

Social robots and L2 teaching for children: a preliminary study on the embodiment of gestures

Robot sociali e insegnamento di L2 ai bambini: uno studio preliminare sull'embodiment della gestualità

Silvia Gasparini^a

^a *Università degli Studi di Trieste, nerone30@libero.it*

Abstract

Experimental research confirms that using social robots with pre-school children as tutors for their learning has important positive effects in terms of children's learning and emotional involvement. In particular, the use of social robots for language learning (L1 and L2) leads to more learned words and a better memory of them. However, many technological limitations prevent from fully implementing the concept of embodiment, a distinctive feature of robotic technology compared to other computer-based technologies, and especially of gestures. In order to contribute to solve the problem, in a preliminary psychological test with Italian children learning English as L2, it was shown that appropriate modifications of the learning environment combined with a more limited use of gestures allow to achieve better results in terms of learning than a larger use of gestures in a less structured environment. If experimentally verified, these results would suggest the development of learning architectures which offer more affordances with respect to the task at hand, while implementing more limited forms of embodiment that are in line with current technological limitations.

Keywords: social robots; L2 learning; embodiment; learning environment.

Abstract

Le ricerche sperimentali confermano che l'utilizzo di robot sociali in funzione di tutor per l'apprendimento di bambini nella fascia di età prescolare ha importanti effetti positivi in termini di apprendimento e di coinvolgimento emotivo dei bambini. In particolare, l'uso di robot in supporto all'apprendimento linguistico (L1 e L2) si traduce in un numero maggiore di parole apprese e in un migliore ricordo delle stesse. Numerosi limitazioni tecnologiche impediscono tuttavia la piena realizzazione nei robot sociali del concetto di embodiment, in particolare per quanto riguarda la gestualità. Proprio per contribuire a dare una soluzione al problema, in una prova psicologica preliminare con bambini italiani che apprendono la lingua inglese come L2 si dimostra che in un ambiente di apprendimento opportunamente strutturato una gestualità più limitata permette di conseguire migliori risultati di apprendimento e ricordo delle parole apprese rispetto a un uso maggiore della gestualità in un ambiente meno organizzato. Se sperimentalmente verificati, tali risultati indirizzerebbero verso lo sviluppo di architetture dell'ambiente di apprendimento più ricche di affordance rispetto al compito che consentano conseguentemente di implementare forme di embodiment più limitate in linea con le limitazioni tecnologiche attuali.

Parole chiave: robot sociali; apprendimento di L2; embodiment; ambiente di apprendimento.

1. Introduzione

Negli ultimi anni gli sforzi per costruire robot sociali in grado di supportare attività di apprendimento si sono intensificati. In particolare, nella fascia di età prescolare (3-5 anni) le recenti tecnologie Kindergarten Social Assistive Robotics (d'ora in poi KindSAR) consentono di sperimentare l'utilizzo di robot in contesti di gioco educativo e in situazioni di apprendimento cooperativo all'interno del paradigma educativo costruttivistico. Per quanto riguarda lo sviluppo linguistico, molti programmi sono finalizzati allo sviluppo delle competenze nella L1 e L2.

A fronte dei numerosi punti di forza che i robot sociali hanno dimostrato in termini di apprendimento nei bambini, si registrano tuttavia alcuni punti di debolezza legati in particolare all'impossibilità tecnica di implementare pienamente nella tecnologia robotica il concetto di *embodiment*, inteso come la possibilità di agire e apprendere in modo autonomo grazie alle interazioni senso-motorie con l'ambiente in analogia con i processi cognitivi umani. Le soluzioni proposte per aggirare gli ostacoli sono diverse. Ad esempio, nell'ambito del tutoraggio per l'apprendimento della L2 nei bambini, una soluzione frequentemente adottata consiste nell'affiancare l'uso dei robot ad altre tecnologie. Tale soluzione non appare pienamente convincente dal momento che l'uso combinato di più tecnologie rende l'ambiente di apprendimento più complesso, svuotando almeno in parte di significato le funzioni svolte dal robot. Ci si chiede se in alternativa sia possibile usare forme semplificate di *embodiment* agendo contemporaneamente sull'architettura dell'ambiente di apprendimento, in modo da offrire affordance fruibili dai bambini a fronte di aiuti più limitati da parte del robot. Tale prospettiva viene esplorata preliminarmente da un punto di vista psico-pedagogico con riferimento all'apprendimento di L2 e all'uso della gestualità. Il lavoro si articola pertanto nei seguenti punti:

- individuazione dei tratti distintivi del paradigma di riferimento;
- individuazione delle criticità;
- focalizzazione sulla psicologia della gestualità come strumento di intervento;
- presentazione e analisi a livello qualitativo di una prova di apprendimento di parole in L2 con il supporto di forme di gestualità semplificata;
- discussione finale.

2. Costruttivismo, apprendimento e tecnologie robotiche

In base alle teorie costruttivistiche un apprendimento efficace richiede che l'apprendente sia attivamente coinvolto in un processo di costruzione della conoscenza che si caratterizza per una forte componente di interazione sociale (fra pari e con adulti) e di co-costruzione di concetti e procedure di lavoro (Duffy, Lowick & Jonassen, 1993; Wilson, 1996). Per i bambini di età prescolare il concetto di apprendimento costruttivistico riferito al linguaggio è basato principalmente sul gioco interattivo alla scoperta dei concetti e dei significati delle parole e sulla costruzione di abilità sociali attraverso la sperimentazione di giochi linguistici finalizzati allo sviluppo di abilità comunicative riferite tanto alla L1 che alla L2 (Oldfather, West, White & Wilmarth, 1999).

