zet de *heavy* olifant achter het hek’) als expliciet (‘*Heavy* betekent ‘zwaar’ in het Engels’) geïnstrueerd. Elk woord kwam gedurende het spel tien keer aan bod.

Resultaten

Middels SPSS versie 21 zijn de hypothesen getoetst. Vanwege kleine groepsgroottes dienen onderstaande resultaten met voorzichtigheid geïnterpreteerd te worden. Voorafgaand aan alle analyses zijn de assumpties van normaliteit en gelijke varianties getoetst middels een histogram, een Q-Q plot en de Levene’s test. Op basis hiervan is aangenomen dat de afhankelijke variabelen voldoen aan de assumpties. De variabele Pretest vormt hierop een uitzondering, gezien de assumptie van normaliteit is geschonden, $F(2, 65) = 10.86, p < .001$.

Correlaties, zichtbaar in Tabel 2, tonen significante samenhang tussen scores op de vertaaltaken tijdens de post- en late test. Hieruit volgt dat een hoge score op één van de vertaaltaken gepaard gaat met een hoge score op iedere andere vertaaltaak. Een gepaarde t-test toont dat de kinderen in de gehele steekproef, en in elke conditie apart, gemiddeld significant hoger scores tijdens de post- en late test dan op de pretest ($p < .001$).

Tabel 2

Pearson Correlatiecoëfficiënten Vertaaltaken

	1	2	3	4
1 EN-NL Post	-			
2 NL-EN Post	.46**	-		
3 EN-NL Late	.46**	.40*	-	
4 NL-EN Late	.42**	.48**	.62**	-

Noot. * $p = .001$, tweezijdig. ** $p < .001$, tweezijdig. EN-NL Post/Late = Score op vertaaltak EN-NL tijdens de post-/late test, NL-EN Post/Late = Score op vertaaltak NL-EN tijdens de post-/late test.

Op de pretest behaalt de gehele steekproef, op twee kinderen met een score van 1 na, een score van 0 ($M_{\text{kind}} = .00$, $SD_{\text{kind}} = .00$; $M_{\text{peer}} = .09$, $SD_{\text{peer}} = .29$; $M_{\text{robot}} = .00$, $SD_{\text{robot}} = .00$). Hiertoe is aangenomen dat behaalde scores op de vertaaltaken tijdens de post- en late test gelijk zijn aan de leeropbrengst. Tevens is, op basis van de reeds benoemde schending van normaliteit, variabele Pretest niet meegenomen in aankomende analyses.

Conditie

Om te onderzoeken of het effect van (1) leren met een tablet en robot op de leeropbrengst groter is dan het effect van (2) leren met enkel een tablet en (3) leren met een tablet en leeftijdgenoot op de leeropbrengst, is een MANCOVA (F-toets) uitgevoerd. Er is gecontroleerd voor Leeftijd. Beschrijvende statistieken van de scores op de vertaaltaken zijn per conditie weergegeven in Tabel 3.

Tabel 3

Gemiddelden (M) en Standaardvariëties (SD) van Scores op de Vertaaltaken

	Kind		Peer		Robot	
	<i>N</i>	<i>M (SD)</i>	<i>N</i>	<i>M (SD)</i>	<i>N</i>	<i>M (SD)</i>
EN-NL Post	22	2.34 (1.10)	22	2.55 (1.16)	23	2.48 (1.13)
NL-EN Post	22	2.59 (1.05)	22	2.30 (1.03)	23	2.74 (1.17)
EN-NL Late	23	2.82 (1.39)	20	2.95 (0.99)	23	2.83 (1.32)
NL-EN Late	23	3.00 (0.96)	20	2.83 (1.09)	23	2.94 (1.11)

Noot. EN-NL Post/Late = Score op vertaaltak EN-NL tijdens de post-/late test, NL-EN post/late = Score op vertaaltak NL-EN tijdens de post-/late test.

