

Cognizione sociale e linguaggio: il ruolo del processamento dei volti nello sviluppo della capacità linguistica

Giuliana Pulvirenti

Dipartimento di Scienze Cognitive, Pedagogiche, Psicologiche e degli Studi Culturali (COSPECS) – Università di Messina
giuliana.pulvirenti@unime.it

Abstract The present work defends the importance of developmental studies to shed light on the evolutionary roots of human language uniqueness and claims that language learning processes are multimodal in nature and socially gated. Recent studies have highlighted the role played by statistically based pattern recognition and extraction mechanisms, as well as that one of neuro-cognitive processes that link perceptual and motor abilities to the usage of social cues. Species-specific aspects of language may derive from both the anatomo-morphological development of sapiens' vocal tract and the interaction of domain-general sensory-motor and cognitive abilities. An analysis of the different strategies employed by infants to process faces during early linguistic learning stages appears to be particularly informative. Indeed, it is experimentally proved that children preferentially look at different regions of the face in different times of their early development. These changes may be functionally correlated to language learning, favouring imitative processes, presumably involved in the development of oro-articulatory abilities, gaze triangulation and sense convergence. Therefore, a focus on the ontogenetic mechanisms involved in language acquisition could act as a magnifying lens on functions until now treated in a generic fashion and inform developmental and evolutionary theories of language.

Keywords: language acquisition, infancy, language ontogeny, face processing, audio-visual integration, perceptual narrowing

Received 1 September 2017; accepted 31 December 2017.

1. Nuove ipotesi sui meccanismi di acquisizione e sviluppo della capacità linguistica.

Le teorie sull'acquisizione del linguaggio (ontogenesi) sono state per parecchio tempo dominate dalla contrapposizione dei concetti, dalla lunga tradizione filosofica, di "innato" e "appreso", quelle sulla sua origine invece (filogenesi) dalla contrapposizione di tesi volte a sottolinearne l'aspetto di alterità qualitativa rispetto agli altri sistemi di comunicazione animale e di altre maggiormente prone ad evidenziarne gli aspetti di continuità.

Il più illustre esponente della corrente innatista moderna è il linguista e filosofo Noam Chomsky, il quale ha sviluppato le proprie posizioni in diretta risposta alle teorie comportamentiste che prevedevano un meccanismo generale di acquisizione della capacità linguistica completamente basato sui meccanismi di condizionamento classico. In particolar modo Chomsky ha individuato nella sintassi (sistema ricorsivo-computazionale) il nucleo della specie-specificità linguistica umana ed ha ipotizzato l'esistenza di un modulo neurale innato dedicato esclusivamente al processamento degli input acustici e degli output motori di tipo linguistico.

In questo contesto, l'analisi linguistico-formale del linguaggio si è tradizionalmente impegnata nell'isolamento delle sue proprietà e componenti costitutive (arbitrarietà, creatività, semantica, sintassi, combinatorietà, capacità d'astrazione) sottolineandone l'intrinseca complessità. Ciò ha condotto all'assunto che un prodotto cognitivo così elaborato e strutturato debba necessariamente essere innato, predisponibile al bambino, per potersi sviluppare compiutamente. In effetti, il bambino, per penetrare il codice linguistico e acquisire lessico e grammatica, deve fronteggiare una serie di problemi, come ad esempio segmentare il flusso fonatorio continuo, apprendere l'utilizzo corretto delle regole grammaticali e sintattiche, assegnare etichette linguistiche a oggetti immediatamente percepibili e non, che sarebbe un errore giudicare banali (GUAISTI 2007). Nonostante tali difficoltà, i bambini non solo riescono in tutti questi compiti, ma lo fanno anche abbastanza velocemente. Dalla nascita all'età di circa 4 mesi il livello cognitivo dei bambini appare simile a quello degli esemplari adulti di altre specie di primati, ma successivamente alla discesa della laringe e all'allungamento della cavità orale (intorno ai 4 mesi d'età) inizia lo sviluppo della vera e propria capacità linguistica. Fra i 6 e i 18 mesi il bambino entra nel periodo della cosiddetta "lallazione" (vocale, manuale nei soggetti sordi) i cui prodotti, seppur lontani dall'essere vere e proprie parole, presentano un'organizzazione sillabica e l'uso specifico dei sottoinsiemi di suoni impiegati nella lingua nativa. A 12 mesi comincia la produzione di proto-parole fino ad arrivare alle parole vere e proprie riferite a oggetti precisi ed utilizzate in contesti diversi allo scopo di comunicare esplicitamente qualcosa a qualcuno. Dopo il secondo anno d'età si assiste all'esplosione del vocabolario, cioè alla veloce e improvvisa acquisizione di forme lessicali e sintattiche¹.

La rapida progressione di questi eventi è spesso fonte di sbalordimento negli adulti e viene pertanto solitamente portata come dato a favore di concezioni innatiste e discontinuiste (HAUSER *et al.* 2002).

L'obiettivo del presente articolo è quello di vagliare la plausibilità delle tesi, storicamente influenti e ancora centrali all'interno del dibattito accademico, sopra schematicamente delineate, alla luce di nuovi dati sperimentali.