Benché le sperimentazioni in tal senso siano poche, i risultati indicano che in ambienti di apprendimento così definiti, la presenza del robot con funzione di assistente per gruppi

ristretti di bambini o in interazioni personalizzate del tipo one-to-one è fruttuosa. Così, in uno studio su bambini di età prescolare Fridin (2014) dimostra che la presenza del robot in classe durante un'attività di story-telling aumenta il divertimento, la curiosità, il coinvolgimento emotivo dei bambini, che a loro volta si traducono in un maggior desiderio di apprendere e in migliori risultati in termini di rafforzamento delle abilità linguistiche proposte come obiettivo.

In modo simile, anche la sperimentazione della tecnologia nell'ambito dell'apprendimento di L2 a diversi livelli comincia a produrre risultati significativi. Così, uno studio durato due settimane con un robot in classe ha rivelato una correlazione positiva tra l'interazione con il robot e l'acquisizione del vocabolario (Kanda, Hirano, Eaton & Ishiguro, 2004). In modo simile, Alemi, Meghdari e Ghazisaedy (2014) dimostrano che la presenza di un assistente-robot in classe per un periodo di cinque settimane produce migliori risultati in termini di apprendimento e di ricordo delle parole apprese rispetto alla sola presenza dell'insegnante (si veda anche Tanaka & Matsuzoe, 2012). In Movellan, Eckhardt, Virnes e Rodriguez (2009) viene assegnato al robot il compito di insegnare dieci parole durante un periodo di permanenza in classe di 12 giorni. I risultati dello studio confermano un significativo aumento nel ricordo delle parole insegnate dal robot. Se negli studi esaminati è la comprensione e il ricordo delle parole a migliorare, Lee et al. (2011) dimostrano che l'utilizzo della tecnologia robotica per la seconda lingua migliora anche la pronuncia delle parole e, in generale, la fiducia nel compito.

In particolare, nell'ambito dell'applicazione della robotica all'apprendimento della L2 nei bambini risulta interessante analizzare il progetto europeo L2TOR (<http://www.l2tor.eu>) rivolto alla fascia di età 3-5 anni. Il progetto è finalizzato alla sperimentazione di un robot-tutor per l'apprendimento di L2 in bambini in età prescolare (tre-cinque anni). Esso prevede l'insegnamento dell'inglese a bambini olandesi, tedeschi e turchi e dell'olandese o del tedesco a bambini turchi immigrati che vivono nei Paesi Bassi o in Germania. Nello specifico, il progetto si propone di sviluppare una serie di lezioni che aiutino i bambini ad imparare un vocabolario di base in lingua straniera avendo il robot sociale autonomo SoftBank Robotics NAO come tutor (Baxter Ashurst, Read, Kennedy & Belpaeme, 2017; Belpaeme et al., 2015; Belpaeme et al., 2018). Come è noto, i pattern di apprendimento di una lingua richiedono input adeguati per quantità e soprattutto qualità (Hirsh-Pasek et al., 2015; Konishi, Kanero, Freeman, Golinkoff & Hirsh-Pasek, 2014). I bambini, infatti, non imparano una lingua soltanto ascoltando, ma necessitano di esperienze interattive di carattere socio-pragmatico che comprendono forme di attenzione condivisa, efficaci scambi non verbali, feedback adeguato, nonché contingenze temporali e semantiche che contribuiscano a definire il senso dei messaggi (Ateş-Şen & Kintay; 2015; Bornstein, Tamis-Le Monda, Hahn & Haynes, 2008). Inoltre, nel caso la lingua da imparare sia una L2, anche l'apporto della L1 è oggi riconosciuto come fondamentale. In linea con questi presupposti, le lezioni previste dal programma L2TOR sono costruite su interazioni personalizzate tra i bambini e NAO e sono concentrate sull'apprendimento del vocabolario presentato in forma di cluster semantici omogenei già noti ai bambini in L1. Ogni lezione (della durata di circa 15-20 minuti) è costruita sullo sfondo di uno scenario costituito da una città virtuale che il bambino e il robot esplorano insieme. Esso contiene vari negozi, aree pubbliche, edifici, la cui funzione è peraltro già nota ai bambini, che vengono scoperti uno ad uno nel corso delle lezioni. Le parole vengono così imparate giocando e svolgendo semplici compiti (come contare degli oggetti o accoppiare determinate figure con gli oggetti corrispondenti). L'insegnamento è previsto per un massimo di sei parole per lezione che vengono ripetute variando i contesti per almeno

dieci volte. Alla fine di ogni lezione i bambini vengono premiati allo scopo di migliorare il coinvolgimento nel compito.