De MANCOVA levert geen significant resultaat op, $F(8, 112) = 0.67$, $p = .72$, $\Lambda = .91$, $\eta_p^2 = .046$. Hieruit volgt geen significant verschil in leeropbrengst tussen de drie condities. Uit de resultaten blijkt tevens geen significant verband tussen leeftijd (covariaat) en de leeropbrengst, $F(4, 56) = 1.41$, $p = .24$, $\Lambda = .91$, $\eta_p^2 = .092$. Op basis van deze resultaten wordt de eerste hypothese verworpen betreffende zowel de posttest, als late test.

Vanwege de kleine steekproefgrootte en daardoor lage power zijn, ondanks het niet significant verschil, alsnog bijbehorende ANCOVA's bestudeerd (zie Tabel 4). Ook hierin blijkt geen sprake van significante en/of middelmatige of grote effecten van conditie op de leeropbrengst ($p > .05$). Covariaat Leeftijd vertoont daarentegen een significant effect op EN-NL Late, $F(1, 59) = 5.61$, $p = .021$, $\eta_p^2 = .087$.

Tabel 4

ANCOVA's voor Effect van Conditie op de Leeropbrengst

	<i>df</i>	<i>df_{error}</i>	<i>F</i>	η_p^2	<i>p</i>
EN-NL Post	2	59	0.24	.008	.79
NL-EN Post	2	59	0.57	.019	.57
EN-NL Late	2	59	0.57	.019	.57
NL-EN Late	2	59	0.34	.011	.72

Noot. EN-NL Post/Late = Score op vertaaltaak EN-NL tijdens de post-/late test, NL-EN Post/Late = Score op vertaaltaak NL-EN tijdens de post-/late test.

Sekse

Eenzelfde MANCOVA is uitgevoerd om te onderzoeken of binnen de *robotconditie* de leeropbrengst groter is voor meisjes dan voor jongens. Er is opnieuw gecontroleerd voor Leeftijd. Beschrijvende statistieken in Tabel 5 tonen gemiddelde scores op de vertaaltaken van zowel jongens, als meisjes in de *robotconditie*.

Tabel 5

Gemiddelde en Standaardafwijking van Scores op de Vertaaltaken

	N	EN-NL Post <i>M (SD)</i>	NL-EN Post <i>M (SD)</i>	EN-NL Late <i>M (SD)</i>	NL-EN Late <i>M (SD)</i>
Jongens ^a	12	2.04 (0.84)	2.42 (1.18)	2.58 (1.14)	2.67 (1.03)
Meisjes ^a	11	2.96 (1.25)	3.09 (1.09)	3.09 (1.50)	3.23 (1.17)

Noot. ^a = Kinderen in de robotconditie. EN-NL Post/Late = Score op vertaaltaak EN-NL tijdens de post-/late test, NL-EN Post/Late = Score op vertaaltaak NL-EN tijdens de post-/late test.

Het resultaat van de MANCOVA ten aanzien van het effect van Sekse op behaalde scores is niet significant, $F(4, 16) = 0.41, p = .80, \Lambda = .91, \eta_p^2 = .093$. Hieruit volgt geen significant verschil in leeropbrengst tussen jongens en meisjes binnen de robotconditie. Wel suggereert de effectgrootte een medium effect. Daarnaast blijkt uit deze resultaten geen significant, maar wel een groot, effect van covariaat Leeftijd op behaalde scores op vertaaltaken, $F(4, 16) = 0.79, p = .55, \Lambda = .84, \eta_p^2 = .17$. Op basis van deze resultaten wordt de tweede hypothese verworpen ten aanzien van zowel de post-, als de late test.

Opnieuw zijn alsnog bijbehorende ANCOVA's bestudeerd. Ook uit deze resultaten volgen geen significante effecten van Sekse op leeropbrengst, evenals van Leeftijd op leeropbrengst. De effectgroottes suggereren enkel een medium effect van Sekse op EN-NL Post, $F(1, 19) = 1.62, p = .22, \eta_p^2 = .079$. Wat betreft de covariaat tonen de resultaten dat mogelijk sprake is van een medium effect van Leeftijd op EN-NL Late, $F(1, 19) = 2.04, p = .17, \eta_p^2 = .097$, en een groot effect van Leeftijd op NL-EN Late, $F(1, 19) = 3.19, p = .09, \eta_p^2 = .14$.