Recenti studi cerebrali e comportamentali hanno infatti evidenziato le potenzialità che meccanismi di riconoscimento ed estrazione di pattern statistici dall'input, unitamente all'attività svolta da sistemi neuro-cognitivi che collegano le abilità percettivo-motorie all'utilizzo di indizi sociali provenienti dal contesto socio-ecologico, hanno nello spiegare questi eventi. Un focus sui processi ontogenetici legati agli aspetti percettivi e motori connessi allo sviluppo delle capacità linguistiche sembra pertanto particolarmente promettente, agendo come una lente di ingrandimento su funzioni sinora trattate in maniera generica e "distante" e fornendo

¹ Studi longitudinali e analisi quantitative della produzione lessicale hanno dimostrato tuttavia che solo 1 bambino su 5 attraversa una fase di vera e propria esplosione del vocabolario (BEHME & DEACON 2008).

dati che permetterebbero di testare la validità di teorie evolutive ed evolucionistiche ancora molto speculative.

Da un lato è infatti evidente il notevole grado di affinamento dell'analisi, che si muove a un livello più basso ma anche biologicamente più fondato. In altri termini, l'ambizione di rispondere a interrogativi del tipo tutto-o-nulla viene abbandonata in favore di una parcellizzazione operativa di quelle che un tempo rappresentavano le "grandi questioni" connesse al tentativo di svelare il mistero dell'unicità umana. Dall'altro, poiché l'indagine si concentra sulla dimensione diacronica dei cambiamenti e delle interazioni che avvengono fra i vari sistemi percettivi e sensomotori nell'organismo in fase di sviluppo, una declinazione in senso comparativo di tale approccio permetterebbe di: ottenere una visione più lucida e fine delle differenze cognitivo-comportamentali realmente esistenti fra le specie odierne di primati antropomorfi, individuandone le traiettorie ontogenetiche e quindi i punti di continuità e divergenza, chiarire il rapporto che esiste fra le strutture anatomico-morfologiche e le funzioni che su di (e grazie a) esse si innestano. Sebbene le famose quattro cause (ontogenesi, meccanismi di causazione, filogenesi e valore adattativo) formulate da Nikolaas Tinbergen (1963) per facilitare l'analisi evolucionistica del comportamento animale siano concettualmente distinte e vengano spesso indagate in modo isolato, è fondamentale prendere coscienza che esse si informano reciprocamente. Detto altrimenti, lo studio dei meccanismi biologici che determinano l'esibizione di un comportamento e quello del suo progressivo sviluppo nell'organismo possono fornire elementi importanti per ricostruirne la storia evolucionistica. È infatti possibile che tratti cognitivo-comportamentali, anche complessi, varino significativamente in specie strettamente affini a causa di piccoli cambiamenti nelle traiettorie di sviluppo ontogenetico (ROSATI *et al.* 2014).

Questo genere di ricerca condotta su soggetti ancora giovani ha già prodotto degli sviluppi interessanti creando uno spazio sperimentale tramite cui testare il ruolo che l'esperienza gioca nell'acquisizione della capacità linguistica e verificare quali processi percettivo-motori e neurocognitivi contribuiscono al suo pieno attecchimento (KUHL 2010).

In tale contesto alcune tecniche di indagine non invasiva², che consentono di registrare quali aree del cervello intervengono, e quando, nell'elaborazione di un dato stimolo, appaiono di grande validità in quanto possono essere impiegate senza particolari problemi etici in protocolli che prevedono il coinvolgimento di bambini in fasi ancora precoci di sviluppo.

Un altro modo per indagare l'ontogenesi tipica ed atipica delle abilità percettive, sensomotorie e cognitive è rappresentato dall'utilizzo di dispositivi di tracciamento dei movimenti oculari (*eye tracking*) durante *tasks* di tipo percettivo (ad esempio di processamento di scene, immagini e volti). Le principali misure rilevate in questi casi includono la localizzazione e la sequenza spaziale delle fissazioni, da cui è possibile ricavare il focus d'attenzione dei soggetti esaminati, e la durata di fissazione, che risulta maggiore in relazione alla complessità del compito o alla necessità di reperire informazioni più dettagliate (KARATEKIN 2007). L'utilizzo di queste metodiche ha prodotto risultati che hanno già messo in discussione l'argomento principe solitamente portato a favore delle tesi nativiste: la *povertà dello stimolo* (BEHME & DEACON 2008; GERVAIN & MEHLER 2010). In effetti, se sottoposto ad attento scrutinio, l'argomento in questione appare basato su una serie di assunzioni assurde

² Elettroencefalografia (EEG) e potenziali evento-correlati (ERPs), Magnetoencefalografia (MEG), Risonanza magnetica funzionale (fMRI) e Spettroscopia a raggi infrarossi (NIRS).

intuitivamente al rango di dati certi ma scientificamente deboli in quanto non ricavate da analisi empiriche ripetute e sperimentalmente controllate. Nella fattispecie: siamo sicuri che l'input a disposizione del bambino sia davvero così povero? Perché dare per assodato che l'unico input utile sia quello derivante dall'insegnamento diretto ed esplicito della lingua e non considerare anche gli indizi non verbali e indiretti come possibile fonte di preziose informazioni? L'apprendimento della lingua avviene sicuramente in modo rapido ma, cosa spesso sottovalutata, non *ex abrupto* bensì progressivamente nel corso di mesi.

