



Universiteit Utrecht

Master thesis

Universiteit Utrecht

Master programma Clinical Child, Family and Education Studies

Vershil in het Volgen van Wijsgedrag tussen Mens en Robot bij Kleuters bij een Referentieel
Conflict en de Rol van Woordenschat

Naam: B.M. (Bente Marije) Homan (5958679)

Eerste beoordelaar: Rianne van den Berghe

Tweede beoordelaar: Elma Blom

Datum: 20-06-2018

Voorwoord

Deze thesis is een onderdeel van de masteropleiding Clinical Child, Family and Education Studies (Orthopedagogiek) en bestudeert het verschil in het volgen van wijsgedrag tussen mens en robot bij kleuters bij een referentieel conflict en de rol van woordenschat hierin. Dit onderzoek is onderdeel van het internationale L2TOR-project. De afgelopen maanden heb ik als zeer intensief en stressvol ervaren, maar dit onderzoek is naar mijn idee voor mij vooral erg leerzaam geweest en het verzamelen van de data was ontzettend leuk om te doen. Bij deze wil ik Rianne van den Berghe bedanken voor haar ondersteuning en zeer prettige begeleiding in het gehele proces van mijn masterthesis. Daarnaast wil ik graag Josje Verhagen bedanken, aangezien zij mij ook van nuttige tips en feedback heeft voorzien. Verder bedank ik mijn medestudenten die zich voor de volle honderd procent in hebben gezet voor onze dataverzameling en waar ik altijd met al mijn vragen bij terecht kon. Ten slotte wil ik de kinderen die deel hebben genomen aan het onderzoek bedanken voor hun inzet en de scholen die hebben deelgenomen voor hun enthousiasme en gastvrijheid

Ik wens u veel leesplezier.

Utrecht, 20 juni 2018

Bente Homan

Abstract

Children can pay attention to objects in two different ways, namely by pointing at an object or by naming an object by words. In this current Masterthesis is studied if there is a difference in following pointing between a human and a robot with referential cues (pointing versus words) with conflicting information by children in the age of 4-6 years and what role the vocabulary of the children plays in it. The current study is a replication of an earlier study of Grassmann and Tomasello (2010), which revealed that children follow pointing over words by a conflict-task. The data was collected by executing a conflict-task and was taken by 63 children in the age of 49.51 to 74.58 months ($M = 61.92$, $SD = 6.13$, 38.10% girls). During the conflict-task, a distinction was made between a 'familiar label'-condition and a 'novel label'-condition. Results showed that there were no significant differences for agent (human - robot), for the conditions (familiar label – novel label) and for the combination of agent and condition to the extent that children following pointing between the human and robot. In addition, there was no effect and no significant difference found between children with a large and small vocabulary. Also, there were no interaction effects. The theoretical and practical implications of these findings will be discussed.

Keywords: robot; conflict-task; referential cues; preschoolers; vocabulary

Samenvatting

Kinderen kunnen hun aandacht richten op voorwerpen op twee verschillende manieren, namelijk wanneer er naar een voorwerp gewezen wordt of wanneer een voorwerp met woorden benoemd wordt. In deze thesis wordt bestudeerd of er een verschil is in het opvolgen van wijsgedrag tussen mens en robot bij kleuters van vier tot zes jaar bij een referentieel conflict (wijzen versus benoemen) en welke rol woordenschat hierin speelt. De huidige studie is een replicatie van een eerder experiment van Grassmann en Tomasello (2010), waarbij geconcludeerd werd dat kinderen eerder wijzen volgden dan benoemen. De data werd verzameld door een conflicttaak uit te voeren bij 63 kinderen in de leeftijd van 49.51 tot 74.58 maanden ($M = 61.92$, $SD = 6.13$, 38.10% meisjes). Tijdens de conflicttaak werd er onderscheid gemaakt tussen een ‘bekend-label’-conditie en een ‘onbekend-label’-conditie. De resultaten toonden aan dat er geen significante verschillen waren voor agent (mens - robot), voor de verschillende condities (bekend-label - onbekend-label) en voor de combinatie van agent en conditie in de mate waarin kinderen het wijsgedrag van mens en robot volgden. Bovendien was er geen effect en geen significant verschil tussen kinderen met een grote en kleine woordenschat. Er waren geen interactie-effecten. De theoretische en praktische implicaties van deze bevindingen zullen worden besproken.

Trefwoorden: robot, conflicttaak, referentiële cues, kleuters, woordenschat

Volgen van Wijsgedrag tussen Mens en Robot bij Kleuters tijdens een Referentieel Conflict en de Rol van Woordenschat

Als volwassenen kinderen woorden willen leren, is het nodig om de aandacht van kinderen te trekken. Deze aandacht proberen volwassenen op twee manieren te verkrijgen: door het gebruik van sociale *cues* (gebaren via wijzen of kijken) en door het gebruik van verbale *cues* (benoemen van woorden). Kinderen ervaren deze gebaren en taal vaak samen (Masur, 1997) en wanneer wijzen in combinatie met een woord wordt gebruikt, verwachten kinderen dat ze naar hetzelfde verwijzen (Gliga & Csibra, 2009). Op deze manier kunnen kinderen een koppeling maken tussen een woord en een voorwerp (Grassmann & Tomasello, 2010).