Perceptie

Ten slotte was het streven om te onderzoeken of eventuele sekseverschillen werden verklaard door de perceptie die kinderen ten aanzien van de robot hadden. Hoewel de vorige hypothese is verworpen, is wegens een medium effect van Sekse op leeropbrengst alsnog de derde hypothese bestudeerd. Echter volgt, vanwege te kleine groepsgroottes, een (overwegend) kwalitatieve beschrijving.

Allereerst is een Fisher's Exact Test uitgevoerd om te onderzoeken of Sekse een significant effect heeft op de perceptie van kinderen ten aanzien van de robot. De beschrijvende statistieken betreffende Perceptie is voor jongens en meisjes weergegeven in Tabel 6.

Tabel 6

Frequentietabel Perceptie voor Jongens en Meisjes

	Perceptie robot			
	Ding	Vriendje/vriendinnetje	Meester/juf	Onduidelijke perceptie
Jongens				
<i>F</i>	3	6	3	0
%	25%	50%	25%	0%
Meisjes				
<i>F</i>	2	5	1	3
%	18%	46%	9%	27%

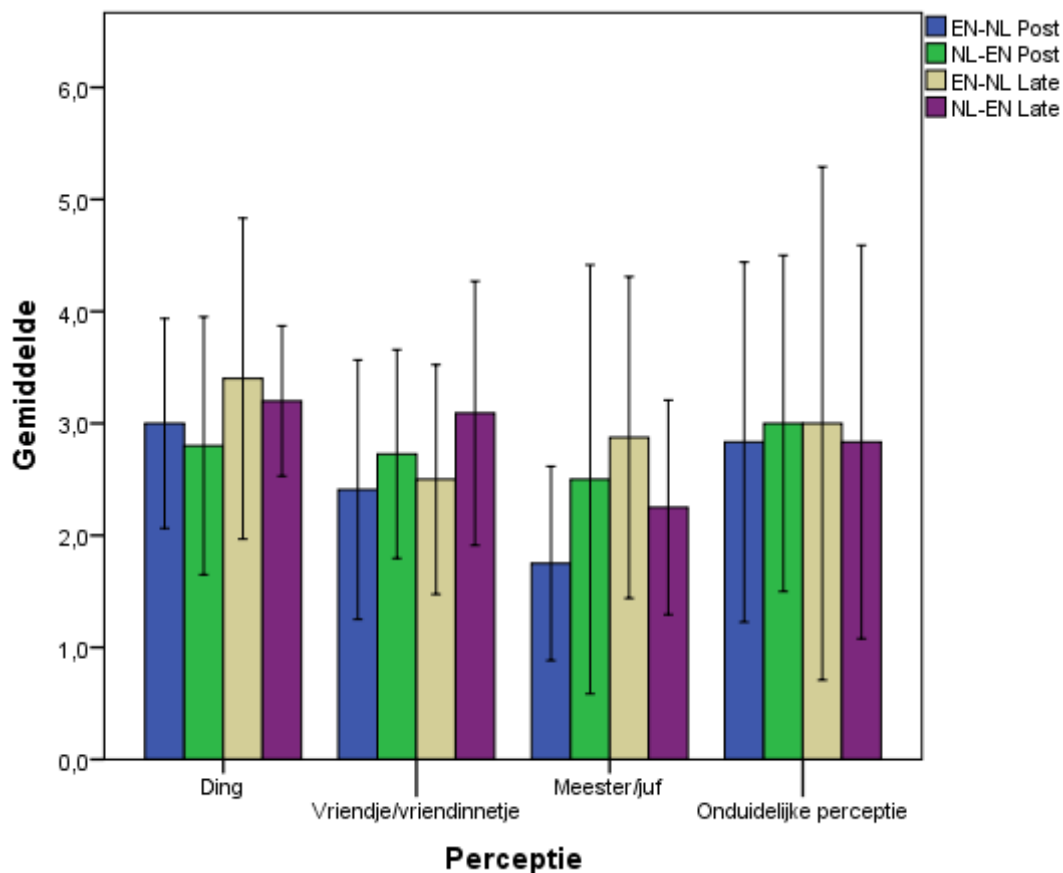
Noot. *f* = Frequentie, % = Percentage van alle jongens of meisjes.