Alla luce di queste considerazioni, il presente lavoro si prefigge lo scopo di valutare il contributo che questo moderno filone di ricerche è in grado di fornire alla comprensione dei vari meccanismi che di fatto determinano la progressiva costituzione di questo tratto tipicamente umano. Tale operazione porta, a nostro avviso, a mitigare la nota contrapposizione fra i concetti di "innato" e "appreso", applicabili in questa prospettiva ad elementi più basilari, dal punto di vista biologico, e in un certo senso "microscopici", del linguaggio. A tal fine, il resto dell'articolo sarà dedicato all'esame dei risultati di alcuni studi ritenuti particolarmente significativi condotti su bambini in fasi precoci di sviluppo volti ad indagare le abilità percettive di discriminazione e processamento di informazioni acustiche di natura prettamente linguistica, per poi passare all'analisi dei dati relativi alle capacità di processamento visivo. La proposta teorica qui abbracciata tende infatti a ravvisare la componente innata sia nelle capacità e nei bias percettivi dei bambini che consentono la percezione dei suoni linguistici coinvolgendo, sin dalle prime fasi di vita, più di un canale sensoriale (vista, oltre che udito) in un contesto di interazione sociale peculiare nel panorama animale (MATSUZAWA 2006)

La specie-specificità della funzione linguistica, nell'ipotesi qui presa in esame, potrebbe costituirsi gradualmente durante lo sviluppo dell'organismo tramite processi di crescita e apprendimento multimodali e integrati, dipendendo dall'interazione dei seguenti fattori:

- sviluppo della configurazione morfo-anatomica specie-specifica del sapiens caratterizzata da: «1) un tratto vocale ricurvo a due canne con una proporzione 1:1 tra la canna verticale e orizzontale; 2) un insieme di muscoli orofacciali che non ostacolano la vocalizzazione fine; 3) una corteccia uditiva iperspecializzata (*Vocal Area*) e un neuroprocessore evolutivo categoriale fondato sul linguaggio articolato, cioè l'area di Broca» (PENNISI 2014: 164).
- integrazione di abilità percettivo-motorie e cognitive di base e *domain general* (percezione categoriale e cross-modale, apprendimento statistico), bias percettivi.
- modalità di interazione col contesto socio-ecologico tipico della specie durante l'ontogenesi, processi imitativi.

1.1 Meccanismi coinvolti nel processamento degli input di natura linguistica.

Una grande mole di dati sperimentali attesta la presenza nei bambini di precoci abilità di discriminazione e processamento linguistico basate sul rilevamento di informazioni di natura statistica presenti nell'input e guidate da bias percettivi generali.

Già a poche ore dalla nascita i bambini sembrano essere istintivamente portati a preferire suoni linguistici a suoni non linguistici (VOULOMANOUS & WERKER 2007a). Sebbene vada ancora chiarita la natura di questo genere di bias (predisposizione innata oppure forme di esperienza avvenute in ambiente intrauterino durante la fase prenatale?) da molti studi è possibile evincere il ruolo che fattori di

natura esperienziale giocano in relazione alle preferenze discriminatorie esibite dal bambino subito dopo la nascita.

I neonati infatti prediligono il suono della voce materna e della lingua nativa a quello di altre donne o a lingue straniere (GERVAIN & MEHLER 2010). Da uno studio condotto da Vouloumanos *et al.* (2010) è emerso che i neonati non mostrano alcuna preferenza fra suoni linguistici e vocalizzazioni di scimmie reso (*Macaca mulatta*), sebbene preferiscano entrambi questi suoni a suoni sintetizzati, mentre soggetti di soli 3 mesi d'età dirigono maggiormente la propria attenzione verso i suoni linguistici.

I neonati sono inoltre in grado di discriminare coppie di lingue “filtrate”, cioè contenenti solo informazioni prosodiche, e mai sentite prima, purché appartenenti a classi ritmiche diverse. Tale capacità (misurata tramite la tecnica della suzione non nutritiva, registrandone velocità e ritmo, parametri che crescono alla presentazione di stimoli percepiti dal soggetto come nuovi) non viene documentata nei confronti di alcune coppie di lingue come lo spagnolo e l'italiano o il tedesco e l'olandese (GUSTI 2007) appartenenti alla stessa classe ritmica³.

Sulla base di ciò è stato ipotizzato che i bambini riescano ad effettuare questo genere di distinzioni basandosi su indizi prosodici, come il ritmo e l'intonazione, il che spiegherebbe i risultati negativi ottenuti negli esperimenti di discriminazione di coppie di lingue caratterizzate da un'organizzazione ritmica molto simile. Le proprietà ritmiche della lingua potrebbero inoltre facilitare l'acquisizione di alcuni pattern morfosintattici (relazioni di accordo: flessioni nominali e verbali). Lingue appartenenti a differenti classi ritmiche hanno infatti proprietà morfosintattiche differenti.

Capacità discriminatorie vengono esibite anche durante la semplice visione di volti parlanti “silenziosi”. In uno studio condotto da Weikum *et al.* (2007), bimbi inglesi di 4, 6 e 8 mesi sono stati sottoposti alla visione di filmati a cui era stato tolto l'audio ed in cui era possibile osservare tre diverse donne bilingue (francese-inglese) che declamavano delle frasi tratte dal libro *Il piccolo principe*. Dopo una prima fase di abituaione al compito in lingua inglese, i soggetti sono stati divisi in due gruppi, uno di controllo ed uno sperimentale. Mentre nel gruppo di controllo le donne articolavano le frasi esclusivamente in lingua inglese, in quello sperimentale è stata utilizzata la lingua francese. Dai risultati è emerso che, al contrario dei bambini di 8 mesi, i bambini appartenenti alle fasce d'età inferiori guardano più a lungo la scena, dimostrando così di percepire il cambiamento e dunque esibendo capacità di discriminazione fra le due lingue impiegate nelle diverse fasi dell'esperimento (MAURER & WERKER 2013). Questi dati sono importanti perché oltre ad attestare la presenza di una forma di specializzazione percettiva a forte base esperienziale, indicano che la percezione della lingua avviene sin dall'inizio in un contesto di tipo cross-modale, tramite l'integrazione di indizi provenienti da vari domini percettivi (nella fattispecie visivo e uditivo).