In de literatuur zijn er voor deze twee verschillende manieren (het gebruik van sociale- en verbale *cues*) twee stromingen te omschrijven. Het uitgangspunt van de eerste stroming is het belang van lexicale principes, waarbij het proces van *mutual exclusivity* een rol kan spelen (Markman & Wachtel, 1988). Bij de tweede stroming zijn vooral de sociaal-pragmatische *cues*, bijvoorbeeld een wijzend gebaar, belangrijk (Tomasello, 2000, 2001, 2008).

Bij *mutual exclusivity* maken kinderen gebruik van hun bestaande woordkennis. Kinderen hebben twee opties wanneer zij twee voorwerpen te zien krijgen en hierbij een onbekend woord te horen krijgen. De eerste optie is dat het kind het bekende voorwerp een nieuwe naam geeft en de tweede optie is dat het kind het onbekende voorwerp koppelt aan het onbekende woord. Kinderen hebben eerder de voorkeur om het onbekende woord aan het onbekende voorwerp koppelen, aangezien zij het bekende voorwerp al gekoppeld hebben aan een woord (Hansen & Markman, 2009; Jaswal & Hansen, 2006; Markman, 1989, 1992; Markman, Wasow, & Hansen, 2003; Markman & Wachtel, 1988; Merriman & Bowman, 1989). Wanneer kinderen moeten kiezen tussen twee voorwerpen kan het benoemen van woorden een belangrijke aanwijzing zijn voor kinderen. Wanneer er uit wordt gegaan van het principe van *mutual exclusivity* zou verwacht worden dat de grootte van de receptieve woordenschat van een kind van invloed is op de reactie op conflicterende verbale en non-verbale informatie. Des te meer woorden de kinderen kennen des te krachtiger zullen zij het principe van *mutual exclusivity* toepassen, omdat het voorwerp of het genoemde woord voor hen dan eerder bekend is. Als kinderen een grotere woordenschat hebben, kunnen ze hun kennis van bekende woorden gebruiken om de betekenis van nieuwe woorden te achterhalen. Eerdere studies hebben een relatie tussen *mutual exclusivity* en woordenschat laten zien: hoe groter de woordenschat, hoe vaker kinderen *mutual exclusivity* toepasten (Houston Price,

Valoghris & Raviglione, 2010; Bion, Borovsky & Fernald, 2013; Verhagen, Grassmann & Küntay, 2016).

Een interessante vraag hierbij is, wat kinderen denken en doen wanneer een gesproken woord en wijzend gebaar conflicteren met elkaar, dat wil zeggen dat zij niet naar hetzelfde verwijzen. In verschillende onderzoeken (Jaswal & Hansen, 2006; Grassmann & Tomasello, 2010) werd onderzocht of bij kinderen de pragmatische informatie of *mutual exclusivity* belangrijker was bij woordleren. Dit werd onderzocht aan de hand van een conflicttaak. De kinderen kregen twee verschillende voorwerpen te zien: een bekend- en onbekend voorwerp. Terwijl de testleider beide voorwerpen aanbood, zei de testleider “Can you give me the blicket?”, waarbij de testleider ondertussen juist naar het andere, bekende, voorwerp wees. Uit onderzoek van Jaswal en Hansen (2006) kwam naar voren dat kinderen voornamelijk het benoemde voorwerp selecteerden. Dit betekent dat de kinderen dachten dat het onbekende woord bedoeld was voor het onbekende voorwerp. Uit dit onderzoek kan geconcludeerd worden dat de kinderen verwachten dat, zelfs wanneer de testleider conflicterende pragmatische aanwijzingen gebruikt (wijzen), woorden *mutually exclusive* zijn. Grassmann en Tomasello (2010) maakten in hun experiment gebruik van ostensief wijzen, waarbij er duidelijk gewezen én gekeken wordt naar het voorwerp, waardoor er extra nadruk op het sociale aspect werd gelegd. Hierbij wees de testleider naar het bekende voorwerp en benoemde de testleider het onbekende woord. Een voorbeeld is de uitgesproken zin (met het onbekende woord) “Give me the modi” waarbij de testleider wees en keek naar een auto (het bekende voorwerp). Grassmann en Tomasello (2010) deden nog een tweede experiment waarbij juist het bekende woord werd genoemd, terwijl er gewezen werd naar het onbekende voorwerp. Een voorbeeld is de uitgesproken zin met het bekende woord “Give me the car”, waarbij de testleider juist wees naar het onbekende voorwerp.

De resultaten van het beide experimenten van Grassmann en Tomasello (2010) waren in contrast met de resultaten uit het onderzoek van Jaswal en Hansen (2006). Grassmann en Tomasello (2010) toonden in beide experimenten aan dat kinderen eerder het wijzen volgden dan het benoemen bij de conflicttaak. Een mogelijke verklaring hiervoor is dat er in het onderzoek van Grassmann en Tomasello (2010) gebruik werd gemaakt van ostensief wijzen en bij Jaswal en Hansen (2006) van niet-ostensief wijzen. Enfield, Kita en de Ruiters (2007) stellen dat niet-ostensief wijzen aangeeft dat de verbale informatie primair is en benadrukt dat de extra informatie niet nodig is waardoor kinderen dit gebaar eerder zullen negeren. Ostensief wijzen geeft echter aan dat het voorwerp waarnaar gewezen wordt primaire informatie bevat die aanduidt dat het aangewezen voorwerp van groot belang is. Wanneer er

sprake was van een conflict met een bekend woord kozen de kinderen het voorwerp dat werd aangewezen minder vaak (Grassmann & Tomasello, 2010), dan wanneer er sprake was van een conflict met een onbekend woord (Jaswal & Hansen, 2006).