3. La tecnologia KindSAR e l'embodiment imperfetto

Molte delle attività previste nell'ambito delle tecnologie KindSAR applicate all'apprendimento delle lingue possono essere espletate anche da altre tecnologie computer-based compatibili con i principi dell'apprendimento costruttivistico. A differenza delle tecnologie computer-based tuttavia, la tecnologia KindSAR mette a disposizione computer umanoidi che incarnano il concetto di embodiment (Shapiro, 2004) considerato il fattore in grado di garantire apprendimenti migliori rispetto agli ambienti di digital learning programmati per tablet e computer (Han, Jo, Jones & Jo, 2008; Leyzberg, Spaulding, Toneva & Scassellati, 2012). Secondo la prospettiva embodied applicata alla robotica avanzata, dal momento che i processi cognitivi sono considerati dipendenti dalla struttura fisica del corpo e dall'interazione con l'ambiente esterno (Borghi, Flumini, Cimatti, Marocco & Scorolli, 2011; Lakoff & Johnson, 1980; 1999; Scorolli et al., 2011), anche l'architettura dei robot deve fondamentalmente rispettare questi principi. In particolare, gli ambienti KindSAR per l'apprendimento delle lingue tendono a realizzare due forme complementari di embodiment. La prima riferita propriamente al robot che, nell'interagire con gli apprendenti, si presenta con caratteristiche tali da tendere all'annullamento delle distanze uomo-macchina per quanto riguarda i tratti della comunicazione verbale e non verbale. In particolare, in una sperimentazione che utilizza un iCat robot per insegnare ai bambini la lingua Toki Pona è stato osservato che i comportamenti embodied del robot riconducibili a interventi mirati a indirizzare l'attenzione, al feedback (verbale e non verbale) fornito, all'empatia e alla comunicatività dimostrata ai bambini, influiscono positivamente sul potenziale di apprendimento (Saerbeck, Shut, Bartneck & Janse, 2010), anche se ne vanno evitati gli eccessi (Herberg, Feller, Yengin & Saerbeck, 2015; Kennedy, Baxter, Senft & Belpaeme, 2016). La seconda forma di embodiment cui si fa riferimento riguarda l'ambiente di apprendimento in cui i bambini possono manipolare oggetti e interagire in un ambiente fisico reale, due aspetti fondamentali per apprendere una lingua e in particolare il vocabolario (Glenberg, 2008; 2010).

A fronte dei numerosi vantaggi riscontrati tuttavia, molti aspetti dell'embodiment non sono a tutt'oggi implementabili con successo. Così, nel progetto L2TOR sopra presentato gli aspetti embodied, riconducibili ai seguenti elementi: (i) primo incontro; (ii) framing; (iii) contesto dell'interazione bambino-robot; (iv) comportamento non verbale del robot; (v) comportamento verbale; (vi) feedback fornito al bambino, sono solo in parte implementati. In particolare, l'aspetto embodied è realizzato pienamente per i primi due aspetti: primo incontro e framing. Nel primo incontro, NAO viene introdotto nella stanza acceso mentre si guarda attivamente intorno, saluta, stringe mani, danza con i bambini. A differenza delle successive sessioni di apprendimento condotte nella forma one-on-one, l'incontro avviene nella forma one-to-many, cioè in piccoli gruppi, seguendo le indicazioni di Fridin (2014). Questo allo scopo di non intimorire i bambini che sono così messi nella condizione di superare eventuali resistenze e paure attraverso il supporto reciproco. Il framing prevede invece per il robot il ruolo di pari e amico che impara la lingua insieme al bambino (pur prevedendo che il robot fornisca organizzazione dei materiali e feedback da adulto esperto).

Per quanto riguarda gli altri aspetti, il computer NAO, pur costruito con tecnologie Human Robot Interaction avanzate, presenta numerose criticità. Per quanto riguarda l'aspetto del comportamento verbale, NAO può essere programmato per parlare diverse lingue rendendo in particolare possibile rivolgersi ai bambini sia in L1 che in L2. Le voci generate tuttavia, essendo sintetiche non hanno accenti particolari e non contengono tratti prosodici, considerati fondamentali specialmente per i bambini (Dominey & Dodane, 2004). Se teniamo presente che il robot NAO manca altresì di espressioni facciali, comprendiamo come i messaggi verbali generati possano pregiudicare la percezione audio-visiva dei bambini (Erber, 1975) potenzialmente compromettendo l'apprendimento della pronuncia delle parole, un aspetto tuttora oggetto di indagini nel quadro del progetto. D'altra parte, anche il riconoscimento automatico per il linguaggio del bambino non è sufficientemente affidabile rendendo l'utilizzo dello stesso sconsigliato (Kennedy et al., 2017). Questo aspetto impedisce al robot di rispondere direttamente nonché di esercitare un controllo sulle produzioni verbali dei bambini (parole pronunciate o risposte di altro genere). L'ostacolo viene aggirato mediante l'introduzione di tablet programmati: a) per mostrare il contesto di apprendimento contenente gli oggetti target; b) per monitorare le risposte dei bambini alle domande. Benché alcune ricerche indichino un'equivalenza in termini di efficacia sull'apprendimento tra l'embodiment di oggetti reali e l'embodiment di oggetti virtuali presentati su schermo interattivo (tablet) (Vlaar, Verhagen, Oudgenoeg-Paz & Leseman, 2017), l'introduzione della triangolazione bambino-tablet-robot quantomeno ridimensiona il ruolo originario del robot complessificando l'ambiente di apprendimento secondo modalità i cui effetti combinati sul bambino non sono al momento noti. Altri fondamentali elementi di criticità, le espressioni facciali e la gestualità che, insieme, caratterizzano il comportamento non verbale e che si attivano in molte forme di feedback. È noto che gli ascoltatori traggono beneficio dal vedere i movimenti delle labbra del parlante (Jesse & Massaro, 2010) tuttavia il robot non li può produrre ed è tuttora allo studio la possibilità di surrogare la limitazione con l'illuminazione degli occhi per indicare gioia, da impiegare come forma di feedback positivo. Anche la gestualità risulta non soddisfacente. In effetti, la possibilità di produrre gesti da parte del robot-tutor rappresenta uno dei principali vantaggi della tecnologia permettendo al robot di ancorare le parole ai concetti in modo intuitivo e diretto. Tuttavia, le limitazioni hardware attuali ne inficiano i potenziali benefici. Così, a causa dei limitati gradi di libertà, il robot non può compiere gesti con lo stesso grado di accuratezza degli umani. Questa limitazione può risultare in una perdita di rappresentatività che condiziona la scelta dei vocaboli da presentare attraverso i gesti alle possibilità del robot (de Wit et al., 2018). In altri casi non risulta possibile presentare il concetto target direttamente, ma solo attraverso la presentazione di gesti aggiuntivi, come ad esempio la simulazione del comportamento di un cavaliere per insegnare il significato di cavallo, essendo il gesto del cavalcare l'unico riproducibile con successo per simulare un cavallo. Ulteriori problemi sono rilevabili con i gesti deittici, in sé più facili in quanto anziché rappresentare si limitano a indicare la posizione spaziale dell'oggetto nell'ambiente. Dal momento che il robot NAO non è in grado di muovere le sue tre dita singolarmente non può indicare o contare nella stessa maniera degli umani. Questo comporta una scarsa comprensibilità del gesto utilizzato per indicare elementi contenuti nei piccoli schermi dei tablet, mentre sono tuttora oggetto di indagine gli effetti sulla comprensibilità del pointing in contesti reali che prevedono l'impiego di oggetti più grandi (Vogt, De Haas, De Jong, Baxter & Kramer, 2017). Senza contare che il robot produce suoni motori durante la produzione dei gesti che, se concomitanti alla pronuncia della parola, possono mascherarla.