³ Secondo la teoria linguistica le lingue sono approssimativamente classificabili in tre classi ritmiche: lingue a isocronismo accentuale basate sul ricorrere regolare degli accenti (inglese, olandese, russo, svedese), lingue a isocronismo sillabico basato sul ricorrere regolare delle sillabe (catalano, francese, greco, italiano, spagnolo), lingue a isocronismo moraic con ritmo basato su un'unità subsillabica detta “mora” (giapponese, tamil, turco). RAMUS NESPOR e MEHLER (1999) hanno dimostrato che tali distinzioni dipendono dalla differente organizzazione di consonanti e vocali, in particolare dalla percentuale di spazio vocalico e dalla diversa distribuzione degli intervalli consonantici.

I bambini nati da pochi giorni possiedono anche la capacità di discernere, al contrario dei soggetti adulti, una gamma molto vasta di “contrastisti di suoni” (si tratta di una forma di percezione categoriale) come ad esempio l/r o b/o (EIMAS *et al.* 1971). Tale abilità è “universale” alla nascita ma durante il primo anno di vita declina per i contrasti di suoni stranieri mentre viene mantenuta e perfino accentuata per quelli della lingua nativa (MELTZOFF *et al.* 2009). Tramite studi longitudinali è stato visto che questo fenomeno di specializzazione percettiva, attraverso cui il bambino si sintonizza con l’ambiente socio-ecologico di sviluppo (*perceptual narrowing*, talvolta descritto anche come *perceptual attunement*), appare in forte correlazione col grado di abilità linguistica a distanza di mesi o anni (KUHL 2006).

Estrarre parole dal flusso continuo del parlato costituisce un’ulteriore sfida per il bambino se si tiene conto dei seguenti fattori: durante la fonazione articolata fra una parola e l’altra non ci sono pause, le caratteristiche fonetiche di un determinato suono sono influenzate da quello che lo precede e da quello che lo segue (coarticolazione), esistono parole incassate in altre parole, i neonati sono privi di lessico. I bambini riescono felicemente anche in questo compito, per cui capire come facciano alla luce delle difficoltà sopraelencate rappresenta un punto di fondamentale importanza.

Uno studio di Saffran *et al.* (1996) ha dimostrato che i bambini di 8 mesi utilizzano meccanismi di apprendimento statistico (implicito, *domain general* e filogeneticamente antico) per assolvere a questo ed altri compiti di processamento dell’input acustico-linguistico tramite il rilevamento di regolarità e pattern distribuzionali (come ad esempio la probabilità transizionale, cioè la probabilità che ha una sillaba di precederne o seguirne un’altra) che, una volta appresi, forniscono uno schema per segmentare lo *speech stream* ed individuare e riconoscere le varie unità linguistiche in esso contenute (KUHL 2010). Ricerche successive hanno inoltre evidenziato come questo meccanismo venga utilizzato durante l’iniziale acquisizione delle forme lessicali principalmente in relazioni alle consonanti. L’estrazione di regolarità statistiche sembra quindi svolgere un ruolo notevole nel processo di progressiva acquisizione del linguaggio (GERVAIN & MEHLER 2010, BEHME & DEACON 2008). Esso tuttavia diventa problematico nel caso di alcune lingue, come il cinese, o del linguaggio semplificato diretto ai bambini (*motherese/child-directed speech*) in cui le parole sono spesso monosillabiche. In questi casi altre proprietà acustiche e fonologiche direttamente correlate ai confini lessicali possono essere utilizzate in modo complementare per assolvere la stessa funzione. Fra questi ad esempio ricordiamo i pattern accentuali e di segmentazione metrica, gli indizi prosodici ed i vincoli fonotattici (distribuzione tipicamente corretta o scorretta di fonemi all’interno di una parola o fra parole adiacenti in una data lingua). Shi *et al.* (1999) hanno dimostrato che i bambini, forse usando questi indizi, riescono a distinguere precocemente anche classi diverse di parole: i funtori, che hanno un’alta frequenza e segnalano la struttura morfosintattica della frase, e le parole di contenuto, caratterizzate da una frequenza più bassa, da proprietà acustiche salienti (accento, struttura multisillabica) e contenenti il significato lessicale. L’utilizzo di meccanismi di estrazione statistica di regolarità, eventualmente supportati dallo sfruttamento di *cues* prosodici, potrebbe giocare una parte fondamentale nell’acquisizione delle prime semplici strutture sintattiche.

La considerazione che sorge spontanea a questo punto è la seguente: dal momento che la realtà esterna contiene una grande varietà di informazioni dotate di pattern distribuzionali potenzialmente estraibili, devono esistere dei vincoli che permettano di dirigere tali meccanismi estrattivi preferenzialmente verso stimoli salienti in relazione alla nicchia socio-ecologica occupata dall’organismo. Quali sono dunque

gli indizi disponibili al bambino e che ruolo svolgono nel processo di acquisizione della capacità linguistica?

Nell'ipotesi qui analizzata lo sviluppo ed il progressivo affinamento di questi processi vengono parzialmente ma significativamente guidati dall'interazione con le figure accudenti di riferimento; sarebbero cioè modellati da dinamiche di tipo sociale. Secondo la *social gating hypothesis* (KUHL 2007) il contesto sociale potrebbe operare a diversi livelli per favorire i processi d'acquisizione linguistica incrementando: attenzione e arousal emotivo, risonanza percettivo-motoria e apprendimento dei pattern motori coinvolti nell'articolazione oro-facciale, possibilità di attribuzione e decodifica semantico-referenziale.