Robotleren

Ook op scholen wordt er veelvuldig aandacht besteed aan het leren van nieuwe woorden waarbij het gebruik van technische hulpmiddelen, zoals computers, tablets en digiborden, populair is (Stockwell, 2007; Baloain, Pino, & Vargas, 2013; Golonka, Bowles, Frank, Richardson, & Freynik, 2012; Haßler, Major, & Hennessy, 2016). Een nieuwe, opkomende vorm van technologie die gebruikt wordt in het onderwijs is het gebruik van sociale robots. In het onderzoek van Bartneck en Forlizzi (2004) worden sociale robots als volgt gedefinieerd: ‘Een sociale robot is een autonome of semiautonome robot die interacteert met mensen door de gedragsnormen te volgen’. In verschillende onderzoeken is gekeken in hoeverre sociale robots gebruikt kunnen worden voor leerdoeleinden, zoals voor het gebruik als tutor (Kennedy, Baxter, Senft & Belpaeme, 2015) en voor het onderwijzen of aanleren van een (tweede) taal (Alemi, Meghdari, & Ghazisaedy, 2014; Kennedy, Baxter, Senft, & Belpaeme, 2016). Uit verschillende studies blijkt dat kinderen ook woorden kunnen leren wanneer zij met een (sociale) robot spelen, die zich gedraagt als een leeftijdsgenoot (Mazzoni & Benvenuti, 2015; Tanaka & Matsuzoe, 2012). Het is daarom van belang om te weten of dergelijke principes, het gebruik van lexicale principes (mutual exclusivity) en het gebruik van sociaal-pragmatische *cues* (wijzen) ook voor robots gelden, zodat we robots zo effectief mogelijk kunnen inzetten in het onderwijs.

Huidig onderzoek

Het huidige onderzoek is onderdeel van een internationaal L2TOR project (<http://www.l2tor.eu>). In dit onderzoek werd gebruikt gemaakt van een NAO-robot. Een NAO-robot is een autonome, programmeerbare mensachtige robot. De NAO-robot die gebruikt werd in dit experiment is een sociale robot, aangezien het sociale gedragingen kan vertonen, zoals praten en bewegen. De uitkomsten van het huidige onderzoek kunnen de kennis op het gebied van robotleren vergroten, omdat het aandacht besteedt aan de mate waarin kinderen non-verbale *cues* van de robot opvolgen, zoals kijken en wijzen, bij het aanbieden van de conflicttaak. De mate waarin kleuters vertrouwen op non-verbale *cues* (wijzen) wordt al benadrukt in de studie van Grassmann en Tomasello (2015) en in de studie van Verhagen, Grassmann en Küntay (2016). Daarnaast, gezien robots al recentelijk worden

ingezet bij (taal)leren (Alemi, Meghdari, & Ghazisaedy, 2014; Kennedy, Baxter, Senft, & Belpaeme, 2016), is het interessant en nuttig om te kijken in hoeverre kinderen ook vertrouwen op non-verbale *cues* van de robot, zoals wijzen, zodat men beter weet hoe robots effectief ingezet kunnen worden in het onderwijs.

De onderzoeksvraag van het huidige onderzoek is als volgt: ‘Is er een verschil in de mate waarin kleuters het wijsgedrag volgen tussen mens en robot bij een referentieel conflict en verschilt dit tussen kleuters met een grote en kleine woordenschat?’ Het huidige onderzoek is een nabootsing van de studie van Grassmann en Tomasello (2010). Uit eerder onderzoek van Jaswal en Hansen (2006) bleek dat kinderen gebruik maakten van *mutual exclusivity*. Daarentegen bleek uit onderzoek van Grassmann en Tomasello (2010) dat sociale pragmatische *cues*, zoals wijzen, juist een sterke aanwijzing waren voor kinderen. Dit gold zowel voor kinderen die net een taal leerden als voor kinderen die al een ver gevorderde taalontwikkeling hadden. Kinderen met een grote woordenschat maakten vooral gebruik van het samenbrengen van het wijzen en benoemen van een nieuw woord. Wanneer er een nieuw woord genoemd werd, maar er werd gewezen naar een bekend voorwerp, dachten de kinderen mogelijk dat de testleider een nieuw aspect van het bekende voorwerp bedoelde (Grassmann & Tomasello, 2010). Daarnaast was een andere mogelijkheid dat kinderen met een grote woordenschat meer gericht waren de pragmatische informatie, dus op het wijsgedrag van de testleider.

De eerste hypothese bij deze onderzoeksvraag is dat verwacht wordt dat kleuters eerder het wijsgedrag van de robot zullen volgen dan van de mens. In het onderzoek van Lin, Liu, Chang en Yeh, (2009) is geconstateerd dat kinderen hogere verwachtingen hebben van robots ten opzichte van de mens tijdens het leren. Zo verwachten kinderen dat de robot een hulpje kan zijn en hen kan helpen met het maken van huiswerk, maar ook dat de robot een vriend kan worden. Door deze hoge verwachtingen, die een robot echter niet waar kan maken, zullen kinderen waarschijnlijk ook eerder afgaan op de non-verbale *cues* van de robot. De tweede hypothese bij deze onderzoeksvraag is dat verwacht wordt dat kleuters met een grote woordenschat meer gebruik zullen maken van sociaal-pragmatische informatie, zoals wijzen.