De Fisher's Exact Test toont geen significant verband tussen Sekse en Perceptie ($p = .29$). Dit suggereert dat jongens en meisjes niet significant verschillen in hun perceptie ten aanzien van de robot. Echter blijkt bij meisjes bijvoorbeeld ook, in tegenstelling tot bij jongens, sprake van een onduidelijke perceptie.

Ten tweede volgt uit beschrijvende statistieken dat behaalde scores op de vertaaltaken verschillen per mogelijke perceptie (zie Figuur 1). Kinderen die de robot als ding zien, scoren bijvoorbeeld het hoogst op drie van de vier vertaaltaken. Kinderen die de robot als meester of juf zien, scoren daarentegen op drie van de vier vertaaltaken het laagst. Echter is een nadere analyse niet mogelijk en speelt leeftijd waarschijnlijk ook hier weer een rol. Deze resultaten zijn dan ook van onvoldoende betekenis en onvoldoende betrouwbaar om een uitspraak te doen betreffende de laatste hypothese.

Figuur 1

Gemiddelden en Standaardvariaties van Scores per Perceptie



Noot. Error Bar = +/- 1 SD

Conclusie

De onderzoeksvraag die in dit onderzoek centraal stond, was: ‘In hoeverre heeft een robot effect op het aanleren van productieve woordenschat in een tweede taal bij kleuters?’. Als onderdeel van het L2TOR project kunnen uitkomsten van huidig onderzoek bijdragen aan optimalisering van zowel Engelse taallessen aan Nederlandstalige kleuters, als de toepassing van sociale robots als leermaatje. Het is bijvoorbeeld één van de eerste onderzoeken die de effecten van een robot als leermaatje op Engels als tweede taal heeft bestudeerd middels een steekproef bestaande uit Nederlandstalige kinderen.

Op basis van de resultaten kan worden geconcludeerd dat het zelfstandig, maar ook samen met een leeftijdgenoot of robot doorlopen van een vocabulairetraining op een tablet de productieve, Engelse woordenschat van Nederlandstalige kleuters significant kan bevorderen. Dit geldt zowel op korte, als langere termijn. Echter, in tegenstelling tot de verwachting, tonen bevindingen dat leren in samenwerking met een robot niet effectiever is dan leren met een tablet of tablet en leeftijdgenoot.

Daarnaast werd specifiek met betrekking tot kinderen die met de robot hebben geleerd, verwacht dat dit bij meisjes een grotere leeropbrengst als gevolg zou hebben dan bij jongens. Meisjes scoren inderdaad hoger op vertaaltaken na afloop van de vocabulairetraining dan jongens en er is sprake van een medium effect. Echter suggereren niet significante bevindingen dat de productieve, Engelse woordenschat van jongens en meisjes op korte en lange termijn in gelijke mate wordt bevorderd middels het leren met een tablet en robot.

Over de mogelijke rol van perceptie betreffende het niet significant, maar medium effect van sekse op de leeropbrengst kan geen conclusie worden getrokken. Een betrouwbare analyse op basis van de huidige steekproef is namelijk niet mogelijk gebleken. Jongens en meisjes verschillen op basis van deze resultaten niet in hun perceptie ten aanzien van de robot. Wel suggereren resultaten dat de manier waarop kinderen de robot waarnemen (als ding, vriendje/vriendinnetje, meester/juf of een combinatie hiervan) bijdraagt aan de mate waarin hun woordenschat is bevorderd.