In generale, le soluzioni consistono nell'aggirare le difficoltà mediante il ricorso ad altri supporti tecnologici. Nel caso delle limitazioni in voce rilevate sopra, ad esempio, si è optato per l'utilizzo in tandem del tablet. In altri casi, si è finito per preferire una tecnologia alternativa, come nel caso di una sperimentazione in cui per evitare il rischio che caratteristiche embodied troppo marcate (altezza e rigidità) del robot NAO potessero trasmettere insicurezza ai bambini, si è optato per l'utilizzo di TEGA, soffice e abbracciabile (Westlund, Martinez, Archie, Das & Breazeal, 2016). Nel caso specifico, prima di prendere in considerazione soluzioni che si propongono di aggirare i limiti della tecnologia dall'esterno, si intende valutare la possibilità di utilizzare forme semplificate di embodiment. A questo punto il problema è comprendere se le modifiche eventualmente apportabili abbiano un fondamento psicologico ai fini dell'apprendimento e a quali condizioni possano essere proficuamente implementate nell'architettura dell'ambiente di apprendimento.

4. Gestualità e linguaggio: iconicità e deissi a confronto

I gesti che accompagnano il discorso sono principalmente distinti in deittici, come l'indicare oggetti presenti nell'ambiente condiviso tra parlante e ascoltatore, o iconici, usati per descrivere caratteristiche fisiche degli oggetti, in particolare la dimensione, o, secondo una logica ancora più rappresentativa, le azioni. In particolare, i gesti figurano come oggetto di indagine nella duplice prospettiva di chi li produce e di chi li riceve. Dal punto di vista della gestualità recepita che qui direttamente ci interessa, si può osservare che benché non vi sia universale consenso riguardo la reale efficacia della gestualità sulla comprensione (Krauss, Chen & Chawla, 1996; Krauss, Dushay, Chen & Rauscher, 1995), molte ricerche pervengono alla conclusione che la comprensione migliora quando l'ascoltatore ha la possibilità di vedere i gesti prodotti dal parlante (Kelly, 2001; Valenzeno, Alibali & Klatzky, 2003). Recentemente, alle ricerche comportamentali si affiancano ricerche che utilizzano le metodologie delle neuroscienze per misurare l'elaborazione della gestualità nel cervello (Kelly, Ward, Creigh & Bartolotti, 2007; Skipper, Goldin-Meadow, Nusbaum & Small, 2007; Wu & Coulson, 2007). Alcune di queste indicano che l'elaborazione semantica dei gesti e del linguaggio condividono gli stessi network cerebrali (Kelly, Özyurek & Maris, 2010; Özyurek, Willems, Kita & Haagort, 2007) costituendo così una prova importante che, essendo elaborati insieme, linguaggio e gestualità siano egualmente informativi al fine della comprensione.

In particolare, la gestualità sembra avvantaggiare alcune tipologie di ascoltatori, fra cui (i) i bambini; (ii) gli apprendenti L2 ad ogni livello; (iii) i soggetti con ritardo mentale o disordini neurologici, sebbene in quest'ultimo caso con un grado minore di consenso dovuto al fatto che condizioni complesse come la sindrome di Down o l'autismo, avendo effetti variabili sulla comprensione del linguaggio, non permettono di condurre ricerche omogenee quanto a campionatura dei gruppi (Groen, Zwiers, van der Gaag & Buitelaar, 2008; Miolo, Chapman & Sindberg, 2005).