Methode

Participanten

De steekproef bestond uit 63 kinderen, namelijk 39 jongens (61.90%) en 24 meisjes (38.10%). Zij waren in de leeftijd van 49.51 tot en met 74.58 maanden ($M = 61.92$ $SD = 6.13$). De kinderen zaten in groep 1 of 2 van een reguliere basisschool en hadden Nederlands

als moedertaal. Tweetaligheid gold als exclusiecriteria. De huidige steekproef ($N = 63$) is een subgroep van een grotere steekproef ($N = 119$) van kinderen die hebben deelgenomen aan het onderzoek. Echter kon van een aantal kinderen geen informatie meegenomen worden over een bepaalde conditie, ofwel omdat ze ziek waren ofwel omdat alle trials mislukt waren. In mensconditie 1 zijn dit vier kinderen, in robotconditie 1 negen kinderen, in mensconditie 2 vijf kinderen en in robotconditie 2 acht kinderen. Daarnaast is er een aantal kinderen ($N = 1$) uit de steekproef gehaald, omdat zij meertalig opgevoed zijn. Het onderzoek is uiteindelijk met $N = 59$ uitgevoerd in mensconditie 1, met $N = 54$ in robotconditie 1, met $N = 58$ in mensconditie 2 en met $N = 55$ in robotconditie 2. Sommige participanten ($N = 7$) hebben alleen aan de menssessie deelgenomen en een aantal ($N = 4$) hebben alleen aan de robotsessie meegedaan. De meeste participanten hebben aan beide sessies deelgenomen ($N = 52$).

De werving van kinderen voor het onderzoek heeft plaatsgevonden op verschillende reguliere basisscholen in bekende kringen van de testleiders. Vervolgens zijn er in verschillende kleutergroepen wervingsbrieven rondgestuurd naar ouders om toestemming te vragen om hun kind te laten deelnemen aan het onderzoek. De eisen hiervoor waren; het kind moest vier of vijf jaar oud zijn en uitsluitend Nederlands als moedertaal hebben.

Onderzoeksinstrumenten

Conflicttaak. In dit experiment, gebaseerd op een eerder experiment van Grassmann en Tomasello (2010), is gekeken hoe kinderen handelen tijdens een conflict tussen referentiële cues (wijzen versus benoemen). De taak werd afgenomen door getrainde studenten en opgenomen met een camera.

Er waren twee condities, van beide vier trials. Tijdens elke trial presenteerde de testleider twee objecten aan het kind op een tablet (een onbekend en een bekend voorwerp), waarbij de testleider zei: 'Kijk! Nu spelen we met de [label A]. Op dit moment wees de testleider naar [label B] en zei vervolgens: 'Druk op de [label A]'. De testleider bleef hierbij wijzen naar het voorwerp [label B] en keek om en om naar het voorwerp en naar het kind. Label A en B staan hierbij voor de twee voorwerpen. Het kind had hierbij de taak om één van de twee voorwerpen te kiezen en hierop te klikken op de tablet. De tablets die gebruikt werden in deze huidige studie vormen een belangrijke wijziging ten opzichte van het oorspronkelijke experiment (Grassmann & Tomasello, 2010) waarbij 'echte' voorwerpen gebruikt werden. De reden waarom deze wijziging is aangebracht is, omdat er in dit

experiment ook getest werd met een robot die geen voorwerpen kan aanpakken en verwisselen.

Conditie 1. Dit is de ‘bekend-label’-‘onbekend-voorwerp’ conditie. Een voorbeeld van de instructie in conditie 1 is: ‘Kijk! Nu spelen we met de auto, druk op de auto’. De testleider wees hierbij naar het onbekende voorwerp dat op de andere tablet te zien was, bijvoorbeeld een tuinslangstuk. Conditie 1 bestond uit vier trials en de bekende labels die werden gebruikt waren: auto, pen, kam en schoen.

Conditie 2. Dit is de ‘onbekend-label’-‘bekend voorwerp’ conditie. Een voorbeeld van de instructie in conditie 2 is: ‘Kijk! Nu spelen we met de modi, druk op de modi’. De testleider wees hierbij naar het bekende voorwerp dat op de andere tablet te zien was, bijvoorbeeld een auto. Conditie 2 bestond uit vier trials en de onbekende labels die gebruikt werden waren: modi, dofoe, toma en bafo. De onbekende voorwerpen die in beide condities te zien waren op één van de tablets waren sluitclip, tuinslangstuk, bouw materiaal en naamlabel.



Figuur 1. Gegeven aanwijzingen ‘bekend label’-conditie en ‘onbekend label’-conditie tijdens de conflict-taak. Aangepast figuur van “Young children follow pointing over words in interpreting acts of reference,” door Grassmann, S., & Tomasello, M., 2010, *Developmental Science*, 13, p.257.

In de afnames van de verschillende proefpersonen veranderden er een aantal zaken: de wijsrichting per trial, de combinatie van objecten en presentatievolgorde van de

objectcombinaties (trials). In de ene trial werd er dus naar rechts gewezen en in de andere trial naar links en werd er altijd gewezen naar het object wat niet benoemd was door de testleider. Daarnaast werd bij het ene kind gestart met de ‘bekend-label’-conditie en bij het andere kind werd gestart met de ‘onbekend-label’-conditie. Allerlaatst zorgden de acht verschillende lijsten (a t/m h) ervoor dat de presentatievolgorde van de trials per kind steeds verschilden.