Eerder onderzoek

De Engelse woordenschat van kleuters kan dus significant worden bevorderd door te leren met een tablet, tablet en leeftijdgenoot en/of tablet en robot. Dit komt overeen met eerder onderzoek, zoals reeds in de inleiding besproken (Baloian et al., 2013; Bulotsky-Shearer et al., 2012; Justice et al., 2011; Lin et al., 2016; Mazzoni & Benvenuti, 2015; Tanaka & Matsuzoe, 2012; Westlund & Breazeal, 2015). Echter is ook gebleken dat een robot waarschijnlijk niet effectiever is dan de andere twee manieren van ondersteuning. Volgens Mazzoni en Benvenuti (2015) daarentegen, is samenwerking met een robot effectiever dan samenwerken met een leeftijdgenoot, vanwege onder andere de structuur die een robot kan bieden. Echter is dit onderzoek, net als huidig onderzoek, door een kleine steekproef niet geheel betrouwbaar. Daarnaast betreft huidig onderzoek enkel productieve woordenschat en hebben de leerinhoud en instructies mogelijk onvoldoende aangesloten bij de kleuters. Dit kan ertoe geleid hebben dat de robot inderdaad niet meer voordeel heeft gehad binnen de ondersteuning dan geen samenwerking of samenwerking met een leeftijdgenoot (Chang et al., 2010).

Wel lijken meisjes in de robotconditie hoger te scoren dan jongens, in overeenstemming met onderzoek van Kennedy en collega's (2015). Echter is dit verschil niet significant. Dat Schermerhorn en collega's (2008) wel een significant sekseverschil betreffende leeropbrengst constateren, is wellicht te verklaren door de kleine steekproef van huidig onderzoek. Het medium effect van sekse op de leeropbrengst wekt namelijk de indruk

dat er wellicht een verschil tussen jongens en meisjes is, mogelijk aantoonbaar middels een grotere steekproef.

Sterke punten

Ondanks dat onbekend is of de betrouwbaarheid en validiteit van de onderzoeksinstrumenten voldoende zijn, is een goede kwaliteit nagestreefd door onder andere de manier van instructie tussen de verschillende onderzoeksters goed af te stemmen. Daarnaast zijn scholen uit verschillende provincies in Nederland in de steekproef opgenomen en kinderen, zover als praktisch mogelijk, aselect ingedeeld in de condities. Dit maakt generaliseerbaarheid van bevindingen waarschijnlijk. Tevens is deze studie een aanvulling op de reeds aanwezige wetenschappelijke literatuur. Dit is bijvoorbeeld het eerste onderzoek dat deze drie mogelijke manieren van ondersteuning om kinderen een tweede taal te leren direct met elkaar vergelijkt. Ook hebben huidige resultaten betrekking op kleuters, iets wat in de inleiding beperkt aanwezig is gebleken (Ortega & Iberri-Shea, 2005). Daarnaast bestond de steekproef uit Nederlandse kinderen, waardoor resultaten waarschijnlijk beter toepasbaar zijn op de Nederlandse populatie.

Kanttekeningen

Deze steekproef is echter klein, wat uitgevoerde analyses minder betrouwbaar maakt. Daarnaast blijkt de mate waarin de woordenschat wordt bevorderd, beïnvloed te worden door de mate waarin vaardigheden van de peer aansluiten bij het niveau van het kind (Hegarty, 2004; Justice et al., 2011; Mashburn, Justice, Downer, & Pianta, 2009). Mogelijk is deze aansluiting beperkt geweest, doordat kinderen willekeurig over de condities zijn verdeeld. Daarnaast betreft huidig onderzoek slechts een éénmalige vocabulairetraining. Hierdoor zijn herhaling en mogelijkheden om een band op te bouwen met de robot slechts beperkt geweest. Echter blijken deze aspecten van belang binnen taalverwervingsprocessen door bij te dragen aan behaalde resultaten (Chang et al., 2010). Ook zijn slechts zes targetwoorden aangeboden en toont eerder onderzoek onder studenten dat het aanbieden van minder targetwoorden samenhangt met een lagere score op productieve vertaaltaken bij het aanleren van een tweede taal (Webb, 2008). Wellicht is dit ook het geval bij kleuters en maakt dit tevens meer spreiding in scores mogelijk, wat kan leiden tot andere bevindingen.