Negli studi che hanno esaminato il ruolo dei gesti nell'acquisizione del linguaggio infantile (Capone & McGregor, 2005; Goodrich, Sauer, & Iverson, 2007; Kelly & Church, 1997) emerge in particolare che nei bambini i gesti aiutano a legare la parola sconosciuta al corrispondente concetto del mondo reale attraverso un legame di tipo iconico o indicativo (Alibali & Nathan, 2007; Ping & Goldin-Meadow, 2008). È anche possibile che i bambini si avvantaggino della gestualità perché i gesti sono particolarmente efficaci nel catturare la loro attenzione. Così, Valenzeno, Alibali e

Klatzky (2003) dimostrano che i bambini si distraggono maggiormente in una videolezione se questa contiene pochi gesti rispetto ad una che ne contiene in misura maggiore. I gesti promuovono anche un migliore rapporto con il parlante predisponendo così a un ascolto migliore (Maricchiolo, Gnisci, Bonaiuto & Ficca, 2009).

Per quanto riguarda gli effetti del gesto negli apprendenti di L2, Sueyoshi e Hardison (2005) hanno dimostrato che la presenza di gesti migliora la comprensione dei soggetti che hanno minore competenza della L2. I dati sono in accordo con quelli che si ricavano dagli studi sul rapporto tra gestualità e bilinguismo infantile che dimostrano una maggiore sensibilità dei bambini bilingui alla gestualità, dovuta alla imperfetta conoscenza di una delle lingue (Brojde, Ahmed, & Colunga, 2012; Wermelinger, Gampe & Daum, 2017; Yow & Markman, 2011).

Come le ricerche esaminate indicano, la gestualità è fondamentale per aumentare la comprensione nei bambini impegnati nell'apprendimento di L2. Non è pertanto pensabile di sopprimerne l'utilizzo per difficoltà tecniche. Nello specifico, con riferimento alla gestualità imperfetta rilevata per NAO nell'ambito del progetto L2TOR, si può ipotizzare di dare la preminenza a forme di gestualità più facilmente riproducibili come il gesto deittico. Nel seguente esperimento pertanto si intende valutare in via preliminare le ricadute psicologiche di una semplificazione dell'embodiment gestuale in due contesti di apprendimento di parole in L2, diversamente strutturati rispetto alle distanze previste fra gli oggetti. Infatti, è ipotizzabile che se un ambiente di apprendimento più caotico offre meno affordance da poter essere sfruttate in autonomia dall'apprendente richiedendo conseguentemente al facilitatore/tutor una gestualità più articolata, un ambiente di apprendimento costruito secondo criteri più vincolanti e in grado di fare emergere a livello percettivo i rapporti concettualmente importanti (Gibson, 1979), potrebbe richiedere minori interventi direttivi esterni, gestualità compresa.

5. Presentazione e analisi di un'osservazione su gestualità e apprendimento del vocabolario L2 nei bambini

Scopo della ricerca è sperimentare l'efficacia di due diverse forme di gestualità in due diversi ambienti di apprendimento a supporto dell'apprendimento del vocabolario in L2 in bambini frequentanti la scuola dell'infanzia. Hanno partecipato all'esperimento 32 bambini di età compresa fra i quattro e i cinque anni (età media = 4,6 anni; ds = 3.2 mesi) il cui piano formativo prevede un avvicinamento graduale alla lingua inglese. L'esperimento prevede che i bambini siano coinvolti nella ricerca di cinque animali giocattolo in peluche (mucca-cow; elefante-elephant; cane-dog; uccello-bird; giraffa-giraffe) posti a distanza ravvicinata (scenario 1) o in alternativa distanziati fra loro (scenario 2) su un tappetino in panno verde di forma quadrata (cm 70 x 70) posto su un tavolino in un angolo tranquillo del salone dei giochi. I bambini, che conoscevano i nomi degli animali in italiano ma non in inglese avendo appena iniziato a familiarizzare con la lingua, sono stati assegnati casualmente alle seguenti condizioni: animali vicini/gesto rappresentativo (otto bambini); animali vicini/gesto deittico (otto bambini); animali distanziati/gesto rappresentativo (otto bambini); animali distanziati/gesto deittico (otto bambini). L'attività veniva presentata ai bambini in forma di gioco. Presi singolarmente, essi erano invitati a ricercare uno alla volta gli animali il cui nome veniva pronunciato in inglese e risultava quindi sconosciuto ai bambini. L'uso della lingua italiana con la sola pronuncia del nome dell'animale in inglese segue la procedura seguita nelle diverse sperimentazioni effettuate nell'ambito del progetto L2TOR. Per evitare che la ricerca