2. Cognizione sociale e linguaggio: la relazione fra *face scanning* e *speech processing*

In relazione all'argomento che stiamo qui brevemente esplorando, risulta estremamente interessante indagare il funzionamento e l'ontogenesi dei meccanismi coinvolti nel processamento dei volti ed il loro legame con quelli impiegati per il processamento dei suoni linguistici, in quanto essi potrebbero svolgere un ruolo centrale nell'acquisizione della capacità articolatoria e fonatoria mediante imitazione o *neural mirroring*.

Tradizionalmente volti e voci sono stati considerati domini separati e solo recentemente sono stati sviluppati paradigmi sperimentali volti a esplorare le abilità d'integrazione di stimoli acustici e visivi ed il ruolo che queste svolgono nello sviluppo del linguaggio. D'altronde, nella vita di tutti i giorni, i bambini esperiscono per lo più volti parlanti (vale a dire stimoli visivi in sincrono con stimoli uditivi) con cui spesso e volentieri interagiscono in vario modo, soprattutto se impersonati da figure accudenti di riferimento.

Come per le unità fonetiche, anche i meccanismi impiegati nel riconoscimento e nella discriminazione dei volti partono da una base piuttosto generica per poi specializzarsi col tempo e l'esperienza, seguendo un andamento evolutivo atto a favorire la sintonizzazione del bambino con l'ambiente sociale di sviluppo (PASCALIS *et al.* 2014, MAURER & WERKER 2013).

Pascalis *et al.* (2002) hanno comparato la capacità di processare uno stimolo *face-like* nuovo, di scimmie (*Macaca fascicularis*) o umani, rispetto a uno familiare di un campione di bambini di 6, 9, 11 mesi e adulti. I risultati di questo studio rivelano che (come accade nel caso dei contrasti fonetici, e più in generale dei suoni *speech-like*) a differenza dei bambini più grandi e dei soggetti adulti, i bimbi di soli 6 mesi mostrano una preferenza per lo stimolo visivo *face-like* nuovo anche se questo non è specie-specifico. Questo fenomeno appare collegato al cosiddetto "other race effect" (ORE), riportato nella popolazione adulta, che fa riferimento alla difficoltà di discriminare in maniera accurata i volti di persone appartenenti a gruppi etnici differenti dal proprio. Tale effetto è stato rilevato anche in studi condotti su bambini di 3 mesi, che esibiscono una predilizione per volti appartenenti al proprio gruppo etnico, ed un "gender bias" correlato presumibilmente al genere della figura accudente primaria, solitamente di sesso femminile, ma in caso di padri single anche maschile (PASCALIS & KELLY 2009). Da uno studio di Kelly *et al.* (2005) volto a valutare, tramite registrazione dei movimenti oculari, la preferenza visiva esibita da un campione di bambini caucasici in due diverse fasce d'età (neonati e 3 mesi) nei confronti di coppie di volti di etnia diversa (uno appartenente al proprio gruppo etnico, l'altro no), è emerso che mentre i neonati non mostrano alcuna preferenza

spontanea verso volti appartenenti al loro o ad un altro gruppo etnico, i bambini di 3 mesi cominciano a mostrarla.

Dunque, il fenomeno del *perceptual narrowing* a cui abbiamo precedentemente fatto riferimento in relazione alle abilità discriminatorie esibite nei confronti dei suoni linguistici, interessa anche il dominio visivo del processamento dei volti. Esso inoltre è stato recentemente osservato anche in ambito cross-modale (*intersensory matching*).

Pons *et al.* (2009) hanno testato la capacità di bambini inglesi e spagnoli, di 6 e 11 mesi, di accoppiare informazioni visive e uditive durante la percezione del contrasto di suono /ba/-/va/ (che non esiste nella lingua spagnola perché le consonanti b e v sono omofone). L'esperimento consisteva in sei fasi: le prime due di presentazione di una coppia di video, posti fianco a fianco, in cui una donna bilingue articola da un lato la sillaba /ba/, dall'altro la sillaba /va/, le altre quattro di test volte a indagare se l'ascolto di una delle sillabe durante una prima fase di familiarizzazione uditiva produceva un aumento del tempo di fissazione sul volto che articolava correttamente la stessa sillaba in una successiva fase di test di preferenza visiva. Ebbene, all'età di 6 mesi sia i bambini inglesi che quelli spagnoli hanno rivolto preferenzialmente il proprio sguardo sul volto che articolava lo stimolo linguistico corrispondente, tuttavia questo comportamento non è stato riscontrato fra i bambini di 11 mesi di origine spagnola.

Lewcovicz & Ghazanfar (2006) hanno indagato la presenza di *intersensory matching* in gruppi di bambini di 4, 6, 8 e 10 mesi utilizzando una procedura di preferenza visiva accoppiata per misurare il tempo di fissazione su due volti dinamici di scimmie (*Macaca mulatta*) vocalizzanti. Ogni bambino è stato sottoposto a due *silent test trials* (coppia video, posti a fianco a fianco, privi di audio, in cui una stessa scimmia articola due differenti vocalizzazioni) e due *in-sound test trials* (due video, come nella prima fase, che però stavolta comprendevano anche un input sonoro presentato in sincrono con entrambi i video ma correttamente accoppiato solo ad uno di essi). Anche in questo caso, i risultati confermano che soltanto i due gruppi di soggetti più giovani sono in grado di superare il compito, mostrando capacità di accoppiamento crossmodale nei confronti di stimoli non specie-specifici.