Receptieve woordenschat. Om receptieve woordenschat de kinderen te meten, is de Nederlandse versie van de Peabody Picture Vocabulary Test-III-NL (PPVT-III-NL; Dunn, Dunn, & Schlichting, 2005) gebruikt. In de PPVT kozen kinderen één afbeelding uit vier afbeeldingen die past bij een mondeling aangeboden woord. Voorafgaand aan de test bepaalde de testleider steeds aan de hand van de leeftijd van het kind bij welke set er ingestapt werd, dit werd de instapset genoemd. De test werd afgebroken wanneer het kind negen fouten of meer maakte binnen een set. Deze set werd de afbreekset genoemd. Na het testen van het kind werd de ruwe score voor ieder kind berekend, die het woordenschatniveau van het kind weergeeft. De ruwe score werd berekend door het aantal gemaakt fouten af te trekken van de score van de afbreekset.

Overige onderzoeksinstrumenten. De Kiraly (2009) en de perceptievragenlijst (gebaseerd op Jipson & Gelman, 2007) maakten ook deel uit van de testbatterij. De Kiraly (2009) is afgenomen om de mate van volgzzaamheid van de kinderen te meten en de perceptievragenlijst (gebaseerd op Jipson & Gelman, 2007) is afgenomen om te meten hoe de kinderen denken over de robot. Echter worden deze onderzoeksinstrumenten niet verder besproken binnen dit onderzoek, omdat ze geen deel uitmaken van de data-analyse.

Procedure

De kinderen werden allemaal individueel getest in een aparte ruimte buiten de klas. Tijdens het onderzoek stond er een camera in de ruimte die de verschillende testen gefilmd heeft. Binnen de mensconditie en de robotconditie werd twee keer (‘bekend-label’-conditie en ‘onbekend-label’-conditie) de conflicttaak afgenomen ieder bestaande uit vier trials. Daarnaast werd tijdens de mensconditie de woordenschattest (PPVT-III-NL) tussen de twee delen van de conflicttaak afgenomen. Bij de robotconditie zijn de overige twee testen, de Kiraly en de perceptievragenlijst, aan bod gekomen. Aan het einde van de testafname kregen de kinderen allemaal een cadeautje of een sticker.

Data analyse

Conflicttaak. De gemiddelde scores op de referentiële conflicttaak werden als volgt berekend: aan ieder item werd een score toegekend van 0 (label gevolgd) of 1 (wijzen gevolgd). Vervolgens werd er voor ieder kind een gemiddelde over de geslaagde items (variërend tussen 1 en 4) per trial berekend. De gemiddelde scores die hieruit kwamen varieerden ook tussen de 0 en de 1. De afhankelijke variabele in het onderzoek was de mate waarin de kinderen het wijsgedrag van mens en robot volgen. De onafhankelijke variabele was de agent (mens of robot). Om te kijken of kinderen eerder kozen voor het wijzen of het benoemen werd er gebruik gemaakt van een one-sample t-test tegen kansniveau. Om te kijken of de scores tussen de afhankelijke en onafhankelijke variabele significant verschilden werd er gebruik gemaakt van een repeated measures anova met de gemiddelde score op het volgen van wijzen in conditie 1 en 2 als within subject variabele. De betrouwbaarheid van de schaalscores van de conflicttaak werd bepaald door de betrouwbaarheid van het coderen te meten. Een onafhankelijke persoon zal een deel van de data (10-20%) coderen en aan de hand hiervan zal de interrater reliability (test-hertest betrouwbaarheid) berekend worden. De Shapiro-Wilk-test liet een significant verschil zien ($p = .000$ dus $p < .05$), wat aangaf dat de data niet normaal verdeeld is. Toch werd er doorgedaan met de analyses, omdat de toetsen (one-sample t-test en repeated measures anova) redelijk robuust zijn met betrekking tot het schenden van normaliteit (Allen & Bennett, 2012).

PPVT. De moderator in het onderzoek is het woordenschatniveau van het kind. Er werd onderzocht of er een significant verschil is in het opvolgen van wijsgedrag tussen kleuters met een grote of kleine woordenschat. Alle kinderen werden op basis van hun ruwe scores op de PPVT-III-NL verdeeld aan de hand van een ‘median-split’ in een groep kinderen met een relatief grote woordenschat en een groep kinderen met een relatief kleine woordenschat. Op basis van de median split ($M = 80$) vallen de kinderen met een ruwe score op de PPVT van ≥ 80 onder de groep kinderen met een ‘grote woordenschat’ en de kinderen met een ruwe score op de PPVT van < 80 onder de groep kinderen met een ‘kleine woordenschat’. Vervolgens werd er een repeated measures anova toegepast waarbij de within subject variabele ‘conditie’ (conditie 1 en 2 mens en robot) werd toegevoegd en ‘woordenschatniveau’ als between subjects variabele. De betrouwbaarheid van de schaalscores van de PPVT is bepaald door middel van de Lamda-2’s coëfficiënt die varieert tussen .89 en .97 ($>.70$) en kan als betrouwbaar gezien worden (Dunn & Dunn, 2005). De interrater reliability (test-hertest betrouwbaarheid) van de PPVT bedroeg 0.941. Er is voldaan

aan de assumpties van normaliteit. De Sharipo-Wilk-test liet geen significant verschil zien ($p = .680$ dus $p > .05$), wat aangeeft dat de data normaal verdeeld is.

Resultaten

Volgen van Wijsgedrag

In tabel 1 staan de percentages van de opvolging van het wijsgedrag van de geslaagde trials van conditie 1 en 2 van de conflicttaak voor beide agents (mens en robot).