Bovendien is de motivatie en vaardigheid in de eerste taal van de kleuters niet in de studie meegenomen, terwijl ook deze factoren van belang blijken in het proces van tweede taalverwerving (Chang et al., 2010; Grøver, Lawrence, & Rydland, 2016; Masgoret & Gardner, 2003). Daarnaast zijn resultaten betreffende receptieve woordenschat niet meegenomen in huidig onderzoek. Wellicht geeft dit een beter beeld van het proces van

tweede taalverwerving bij kleuters. Doorgaans blijkt de verwerving van receptieve woordenschat in een tweede taal namelijk makkelijker dan productieve woordenschat. Dit verschil is tevens groter wanneer er weinig woorden worden geleerd (Mondria & Wiersma, 2004; Webb, 2008). Ten slotte was de robot nog niet in staat om emoties te uiten en om verbale en nonverbale signalen van de kinderen waar te nemen en hierop te reageren. Dit kan echter leiden tot een effectievere samenwerking met de robot en zo de taalverwerving bevorderen (Chang et al., 2010; Tanaka & Matsuzoe, 2012).

Aanbevelingen

Uit deze kanttekeningen volgen verschillende implicaties voor toekomstig onderzoek. Ten eerste dient men gebruik te maken van een grotere steekproef, zodat analyses en resultaten betrouwbaarder zijn. Wellicht leidt dit tot bevestiging van huidige bevindingen, of worden hypothesen alsnog aangenomen. In huidig onderzoek zijn namelijk matige en grote effecten gevonden, welke mogelijk in een grotere steekproef wel significant zijn.

Daarnaast zou toekomstig onderzoek een langere periode, of hogere frequentie, van vocabulairetraining kunnen nastreven. Dit geeft kinderen mogelijkheden tot herhaling en het opbouwen van een band met de robot. Deze robot dient bovendien verder ontwikkeld te worden, betreffende het uiten van emoties en waarnemen van/reageren op signalen van kinderen, om de interactie tussen de robot en kinderen te optimaliseren en effectiviteit te bevorderen. Ook zou het goed zijn om zowel productieve, als receptieve woordenschat in onderzoek te betrekken en meer targetwoorden aan te bieden. Ten slotte dient in toekomstig onderzoek de taalvaardigheden van de peer en robot te worden afgestemd op die van het kind.

Dit alles kan bijdragen aan kennis met betrekking tot het aanleren van een tweede taal aan kleuters. Uiteindelijk biedt dit mogelijkheden om onderwijs te optimaliseren. Bevindingen tot nu toe zijn wellicht niet overtuigend genoeg voor scholen om een robot aan te schaffen, zeker gezien de kosten en praktische tekortkomingen die er nog zijn, dus kan vervolgonderzoek scholen ondersteunen in hun keuzes (Kennedy et al., 2015).

Tot slot

Samenwerking met een robot kan dus zorgen voor een significante bevordering van de productieve woordenschat van kleuters in een tweede taal. Echter tonen bevindingen dat een robot relatief gezien geen voordeel heeft ten opzichte van de mate waarin leren met een tablet enerzijds en leren met een tablet en leeftijdgenoot anderzijds bijdraagt aan de leeropbrengst. Daarnaast is geconcludeerd dat, specifiek betreffende het leren met de robot, jongens en meisjes niet verschillen in leeropbrengst en perceptie ten aanzien van de robot. Reeds aanwezige wetenschappelijke literatuur is schaars en vertoont gebreken. Huidig onderzoek

vult enkele kennishiaten op, echter is sprake van een kleine steekproef waardoor resultaten met voorzichtigheid geïnterpreteerd dienen te worden. Door de hierboven benoemde implicaties in acht te nemen, wordt mogelijk de bijdrage aan deze literatuur verder bevorderd. Kortom, er is nog veel winst te behalen in het (onderzoek naar) ontwikkelen van een sociale robot en het inzicht in de toepassing van een robot in het onderwijs. Dit zal het onderwijs en vaardigheden van kleuters ten goede komen. Want zij, en robots, zijn de toekomst.