fosse casuale, la pronuncia del nome veniva accompagnata da un gesto che poteva essere rappresentativo ovvero deittico. Il primo doveva indicare approssimativamente le dimensioni dell'animale (braccia allargate orizzontalmente a indicare la lunghezza; braccia estese verticalmente ad indicare l'altezza); il secondo consisteva in braccio e dito teso. Si è scelta questa tipologia di gesto rappresentativo piuttosto che forme più iconiche (battito d'ali per l'uccello, etc.) perché potenzialmente generalizzabile ad altri animali. Se il bambino falliva nell'identificazione, la procedura (ripetizione della parola + gesto) veniva ripetuta fino all'identificazione corretta. Al ritrovamento dell'animale designato esso veniva nuovamente nominato in inglese, il bambino poteva prenderlo in mano e veniva invitato a ripeterne il nome per almeno cinque volte con frasi amichevoli del tipo "Te lo ricordi come si chiama?"; "Chi è lui?" etc. Successivamente al bambino venivano impartiti cinque brevi ordini in italiano mentre la parola-obiettivo era in inglese (del tipo "Metti la cow vicino alla giraffa"; "Dammi la cow", etc.) con lo scopo di fissare la parola appresa mediante azioni di manipolazione. Anche in questo caso sono state seguite le indicazioni del progetto L2TOR che suggeriscono una media di dieci ripetizioni per ogni parola nuova. Il giorno dopo, per verificare se ricordavano i nomi degli animali, a ciascun bambino è stato chiesto di trovare gli animali (nominati in inglese) nel medesimo scenario. In questo caso la richiesta è stata formulata senza accompagnamento del gesto. Sono stati valutati il numero di tentativi prima di raggiungere l'animale obiettivo sia per quanto riguarda la prova di apprendimento che per quanto riguarda la prova di ricordo. Per quest'ultima prova i gruppi sono costituiti dai bambini che nella prova precedente avevano appreso il nome dopo un numero di tentativi inferiore o uguale a due (gruppo 1) e dai bambini che avevano fatto più tentativi prima di giungere al nome (gruppo 2).

Apprendimento (scenario 1)		
	Tentativi 0-2	Tentativi \geq 3
Gesto rappresentativo	2 (25%)	6 (75%)
Gesto deittico	4 (50%)	4 (50%)
Apprendimento (scenario 2)		
	Tentativi 0-2	Tentativi \geq 3
Gesto rappresentativo	3 (37,5%)	5 (62,5%)
Gesto deittico	6 (75%)	2 (25%)
Ricordo nomi		
	Tentativi 0-2	Tentativi \geq 3
Gruppo 1 (15 bambini)	10 (67%)	5 (33%)
Gruppo 2 (17 bambini)	8 (47%)	9 (53%)

Figura 1. Percentuali di risposte nelle tre prove.

I risultati, presentati a livello qualitativo nella Figura 1, indicano che nell'ambiente di apprendimento in cui gli animali comparivano vicini (scenario 1) i bambini che hanno avuto la possibilità di vedere i gesti deittici hanno identificato gli animali con un numero di tentativi compresi tra 0 e 2 in misura maggiore rispetto ai bambini che hanno visto il gesto rappresentativo (rispettivamente il 50% contro il 25%). Questo vantaggio è stato mantenuto anche nello scenario 2 (animali distanziati) dove il 75% dei bambini che hanno

visto i gesti deittici hanno identificato l'animale designato in 0-2 tentativi contro il 37,5% dei bambini che hanno visto i gesti rappresentativi. Come si vede, i risultati della prova ambientata nel secondo scenario sono migliori per tutti i gruppi rispetto ai risultati della prova in cui gli animali erano presentati vicini (75% contro 50% gesto deittico; 37,5 contro 25% gesto rappresentativo per i tentativi 0-2). Nella prova di ricordo infine, i bambini che avevano identificato con meno tentativi l'animale nella prova di apprendimento dimostravano di ricordarne il nome più facilmente dei bambini che avevano avuto bisogno di più tentativi (67% contro 47%).

6. Conclusioni

Benché recente, la tecnologia dei robot sociali a supporto dell'apprendimento delle lingue straniere nei bambini di età prescolare ha già dimostrato numerosi vantaggi in termini di coinvolgimento emotivo e risultati di apprendimento. A fronte delle ricadute positive si registrano, tuttavia, alcune criticità legate in particolare a difficoltà implementative del concetto di embodiment. Rendere l'embodiment più semplice e nel contempo aumentare la significatività dell'ambiente di apprendimento potrebbe rappresentare una via per affrontare il problema. Questa possibilità è stata qui esplorata per la gestualità utilizzata a supporto di un compito di apprendimento del vocabolario in L2. I risultati dello studio preliminare tendenzialmente indicano che la gestualità di tipo rappresentativo apporta minori benefici all'apprendimento rispetto all'indicazione, che risulta più diretta. Un contributo importante ai risultati di apprendimento sembra darlo anche l'organizzazione dell'ambiente che, nel caso offra affordance immediatamente percepibili – vedi la maggiore distanza fra gli oggetti – facilita l'apprendimento in tutte le condizioni. D'altra parte, un apprendimento più efficace sembra influenzare anche i risultati del ricordo dei significati a distanza di tempo.

Se la ricerca, replicata nella forma pre-/post-test e con il coinvolgimento di un numero adeguato di soggetti, confermasse a livello sperimentale i risultati della presente osservazione, risulterebbe possibile passare a una sperimentazione con il robot NAO che, come si è ricorda, presenta per il gesto deittico alcune limitazioni. Anche in relazione a questo problema, la ricerca potrebbe essere sviluppata sperimentando, se necessarie, ulteriori, più incisive, forme di strutturazione dell'ambiente di apprendimento non solo fisico, ma anche a livello di interazioni verbali.

Bibliografia

- 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(3), 1–24. <https://doi.org/10.1142/S0219843614500224> (ver. 15.04.2019).
- Alibali, M.W., & Nathan, M. (2007). Teachers' gestures as a means of scaffolding students' understanding: Evidence from an early algebra lesson. In R. Goldman, R. Pea, B. Barron & S.J. Derry (eds.), *Video research in the learning sciences* (pp. 349-365). Mahwah, NJ: Erlbaum.