Conditie 1 en 2. Conditie 1 bevat de trials waarbij de combinaties onbekend label-bekend voorwerp voorkwamen. Conditie 2 bevat de trials waarbij de combinaties bekend label-onbekend voorwerp voorkwamen.

Tabel 1

Percentages Opvolging Wijzen van Geslaagde Trials Conditie 1 en 2

	Conditie 1				Conditie 2			
	Mens		Robot		Mens		Robot	
Respons typen	N	Cumulatieve %	N	Cumulatieve %	N	Cumulatieve %	N	Cumulatieve %
0%	20	33.9	15	27.8	24	4.4	22	40.0
25%	1	35.6	5	37.0	2	44.8	3	45.5
33%	0	35.6	0	37.0	1	46.6	0	45.5
50%	3	40.7	2	40.7	4	53.4	5	54.5
75%	4	47.5	8	55.6	4	60.3	2	58.2
100%	31	100.00	24	100.00	23	100.0	23	100.0
	59		54		58		55	

Opvallend in de resultaten van tabel 1 is dat de kinderen in beide condities, mens- en robotconditie, een duidelijke strategie lijken te hanteren en dat de kinderen tussen de verschillende trials door weinig wisselden van hun strategie. Dit was ook te zien in de verdeling tussen de verschillende percentages in de opvolging van het label (0%) en de opvolging van het wijzen (100%). De meeste kinderen gingen bij alle geslaagde trials voor het volgen van het label (in conditie 1 $N = 20$ en $N = 15$ en in conditie 2 $N = 24$ en $N = 22$) of juist bij alle geslaagde trials voor het volgen van het wijzen (in conditie 1 $N = 31$ en $N = 24$ en in conditie 2 $N = 23$ en $N = 23$). Minder kinderen wisselden tussen de trials tussen het volgen

van wijzen of het volgen van het label. In conditie 1 was dit $N = 8$ en $N = 15$ en in conditie 2 was dit $N = 11$ en $N = 10$.

Wijzen versus Labelen

De one-sample t-test tegen kansniveau werd uitgevoerd met de volgende hypothesen:

H0: De gemiddelde score op de conflicttaak = 0.50

H1: De gemiddelde score op de conflicttaak \neq 0.50

Het kansniveau is 0.50, want als een kind compleet random zou kiezen, zou het ongeveer de helft van de tijd het wijzen volgen en de helft van de tijd het labelen volgen. Wanneer dit het geval zou zijn, kom je uit op de helft van de trials, dus een proportie van 0.50.

Conditie 1. Uit de one-sample t-test bleek dat de gemiddelde score op de conflicttaak bij mensconditie 1 ($M = 0.61$, $SD = 0.46$) niet significant verschilt van de gemiddelde score ($M = 0.50$), $t(58) = 1.76$, $p = .084$. Op basis van het gemiddelde ($M = 0.61$) lijkt het alsof kinderen het wijzen meer volgen bij mensconditie 1, maar er is echter geen significant verschil met kansniveau.

Daarnaast zagen we ook dat de gemiddelde score op de conflicttaak voor robotconditie 1 ($M = 0.60$, $SD = 0.44$) niet significant verschilt van de gemiddelde score ($M = 0.50$), $t(53) = 1.64$, $p = .107$. Hierbij lijkt het ook op basis van het gemiddelde ($M = 0.60$) alsof kinderen het wijzen meer volgen dan het labelen, maar er is echter geen significant verschil met kansniveau.

Conditie 2. Ook in conditie 2 bleek uit de one-sample t-test dat de gemiddelde score op de conflicttaak bij mensconditie 2 ($M = 0.50$, $SD = 0.46$) niet significant verschilt van de gemiddelde score ($M = 0.50$), $t(57) = -.05$, $p = .962$.

Dit gold tevens voor de gemiddelde score op de conflicttaak bij robotconditie 2 ($M = 0.50$, $SD = 0.46$). Deze verschilde ook niet significant van de gemiddelde score ($M = 0.50$), $t(54) = .07$, $p = .942$.

Agent en Conditie

Een repeated measures anova is uitgevoerd om te kijken of er een effect is van agent (mens en robot), van conditie (bekend label – onbekend label) en de van de combinatie agent en conditie. Uit de resultaten van de repeated-measures anova blijkt dat de mate waarin kinderen het wijsgedrag van mens en robot volgden niet significant verschilt voor agent (mens en robot), $F(1,48) = 0.30$, $p = .585$, niet significant verschilt voor de condities, $F(1,48) = 1.79$, $p = .188$ en niet significant verschilt voor de interactie agent en conditie, $F(1, 48) = 0.64$, $p = .427$. Deze resultaten komen dus niet overeen met de vooraf opgestelde hypothese.

Woordenschat

Daarnaast werd een repeated measures anova uitgevoerd om te kijken of er een significant verschil is in het opvolgen van wijsgedrag tussen kleuters met een grote of kleine woordenschat. De hypothese bij deze onderzoeksvraag is dat verwacht wordt dat kleuters met een grote woordenschat meer het wijzen zullen volgen van mens en robot dan kleuters met een kleine woordenschat. In tabel 2 staan de gemiddelden en standaarddeviaties met betrekking tot de mate van volgzzaamheid van het wijzen voor kleuters met een kleine en grote woordenschat. De repeated-measures anova toonde aan dat er geen effect was tussen kinderen met een grote en kleine woordenschat in de mate van opvolging van wijsgedrag $F(1,47) = 0.003, p = .956$ en geen significant verschil tussen kinderen met een grote en kleine woordenschat ($p > .05$). Daarnaast werd er geen effect gemeten tussen de verschillende condities; $F(1, 47) = 1.79, p = .187$ en daarmee ook geen significant verschil tussen de verschillende condities ($p > .05$). Er waren geen interactie-effecten ($p = .956$).