Literatuur

- Baloian, N., Pino, J. A., & Vargas, R. (2013). Tablet gestures as a motivating factor for learning. *Chilean Conference on Human - Computer Interaction, 15*, 98-103.
doi:10.1145/2535597.2535622
- Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner (Eds.), *Handbook of child development*, (pp. 793-828). Hoboken, NJ: Wiley. doi:10.1002/9780470147658.chpsy0114
- Bulotsky-Shearer, R. J., Manz, P. H., Mendez, J. L., McWayne, C. M., Sekino, Y., & Fantuzzo, J. W. (2012). Peer play interactions and readiness to learn: A protective influence for African American preschool children from low-income households. *Child Development Perspectives, 6*, 225-231. doi:10.1111/j.1750-8606.2011.00221.x
- Chang, C. W., Lee, J. H., Chao, P. Y., Wang, C. Y., & Chen, G. D. (2010). Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *Educational Technology & Society, 13*, 13-24. Verkregen op 14 juni, 2017, van <http://web.b.ebscohost.com.proxy.library.uu.nl/ehost/pdfviewer/pdfviewer?sid=32098374-bcc9-4030-81c0-665eed11afb5%40sessionmgr102&vid=1&hid=118>
- Chanprasert, C., & Han, H. (2013). Learning on the move: The use of mobile technologies for language skill development. *Executive Journal, 34*, 98-107. Verkregen op 29 oktober, 2016, van http://www.bu.ac.th/knowledgecenter/executive_journal/jan_june_14/pdf/aw10.pdf
- Chinnery, G. (2006). Emerging technologies going to the MALL: Mobile assisted language learning. *Language Learning & Technology, 10*, 9-16. Verkregen op 13 juni, 2017, van <http://lt.msu.edu/vol10num1/pdf/emerging.pdf>
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2012). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*, 70-105.
doi:10.1080/09588221.2012.700315
- Grøver, V., Lawrence, J., & Rydland, V. (2016). Bilingual preschool children's second-language vocabulary development: The role of first-language vocabulary skills and second-language talk input. *International Journal of Bilingualism, 1*, 1-17.
doi:10.1177/1367006916666389

- Haßler, B., Major, L., & Hennessy, S. (2016). Tablet use in schools: A critical review of the evidence for learning outcomes. *Journal of Computer Assisted Learning*, 32, 139-156. doi:10.1111/jcal.12123
- Hegarty, M. (2004). Dynamic visualizations and learning: Getting to difficult questions. *Learning and Instruction*, 14, 343-351. doi:10.1016/j.learninstruc.2004.06.007
- Justice, L. M., Petscher, Y., Schatschneider, C., & Mashburn, A. (2011). Peer effects in preschool classrooms: Is children's language growth associated with their classmates' skills?. *Child Development*, 82, 1768-1777. doi:10.1111/j.1467-8624.2011.01665.x
- Kennedy, J., Baxter, P., & Belpaeme, T. (2015). The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction* (pp. 67-74). Verkregen op 30 oktober, 2016, van <http://dl.acm.org.proxy.library.uu.nl/citation.cfm?id=2696457>
- Kose-Bagci, H., Ferrari, E., Dautenhahn, K., Syrdal, D. S., & Nehaniv, C. L. (2009). Effects of embodiment and gestures on social interaction in drumming games with a humanoid robot. *Advanced Robotics*, 23, 1951-1996. doi:10.1163/016918609X12518783330360
- Lin, T. J., Justice, L. M., Paul, N., & Mashburn, A. J. (2016). Peer interaction in rural preschool classrooms: Contributions of children's learning-related behaviors, language and literacy skills, and problem behaviors. *Early Childhood Research Quarterly*, 37, 106-117. doi:10.1016/j.ecresq.2016.04.001
- Masgoret, A. M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, 53, 123-163. doi:10.1111/1467-9922.00212
- Mashburn, A. J., Justice, L. M., Downer, J. T., & Pianta, R. C. (2009). Peer effects on children's language achievement during pre-kindergarten. *Child Development*, 80, 686-702. doi:10.1111/j.1467-8624.2009.01291.x
- Mazzoni, E., & Benvenuti, M. (2015). A robot-partner for preschool children learning English using socio-cognitive conflict. *Educational Technology & Society*, 18, 474-485. Verkregen op 2 februari, 2016, van http://www.ifets.info.proxy.library.uu.nl/journals/18_4/36.pdf
- Mondria, J. A., & Wiersma, B. (2004). Receptief, productief, en receptief & productief woorden leren: Wat maakt het uit? *Levende Talen Tijdschrift*, 5, 29-37. Verkregen op