Altri studi recenti impegnati nell'analisi dei meccanismi sensori-motori e cognitivi grazie ai quali il bambino impara a produrre ed articolare in maniera appropriate le unità fonetiche della propria lingua, hanno dimostrato che il rilevamento di regolarità statistiche⁴ presenti nell'input acustico è influenzato in maniera significativa dal processamento di stimoli visivi correlati, come ad esempio l'osservazione della fonte da cui proviene lo stimolo linguistico (KUHL 2010). Da un esperimento di discriminazione fonetica in presenza di stimoli audiovisivi, condotto da Teinonen *et al.* (2008) su un campione di quarantotto bambini di 6 mesi, è emerso che l'osservazione sincronica di indizi visivi (presentazione su schermo di un volto che articola sillabe) congruenti con l'input sonoro (sillabe /ba/ e /da/) consente ai bambini di superare il *task*. Viceversa, i soggetti sperimentali a cui era stato fornito un unico tipo di indizio visivo durante la presentazione dei due differenti input sonori fallivano il compito di discriminazione.

⁴ La percezione fonetica in fase di sviluppo, ed in particolare fra i 6 e i 12 mesi, dipende ed è influenzata dai pattern distribuzionali di frequenza di un dato fonema, cioè dalla probabilità più o meno elevata di incontrare un determinato suono nel contesto linguistico d'appartenenza (KUHL 2010).

Allo stesso modo l'interazione verbale potrebbe svolgere un ruolo importante nei processi di discriminazione visiva postnatale. Sai *et al.* (2005) hanno indagato questa ipotesi su un campione di bambini appena nati, incoraggiando soltanto metà delle madri ad interagire verbalmente con i propri piccoli e precludendo ogni forma di interazione linguistica alla restante metà. In una successiva fase di test, avvenuta dopo poche ore, i bambini sono stati sottoposti alla visione della madre e di un'altra donna ma soltanto i bambini che avevano potuto usufruire dell'interazione verbale materna sono stati in grado di discriminare il volto.

La possibilità di accedere a indizi sincronici appartenenti a più modalità sensoriali sembra dunque aumentare la salienza degli stimoli presentati facilitando così il riconoscimento dei volti.

Ai fini della presente discussione, appare di primaria importanza esplorare quali siano le aree del volto su cui il bambino focalizza la propria attenzione durante il processamento dei suoni linguistici e verificare la presenza di cambiamenti nella direzione dello sguardo connessi alle varie fasi del suo sviluppo percettivo-motorio. A partire dal sesto mese d'età il bambino entra infatti nella fase della cosiddetta lallazione, periodo in cui tenta, sebbene goffamente, di riprodurre i suoni linguistici di volta in volta percepiti. La competenza motoria del bambino è ancora molto limitata, ma, proprio per questo motivo, l'eventuale osservazione dei movimenti che interessano la cavità orale potrebbe svolgere un ruolo centrale nei processi di imitazione oro-articolatoria in funzione della progressiva acquisizione della capacità di produzione linguistica.

In uno studio del 2012, Lewkowicz e Hansen-tift hanno registrato i movimenti oculari esibiti da bambini inglesi appartenenti a diverse fasce d'età (4, 6, 8, 10 e 12 mesi) e adulti durante l'osservazione di un video in cui una donna produce un monologo in lingua straniera. I risultati mostrano che nei bambini di età compresa fra i 6 e i 12 mesi, sottoposti all'osservazione di volti che articolano suoni linguistici, si verifica un cambiamento nell'orientamento preferenziale dello sguardo verso l'area della bocca. Dai 12 mesi in poi lo sguardo viene ridiretto verso l'area degli occhi se l'input linguistico è familiare (lingua madre), ma non in caso contrario (lingua straniera).

Secondo l'*intersensory redundancy hypothesis* proposta da questi ricercatori, i bambini, dopo i 6 mesi, dirigono lo sguardo maggiormente verso la bocca soprattutto durante la presentazione di stimoli audio-visivi fra loro congruenti per meglio apprendere la lingua. Una volta acquisito un livello di competenza sufficiente il bambino può nuovamente spostare lo sguardo verso l'area degli occhi per avere accesso a indizi di natura più specificamente sociale.

Da altri studi è possibile ricavare l'importanza che l'aspetto interattivo riveste nei processi di apprendimento. Patricia Kuhl (2003) ha testato le capacità discriminative di un campione di bambini americani di 9 mesi (una fase in cui i sistemi percettivi deputati al processamento degli input linguistici sono già "accordati" con le proprietà statistico-fonetiche della lingua nativa) esposti a input linguistici stranieri (mandarino) in condizioni sperimentali diverse. I bambini che hanno avuto la possibilità di ascoltare un parlante straniero in carne ed ossa hanno poi riportato risultati di gran lunga migliori rispetto al gruppo di controllo, composto da soggetti non esposti ai suoni stranieri; al contrario, i soggetti sottoposti allo stesso genere di stimoli tramite apparati televisivi o registrazioni non hanno mostrato segni d'apprendimento e la loro performance è risultata identica a quella del gruppo di controllo. Questo esperimento, oltre a portare robusti dati a favore della plasticità delle aree deputate al processamento degli input linguistici (se non fosse così

d'altronde non potremmo imparare una seconda lingua da adulti), mette in luce l'impatto che l'interazione sociale dal vivo ha su queste dinamiche.

Per quali motivi essa appare tanto vitale per l'apprendimento della lingua?