- Ateş-Şen, A.B., & Küntay, A.C. (2015). Children's sensitivity to caregiver cues and the role of adult feedback in the development of referential communication. *The acquisition of reference*, 241–262.
- Baxter, P., Ashurst, E., Read, R., Kennedy, J., & Belpaeme, T. (2017). Robot education peers in a situated primary school study: personalisation promotes child learning. *PLoS One*, 12(5), 1–23.
- Belpaeme, T., Kennedy, J., Baxter, P., Vogt, P., Krahmer, E.J., Kopp, S., ... & Deblieck, T. (2015). L2TOR-second language tutoring using social robotics. *Proceedings of first international workshop on educational robotics*. Paris, France.
- Belpaeme, T., Vogt, P., van den Berghe, R., Bergmann, K., Göksun, T., de Haas, M., ... Pandey, A.K. (2018) Guidelines for designing social robots as second language tutors. *International Journal of Social Robotics*, 10, 325–341. <https://doi.org/10.1007/s12369-018-0467-6> (ver. 15.04.2019).
- Borghi, A.M., Flumini, A., Cimatti, F., Marocco, D., & Scorolli, C. (2011). Manipulating objects and telling words: a study on concrete and abstract words acquisition. *Frontiers in Psychology*, 2, 1–14.
- Bornstein, M.H., Tamis-LeMonda, C.S., Hahn, C.S., & Haynes, O.M. (2008). Maternal responsiveness to young children at three ages: longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, 44(3), 867–874.
- Brojde, C.L., Ahmed, S., & Colunga, E. (2012). Bilingual and monolingual children attend to different cues when learning new words. *Frontiers in Psychology*, 3, 155. <https://doi.org/10.3389/fpsyg.2012.00155> (ver. 15.04.2019).
- Capone, N.C., & McGregor, K.K. (2005). The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, and Hearing Research*, 48, 1468–1480.
- de Wit, J., Schodde, T., Willemsen, B., Bergmann, K., de Haas, M., Kopp, S., ... Vogt, P. (2018). The effect of a robot's gestures and adaptive tutoring on children's acquisition of second language vocabularies. *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 50–58. <https://research.tilburguniversity.edu/en/publications/the-effect-of-a-robots-gestures-and-adaptive-tutoring-on-children> (ver. 15.04.2019).
- Dominey, P.F., & Dodane, C. (2004) Indeterminacy in language acquisition: the role of child directed speech and joint attention. *Journal of Neurolinguistics*, 17(23), 121–145.
- Duffy, T.M., Lowick, J., & Jonassen, D.H. (eds.). (1993). *The design of constructivistic learning environments: implications for instructional design and the use of technology*. Heidelberg: Springer-Verlag.
- Erber, N.P. (1975). Auditory-visual perception of speech. *Journal of Speech and Hearing Disorders*, 40(4), 481–492.
- Fridin, M. (2014). Storytelling by a kindergarten social assistive robot: a tool for constructive learning in preschool education. *Computer Education*, 70, 53–64.
- Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

- Glenberg, A.M. (2008). Embodiment for education. In P. Calvo & Gomila, A. (eds.), *Handbook of cognitive science: an embodied approach* (pp. 355-372). San Diego, CA: Elsevier.
- Glenberg, A.M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(4), 586–596.
- Goldin-Meadow, S., Goodrich, W., Sauer, E., & Iverson, J. (2007). Young children use their hands to tell their mothers what to say. *Developmental Science*, 10, 778–785.
- Groen, W.B., Zwiers, M. P., van der Gaag, R.J., & Buitelaar, J.K. (2008). The phenotype and neural correlates of language in autism: An integrative review. *Neuroscience & Biobehavioral Reviews*, 32, 1416–1425.
- Han, J.H., Jo, M.H., Jones, V., & Jo, J.H. (2008). Comparative study on the educational use of home robots for children. *Journal of Information Processing Systems*, 4(4), 159–168.
- Herberg, J., Feller, S., Yengin, I., & Saerbeck, M. (2015). Robot watchfulness hinders learning performance. *Proceedings of the 24th IEEE international symposium on robot and human interactive communication*, 153–160.
- Hirsh-Pasek, K., Adamson, L.B., Bakeman, R., Owen, M.T., Golinkoff, R.M., Pace, A., ...Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological science*, 26(7), 1071–1083. <https://doi.org/10.1177/0956797615581493> (ver. 15.04.2019).
- Jesse, A., & Massaro, D.W. (2010). The temporal distribution of information in audiovisual spoken-word identification. *Attention, Perception, & Psychophysics*, 72, 209–225. <https://doi.org/10.3758/APP.72.1.209> (ver. 15.04.2019).
- Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Journal of Human-Computer Interaction*, 19(1), 61–84.
- Kelly, S.D. (2001). Broadening the units of analysis in communication: Speech and nonverbal behaviours in pragmatic comprehension. *Journal of Child Language*, 28, 325–349. <https://doi.org/10.1017/S0305000901004664> (ver.15.04.2019).
- Kelly, S.D., & Church, R.B. (1997). Can children detect conceptual information conveyed through other children's nonverbal behaviors?. *Cognition and Instruction*, 15, 107–134.
- Kelly, S.D., Ward, S., Creigh, P., & Bartolotti, J. (2007). An intentional stance modulates the integration of gesture and speech during comprehension. *Brain & Language*, 101, 222–233.
- Kelly, S. D., Özyurek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture manually interact to enhance comprehension. *Psychological Science*, 21, 260–267.
- Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2016). Social robot tutoring for child second language learning. *Proceedings of the 11th ACM/IEEE international conference on human-robot interaction*, ACM, 67–74.