Tabel 2

Gemiddelden (M) en Standaarddeviaties (SD) Opvolgen Wijzen Kleuters met Kleine en Grote Woordenschat

	Kleine woordenschat (N = 21)	Grote woordenschat (N = 28)
	M (SD)	M (SD)
Mensconditie 1	0.65 (0.48)	0.61 (0.47)
Robotconditie 1	0.56 (0.44)	0.62 (0.45)
Mensconditie 2	0.55 (0.47)	0.48 (0.47)
Robotconditie 2	0.51 (0.48)	0.52 (0.46)

Conclusie en Discussie

De onderzoeksvraag die in deze studie centraal stond was: 'Is er een verschil in de mate waarin kleuters het wijsgedrag volgen tussen mens en robot bij een referentieel conflict en verschilt dit tussen kleuters met een grote en kleine woordenschat?' Het belang van deze studie was om te onderzoeken of kleuters bij het leren met een sociale robot ook gebruik maken van lexicale principes (mutual exclusivity) en van sociaal-pragmatische cues (wijzen) en of dit verschilt met de mens. Deze twee principes spelen namelijk een belangrijke rol wanneer volwassen kinderen iets willen leren. Het is daarom nuttig en interessant om te weten of

dergelijke principes ook voor robots gelden, zodat men robots zo effectief mogelijk kan inzetten binnen het onderwijs.

De eerste hypothese bij de onderzoeksvraag was dat kleuters eerder het wijsgedrag van de robot zouden volgen dan van de mens. Deze hypothese kwam echter niet overeen met de resultaten. Er werden geen significante verschillen gevonden in de mate waarin kleuters het wijsgedrag van mens en robot volgden voor agent (mens en robot), voor de verschillende condities (bekend-label en onbekend-label) en voor de combinatie tussen agent en conditie. Ook kwamen de resultaten niet overeen met de resultaten vanuit eerder onderzoek. Jawsal en Hansen (2006) concludeerden namelijk dat kinderen voornamelijk het voorwerp selecteerden dat benoemd was met woorden. Echter kwam uit de resultaten van het experiment van Grassmann en Tomasello (2010) dat kinderen eerder het wijzen volgden dat het benoemen bij de conflicttaak. Beide conclusies (Jawsal & Hansen, 2006; Grassmann & Tomasello, 2010) kwamen dus niet overeen met de resultaten uit deze studie, aangezien er geen significante verschillen gevonden zijn voor de opvolging wijzen en benoemen.

De tweede hypothese was dat verwacht werd dat kleuters met een grote woordenschat meer gebruik zullen maken van sociaal-pragmatische informatie, zoals wijzen. Vanuit eerder onderzoek werd geconcludeerd dat kinderen met een grote woordenschat waarschijnlijk meer gericht waren op de bedoeling van de omgeving, de pragmatische informatie (wijzen). Deze hypothese werd echter ook niet bevestigd door de resultaten van deze studie. Er werd geen effect gemeten tussen kinderen met een grote en kleine woordenschat in de mate van opvolging van wijsgedrag, ook niet tussen de verschillende condities. Er werden ook geen interactie-effecten gemeten. Deze resultaten kwamen niet overeen met de resultaten vanuit eerder onderzoek. Verhagen, Grassman en Küntay (2016) stelden dat in de meeste gevallen gold; hoe groter de woordenschat van kinderen, hoe vaker zij mutual exclusivity toepasten. Echter worden er in huidige studie geen significante verschillen gevonden tussen de kinderen met een grote woordenschat en de kinderen met een kleine woordenschat met betrekking tot de opvolging van wijsgedrag bij een conflicttaak.

Het huidige onderzoek kent een aantal sterke punten en enkele beperkingen. Het gebruik van de robot is een voorbeeld van een sterk punt, aangezien dit nog nauwelijks onderzocht is tijdens onderzoek naar kleuters in combinatie met een referentiële conflicttaak. Een ander sterk punt is dat de betrouwbaarheid en validiteit van de PPVT goed is. Daarnaast is, ondanks dat de betrouwbaarheid en validiteit van de conflicttaak nog onbekend zijn, wel goede kwaliteit

nagestreefd door onder andere de manier van instructie geven tussen de verschillende testleiders goed op elkaar af te stemmen, zodat iedereen hetzelfde deed en zei.

Een beperking van het huidige onderzoek is dat de steekproef relatief klein was. Er viel een groot deel van de steekproef uit (missings), aangezien van deze kinderen geen informatie meegenomen kon worden over een bepaalde conditie, ofwel omdat ze ziek waren ofwel omdat alle trials mislukt waren. Dit maakt de uitgevoerde analyses minder betrouwbaar. Tevens werd voordat of nadat de conflicttaak werd afgenomen ook een andere taak afgenomen (PPVT, Kiraly of perceptievragenlijst). Dit kan van invloed zijn geweest op de resultaten, aangezien kinderen in mindere het wijzen opgevolgd hebben dan in vergelijkbare studies (Grassmann & Tomasello, 2010). Tevens betreft dit onderzoek voornamelijk kinderen van hoogopgeleide ouders (44.3% WO-niveau), waardoor de resultaten niet generaliseerbaar zijn naar de gehele populatie. Allerlaatst was de robot die gebruikt werd in het experiment nog niet in staat om verbale en non-verbale signalen van kinderen waar te nemen en hierop te reageren of om emoties te uiten. Wanneer dit wel aanwezig geweest zou zijn, zou dit leiden tot een effectievere samenwerking met de robot wat van invloed zou kunnen zijn op de resultaten (Chang et al., 2010; Tanaka & Matsuzoe, 2012).