Sebbene non sia al momento possibile avanzare tesi univoche circa il suo ruolo è lecito ipotizzare che essa abbia delle proprietà tali da catturare l'attenzione del bambino incanalandone le risorse computazionali verso stimoli salienti, oltre che una funzione informativa, consentendo l'accesso a informazioni che altrimenti non verrebbero rilevate e processate (equivalenza cross-modale tra percezione e azione, triangolazione, referenzialità).

In relazione a quest'ultimo punto mi sembra opportuno concludere questa breve rassegna con un accenno alle ricerche condotte da Andrew A. Meltzoff, volte ad indagare i meccanismi neurali che sottostanno alle esperienze corporee condivise. Questi infatti potrebbero costituire la base concreta del cosiddetto senso di "intersoggettività" (il sentimento di comprensione e comunicazione reciproca fra il sé e l'altro) spesso invocato nell'ambito della filosofia della mente, e svolgere un ruolo fondamentale nei processi di apprendimento a carattere imitativo. In un esperimento (MARSHALL *et al.* 2011) condotto mediante l'utilizzo di tecniche di neuroimmagine (Elettroencefalografia/EEG e analisi delle onde mu), un campione di trentacinque bambini di 14 mesi (esaminati singolarmente) è stato coinvolto in un *turn-taking game* in cui sia i soggetti (*execution condition*) che lo sperimentatore (*observation condition*) dovevano a turno agire su un oggetto, utilizzando di volta in volta diversi effettori (piede, mano). Le misure di attivazione neurofisiologica registrate nelle varie fasi dell'esperimento hanno evidenziato la presenza di pattern somatotopici di attivazione corticale sia nella condizione di osservazione che in quella di esecuzione dell'azione. Una sorta di "ponte" corporeo-neurale fra la percezione e l'azione, in grado di rendere "commensurabili" due entità distinte come il sé e l'altro.

3. Conclusioni

I nuovi approcci delle moderne scienze cognitive, insieme ai dati da essi prodotti, sembrano imporre una revisione, perlomeno parziale, dei modelli teorici sinora proposti per spiegare la specie-specificità della capacità linguistica. Essi mettono in dubbio sia l'ipotesi chomskiana, che concepisce questo tratto come innato attribuendone il funzionamento ad un modulo neurale deputato esclusivamente al processamento di input e output linguistici, che quella basata sul modello behaviorista classico di apprendimento. Al contrario, il processo di acquisizione del linguaggio nelle primissime fasi di vita sembra dipendere sia da meccanismi di apprendimento *domain general*, presenti anche in altre specie animali (es.: percezione categoriale, *statistical learning*) che dal funzionamento integrato di diversi sistemi percettivi che si sviluppano influenzandosi reciprocamente. Le capacità percettive del bambino appaiono infatti interconnesse sin dalla nascita, come dimostrato dai fenomeni di *intersensory matching* e *perceptual narrowing* cross-modale. L'intero processo appare guidato da bias percettivi che indirizzano le risorse attentive del bambino, coinvolgendo più di una modalità sensoriale, verso stimoli socialmente rilevanti come volti e suoni linguistici. Questi domini, appartenenti all'ambito della cognizione sociale, sono soggetti a dinamiche evolutive simili, a collo di bottiglia. Durante i primi mesi di vita infatti i bias attenzionali e le capacità discriminatorie non riguardano esclusivamente stimoli specie-specifici. Tuttavia col passare del tempo e tramite la massiccia esposizione agli input presenti

nell'ambiente di sviluppo si assiste a un fenomeno di specializzazione percettiva (*perceptual narrowing*) grazie al quale il bambino si sintonizza col mondo sociale presente intorno a lui, parzialmente perdendo le vaste e più indifferenziate abilità discriminatorie di cui è inizialmente dotato. All'interno di questo articolato scenario l'interazione sociale sembra essere un fattore determinante affinché il processo di acquisizione linguistica vada a buon fine. Chiarire in che modo e perché sia così fondamentale rappresenta una sfida per le moderne scienze cognitive del linguaggio. Alla luce di questi dati si ritiene che un approccio comparativo applicato allo studio dei meccanismi di causazione e all'ontogenesi di questa capacità esclusivamente umana possa illuminare le ragioni della sua presenza nella nostra specie.

Bibliografia

BEHME Christina & DEACON Hélèn S (2008), «Language learning in infancy: does the empirical evidence support a domain specific language acquisition device?», in *Philosophical Psychology*, October 2008, vol. 21, n. 5, pp. 641-671.

EIMAS Peter D *et al.* (1971), «Speech perception in infants», in *Science*, vol. 171, pp. 303-306.

FALZONE Alessandra (2014), «Structural constraints on language», in *Reti, saperi, linguaggi*, vol. 1(2), pp. 247-266.

FALZONE Alessandra & CARDELLA Valentina (2014), «Per una natura linguistica della mente umana: la cognizione sociale alla prova delle specie-specificità verbale», in *RIFL/SFL* 2014, pp. 164-178.

GERVAIN Judit & WERKER Janet F (2008), «How infant speech perception contributes to language acquisition», in *Language and linguistic compass*, vol. 2/6, pp.1149-1170.

GERVAIN Judit & MEHLER Jacques (2010), «Speech perception and language acquisition in the first year of life», *Annu. Rev. Psychol.*, vol.61, pp. 191-218.

GUASTI Maria Teresa (2007), *L'acquisizione del linguaggio: un'introduzione*. Raffaello Cortina, Milano.

HAUSER Marc D *et al.* (2002), «The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?», in *Science*, 22 Nov. 2002, vol. 298, Issue 5598, pp. 1569-1579.