- Kennedy, J., Lemaignan, S., Montassier, C., Lavalade, P., Irfan, B., Papadopoulos, F, ...Belpaeme, T. (2017). Child speech recognition in human-robot interaction: evaluations and recommendations. *Proceedings of the 12th ACM/IEEE international conference on human-robot interaction, ACM*, 82–90.
- Konishi, H., Kanero, J., Freeman, M.R., Golinkoff, R.M., & Hirsh-Pasek, K. (2014) Six principles of language development: implications for second language learners. *Developmental Neuropsychology*, 39(5), 404–420.
- Krauss, R.M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us?. *Advances in Experimental Social Psychology*, 28, 389–450.
- Krauss, R.M., Dushay, R.A., Chen, Y., & Rauscher, F. (1995). The communicative value of communicative hand gestures. *Journal of Experimental Social Psychology*, 31, 533–552.
- L2TOR. Second language tutoring using social robot. <http://www.l2tor.eu> (ver. 15.04.2019).
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: the embodied mind and its challenge to western thought*. New York, NY: Basic Books.
- Lee, S., Noh, H., Lee, J., Lee, K., Lee, G.G., Sagong, S., & Kim, M. (2011) On the effectiveness of robot-assisted language learning. *ReCALL*, 23(01), 25–58.
- Leyzberg, D., Spaulding, S., Toneva, M., & Scassellati, B. (2012). The physical presence of a robot tutor increases cognitive learning gains. *Proceedings of the 34th annual conference of the cognitive science society*. New Haven, USA.
- Maricchiolo, F., Gnisci, A., Bonaiuto, M., & Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers' evaluations. *Language and Cognitive Processes*, 24, 239–266.
- Miolo, G., Chapman, R.S., & Sindberg, H.A. (2005). Sentence comprehension in adolescents with Down Syndrome and typically developing children: Role of sentence voice, visual context, and auditory-verbal short-term memory. *Journal of Speech, Language, and Hearing Research*, 48, 172–188.
- Movellan, J.R., Eckhardt, M., Virnes, M., & Rodriguez, A. (2009) Sociable robot improves toddler vocabulary skills. *Proceedings of the 4th ACM/IEEE international conference on human-robot interaction*, 307–308.
- Oldfather, P., West, J., White, J., & Wilmarth, J. (1999). *Learning through children's eyes: Social constructivism and the desire to learn*. Washington, DC: APA.
- Özyurek, A., Willems, R.M., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: Insights from event-related potentials. *Journal of Cognitive Neuroscience*, 19, 605–616. <https://repository.ubn.ru.nl/bitstream/handle/2066/56451/56451.pdf?sequence=1> (ver.15.04.2019).

- Ping, R.M., & Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychobiology*, *44*, 1277–1287.
- Saerbeck, M., Schut, T., Bartneck, C., & Janse, M.D. (2010). Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. *Proceedings of the SIGCHI conference on human factors in computing systems, ACM*, 1613–1622.
- Scorolli, C., Binkofski F., Buccino, G., Nicoletti, R., Riggio, L., & Borghi, A.M. (2011). Abstract and Concrete Sentences, Embodiment, and Languages. *Frontiers in Psychology*, *2*, 1–11.
- Shapiro, L.A. (2004). *The mind incarnate*. Cambridge, MA: MIT Press.
- Skipper, J.I., Goldin-Meadow, S., Nusbaum, H.C., & Small, S.L. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language*, *101*, 260–277.
- Sueyoshi, A., & Hardison, D.M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*, 661–699.
- 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*(1), 78–95.
- Valenzeno, L., Alibali, M.W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28*, 187–204.
- Vlaar, R., Verhagen, J., Oudgenoeg-Paz, O., & Leseman, P. (2017). Comparing L2 word learning through a tablet or real objects: what benefits learning most?. *Proceedings of the R4L workshop at HRI'17*. Vienna, Austria.
- Vogt, M., De Haas, C., De Jong, P., Baxter, P., & Kramer, E. (2017). Child-robot interactions for second language tutoring to pre-school children. *Front. Hum. Neuroscience*, *11*, 1–7. <https://doi.org/10.3389/fnhum.2017.00073> (ver. 15.04.2019).
- Wermelinger, S., Gampe, A., & Daum, M.M. (2017). Bilingual toddlers have advanced abilities to repair communication failure. *Journal of experimental child psychology*, *155*, 84–94.
- Westlund, K.J.M., Martinez, M., Archie, M., Das, M., & Breazeal, C. (2016). Effects of framing a robot as a social agent or as a machine on children's social behavior. *Proceedings of the 25th IEEE international symposium on robot and human interactive communication, IEEE*, 688–693.
- Wilson, B.G. (1996). *Designing constructivist learning environments*. Englewood Cliffs, NJ: Educational Technology Publications.
- Wu, Y.C., & Coulson, S. (2007). How iconic gestures enhance communication: An ERP study. *Brain & Language*, *101*, 234–245.
- Yow, W.Q., & Markman, E.M. (2011). Young bilingual children's heightened sensitivity to referential cues. *Journal of Cognition and Development*, *12*(1), 12–31.