Uit deze beperkingen volgen verschillende implicaties voor toekomstig onderzoek. Ten eerste moet er gebruik gemaakt worden van een grotere steekproef, zodat de analyses en resultaten een betrouwbaarder beeld weer kunnen geven. Dit kan wellicht leiden tot een bevestiging van de huidige bevindingen of juist het aannemen van vooraf opgestelde hypothesen. Daarnaast dient de robot verder ontwikkeld te worden, zodat het ook emoties kan uiten en kan reageren op de signalen van kinderen. Op deze manier kan de interactie tussen de robot en het kind bevorderd worden wat van invloed kan zijn op de effectiviteit van de conflicttaak. Dit alles kan een bijdrage leveren aan het vergroten van kennis om robots op een zo effectief mogelijke manier in te zetten binnen het onderwijs.

Referenties

- Alemi, M., Meghdari, A., & Ghazisaedy, M. (2014). Employing humanoid robots for teaching English language in Iranian junior high-schools. *International Journal of Humanoid Robotics, 11*, 1450022-1-1450022–24. doi: 10.1142/S0219843614500224
- Allen, P., & Bennett, K. (2012). *SPSS Statistics, a practical guide version 20*. Australia: Cengage Learning Australia.
- 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
- Bartneck, C., & Forlizzi, J. (2004). A Design-Centred Framework for Social Human-Robot Interaction. *Proceedings of the Ro-Man, Kurashiki*, (pp. 591-594). doi: 10.1109/ROMAN.2004.1374827
- Bion, R.A.H., Borovsky, A., Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novelword- object mappings in relation to vocabulary learning at 18, 24 and 30 months. *Cognition, 126*, 39-53.
- 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.
- Dunn, L. M., Dunn, L. M., & Schlichting, L. (2005). *Peabody picture vocabulary test-III-NL*. Amsterdam: Pearson.
- Enfield, N.J., Kita, S., & de Ruiter, J.P. (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics, 39* (10), 1722–1741.
- Gliga, T., & Csibra, G. (2009). One-year old infants appreciate the referential nature of deictic gestures and words. *Psychological Science, 20*(3), 347-353.
- 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
- Grassmann, S. & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science, 13*, 252-263. doi:10.1111/j.1467-7687.2009.00871.x

- Hansen, M. B., & Markman, E. M. (2009). Children's use of mutual exclusivity to learn labels for parts of objects. *Developmental psychology*, *45*, 592-596. doi: 10.1037/a0014838.
- 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
- Houston- Price, C., Caloghris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, *15*, 125-150.
- Jaswal, V. K., & Hansen, M. B. (2006). Learning words: Children disregard some pragmatic information that conflicts with mutual exclusivity. *Developmental Science*, *9* (2), 158-165. doi: 10.1111/j.1467-7687.2006.00475.x
- Jipson, J., & Gelman, S. (2007). Robots and Rodents: Children's Inferences about Living and Nonliving Kinds. *Child Development*, *78*(6), 1675-1688.
- Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2015). Higher nonverbal immediacy leads to greater learning gains in child-robot tutoring interactions. *Proceedings of the International Conference on Social Robotics*, 327–336. doi: 10.1007/978-3-319-25554-5
- Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2016). *Social robot tutoring for child second language learning*. In 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 231–238).
- Kiraly, I. (2009). Memories for events in infants: Goal-relevant action coding. In T. Striano & V. Reid (Eds.), *Social cognition development, neuroscience, and autism* (pp. 113-128).
- Lin, Y. C., Liu, T.Z., Chang, M., & Yeh, S.P. (2009). *Exploring Children's Perceptions of the Robots*. Proceedings of the 4th International Conference on E-Learning and Games: Learning by Playing. Game-based Education System Design and Development. Banff, Alberta, Canada, Springer-Verlag: p. 512-517.
- Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.
- Markman, E. M. (1992). Constraints on word learning: Speculations about their nature, origins and domain specificity. In M. R. Gunnar & M. Maratsos (Eds.), *Modularity and constraints in language and cognition*, *25*, 59-101.

- Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, *47*, 241-275. doi:10.1016/S0010-0285(03)00034-3
- Markman, E.M., & Wachtel, G.F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*, 121-157.
- Masur, E.R. (1997). Maternal labeling of novel familiar objects: Implications for children's development of lexical constraints. *Journal of Child Language*, *24*(2), 427-439.
- Mazzoni, E., & Benvenuti, M. (2015). A robot-partner for preschool children learning English using socio-cognitive conflict. *Educational Technology & Society*, *18*, 474-485.
- Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Developmental Review*, *11*, 137-163. doi: 10.2307/1166130
- 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 HumanRobot Interaction*, *1*, 78-95. doi:10.5898/JHRI.1.1.Tanaka
- Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics*, *10*, 401-414.
- Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (132-158). Cambridge: Cambridge University Press.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
- Verhagen, J., Grassmann, S., & Küntaya, A.C. (2016). Monolingual and bilingual children's resolution of referential conflicts: Effects of bilingualism and relative language proficiency. *Cognitive Development*, *41*, 10-18. doi: 10.1016/j.cogdev.2016.10.003