KARATEKIN Canan (2007), «Eye tracking studies of normal and atypical development», in *Developmental Review*, vol. 27, pp. 283-348.

KELLY David J *et al.* (2005), «Three-month-olds, but not newborns, prefer own-race faces» *Developmental Science*, November 2005, vol. 8(6), pp. F31-F36.

KUHL Patricia K *et al.* (2003), «Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning», *PNAS*, 100(15), pp. 9096-9101.

KUHL Patricia K *et al.* (2006), «Infants show a facilitation effect for native language phonetic perception between 6 and 12 months», *Developmental Science*, vol. 9, pp. F13-F21.

KUHL Patricia K (2007), «Is speech learning ‘gated’ by the social brain?» *Developmental Science* vol. 10(1), pp 110-120.

KUHL Patricia K (2010), «Brain mechanisms in early language acquisition», *Neuron*, September 2010, vol. 67(5), pp. 713-727.

KUHL Patricia K & MELTZOFF Andrew N (1982), «The bimodal perception of speech in infancy», *Science*, vol. 218, pp. 1138-1141.

LEWKOWICZ David J & GHAZANFAR Asif A (2006), «The decline of cross-species intersensory perception in human infants», *PNAS*, vol. 103(17), pp. 6771-6774.

LEWKOWICZ David J & HANSEN-TIFT Amy M (2012), «Infants deploy selective attention to the mouth of a talking face when learning speech», *PNAS*, vol. 109(5), pp. 1431-1436.

MARSHALL Peter J & MELTZOFF Andrew N (2011), «Neural mirroring systems: Exploring the EEG mu rhythm in human infancy», *Developmental Cognitive Neuroscience*, vol. 1, pp. 110-123.

MATSUZAWA Tetsuro (2006), *Evolutionary origins of the human mother-infant relationship*, in MATSUZAWA Tetsuro, TOMONAGA Masaki, TANAGA Masayuki (2006), *Cognitive development in chimpanzees*, Springer-Verlag, Tokyo.

MAURER Daphne & WERKER Janet F (2013), «Perceptual narrowing during infancy: a comparison of language and faces», *Developmental Psychobiology*, vol. 56, pp. 154-178.

MELTZOFF Andrew N *et al.* (2009), «Foundations for a new science of learning», *Science* July 2009, vol. 17, 325(5938), pp. 284-288.

NESPOR Marina *et al.* (2003), «On the different roles of vowels and consonants in speech processing and language acquisition», in *Lingue e Linguaggio*, vol. 2, pp. 223-229.

PASCALIS Olivier *et al.* (2002), «Is Face Processing Species-Specific During the First Year of Life?», *Science*, vol 296, pp. 1321-1322.

PASCALIS Olivier & KELLY David (2009), «The origins of face processing in humans: phylogeny and ontogeny», *Perspectives on Psychological Science*, vol. 4(2), pp. 200-209.

PASCALIS Olivier *et al.* (2014), «On the links among face processing, language processing and narrowing during development», *Child Development Perspectives*, vol. 8(2), pp. 65-70.

PENNISI Antonio (2014), *L'errore di Platone: biopolitica, linguaggio e diritti civili in tempo di crisi*. Bologna: Il Mulino.

PENNISI Antonio & FALZONE Alessandra (2015), «Nuovi approcci epistemologici ad una filosofia naturalistica del linguaggio», *RIFL/SFL*, vol. 1, pp. 92-105.

PONS Ferran *et al.* (2009), «Narrowing of intersensory speech perception in infancy», *PNAS*, vol. 106, pp. 10598-10602.

RAMUS Franck, NESPOR Marina, MEHLER Jacques (1999), «Correlates of linguistic rhythm in the speech signal», in *Cognition*, vol. 72, pp. 265-292.

ROSATI Alexandra G *et al.* (2014), «Comparative developmental psychology: how is human cognitive development unique?», *Evolutionary Psychology*, 12(2), pp. 448-473.

SAFFRAN Jenny R (1996), «Statistical learning by 8 months old infants», *Science*, New series, vol. 274(5294), pp. 1926-1928.

SAI FZ (2005), «The Role of the Mother's Voice in Developing Mother's Face Preference: Evidence for Intermodal Perception at Birth», *Inf. Child Dev.* vol. 14, pp. 29-50.

SHI Rushen *et al.* (1999), «Newborn infants' sensitivity to perceptual cues to lexical and grammatical words», *Cognition*, vol. 72B, pp. 11-21.

TEINONEN Tuomas *et al.* (2008), «Visual speech contributes to phonetic learning in 6 months old infants», *Cognition*, vol. 108, pp. 850-855.

TINBERGEN Nikolaas (1963), «On aims and method of ethology», *Zeitschrift für Tierpsychologie*, vol. 20, pp. 410-433.

VALLAURI Edoardo Lombardi (2004), «The relation between mind and language: The Innateness Hypothesis and the Poverty of the Stimulus», *The Linguistic Review*, vol. 21, pp. 345-387.

VOULOUMANOS Athena & WERKER Janet F (2007a), «Listening to language at birth: evidence for a bias for speech in neonates», *Developmental Science*, vol.10:2 (2007), pp 159-171.

VOULOUMANOS Athena *et al.* (2010), «The Tuning of Human Neonates' Preference for Speech», *Child Development*, March/April 2010, vol. 81, n.2, pp. 517-527.

WEIKUM Whitney M *et al.* (2005), «Visual Language Discrimination in Infancy», *Science*, 25 May 2007 vol. 316, p. 1159.