- 26 november, 2017, van <http://www.lt-tijdschriften.nl/ojs/index.php/ltt/article/view/501/494>
- Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2008). Prediction of human behavior in human-robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE Transactions on Robotics*, *24*, 442-451. doi:10.1109/TRO.2007.914004
- Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, *25*, 26-45. doi:10.1017/S0267190505000024
- Schermerhorn, P., Scheutz, M., & Crowell, C. R. (2008). Robot social presence and gender: Do females view robots differently than males? In *Proceedings of the 3rd ACM/IEEE international conference on human robot interaction* (pp. 263-270). Verkregen op 16 november, 2016 van <http://dl.acm.org.proxy.library.uu.nl/citation.cfm?id=1349857>
- Stockwell, G. (2007). A review of technology choice for teaching language skills and areas in the CALL literature. *European Association for Computer-Assisted Language Learning (ReCALL)*, *19*, 105-120. doi:10.1017/S0958344007000225
- Tanaka, F., & Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction*, *1*, 78-95. doi:10.5898/JHRI.1.1.Tanaka
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, *30*, 79-95. doi:10/10170S02722263108080042
- Westlund, J. K., & Breazeal, C. (2015). The interplay of robot language level with children's language learning during storytelling. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction extended abstracts* (pp. 65-66). Verkregen op 5 november, 2016, van <http://dl.acm.org.proxy.library.uu.nl/citation.cfm?id=2701989>

Bijlage 1
Vertaaltaak Engels-Nederlands

Instructie PRETEST

Zo meteen gaan we spel over een diertuin spelen, en daarbij ga je wat Engelse woordjes leren. Maar misschien ken je ze al, en dan kan je ze natuurlijk niet meer leren! Daarom gaan we nu even kijken of je de woorden al kent. Ik hoop dat je ze nog niet weet, want dan kan ik je ze zo meteen leren! Dus denk maar goed na of je ze kent, maar het is dus helemaal niet erg als je ze niet kent.

Instructie POSTTEST

In de diertuin werd natuurlijk Engels gepraat en jij hebt misschien ook wat Engelse woordjes geleerd. Het waren een hoop woorden. Laten we eens kijken welke je nog weet! Ik ga zo meteen steeds een woord in het Engels zeggen. Jij mag dan het Nederlandse woord zeggen.

Instructie LATE POSTTEST

Vorige week werd er natuurlijk in de diertuin Engels gepraat, en jij hebt misschien ook wat Engelse woordjes geleerd! Laten we eens gaan kijken welke je nog weet! Ik ga zo meteen steeds een woord in het Engels zeggen. Jij mag dan het Nederlandse woord zeggen.

Testvragen – lijst 1

	Goed	Goed na prompt	Fout/geen antwoord	Score per item
Wat is “light”?				
Wat is “full”?				
Wat is “in front of”?				
Wat is “heavy”?				
Wat is “behind”?				
Wat is “empty”?				
<i>Totaalscore</i>				

Bijlage 2

Vertaaltaak Nederlands-Engels

Instructie POSTTEST

Je deed net al heel goed je best toen ik de Engelse woordjes zei en jij de Nederlandse woordjes moest zeggen. Nu gaan we het andersom doen! Ik ga zo meteen steeds een woord in het Nederlands zeggen. Jij mag dan het Engelse woord zeggen. We gaan gewoon kijken welke woordjes je nog weet!

Instructie LATE POSTTEST

Je deed net al heel goed je best toen ik de Engelse woordjes zei en jij de Nederlandse woordjes moest zeggen. Nu gaan we het andersom doen! Ik ga zo meteen steeds een woord in het Nederlands zeggen. Jij mag dan het Engelse woord zeggen. We gaan gewoon kijken welke woordjes je nog weet!

Testvragen – lijst 1

	Goed	Goed na prompt	Fout/geen antwoord	Score per item
Wat is "licht"?				
Wat is "vol"?				
Wat is "voor"?				
Wat is "zwaar"?				
Wat is "achter"?				
Wat is "leeg"?				
<i>Totaalscore</i>				