

Michael Tomasello

Le origini culturali
della cognizione
umana

COMUNICAZIONE LINGUISTICA E RAPPRESENTAZIONE SIMBOLICA

Ogni particolare notazione mette in rilievo un particolare punto di vista.

Ludwig Wittgenstein

Nelle discussioni sulla filogenesi della cognizione umana il *linguaggio* è spesso citato come una causa dell'unicità cognitiva dell'uomo. Ma considerare il linguaggio una causa evolutiva della cognizione umana è un po' come considerare il denaro una causa evolutiva dell'attività economica umana. Non vi è dubbio che acquisire e usare un linguaggio naturale influenzi, e perfino trasformi, la natura della cognizione umana, proprio come il denaro trasforma la natura dell'attività economica umana. Ma il linguaggio non nasce dal nulla. Non è piombato sulla Terra come un asteroide vagabondo, né, a dispetto di quel che credono studiosi contemporanei come Chomsky [1980], è il frutto di una bizzarra mutazione genetica indipendente dagli altri aspetti della cognizione e della vita sociale dell'uomo. Proprio come il denaro è una istituzione sociale di natura simbolica che è nata storicamente da attività economiche preesistenti, il linguaggio naturale è una istituzione sociale di natura simbolica che è nata storicamente da attività sociocomunicative preesistenti.

Affinché i bambini apprendano a usare i simboli linguistici o monetari secondo le convenzioni delle loro società, c'è bisogno anzitutto di un equivalente ontogenetico di quelle attività comunicative ed economiche storicamente primarie. Nel caso del linguaggio, naturalmente, l'equivalente ontogenetico è dato dalle attività comunicative di attenzione congiunta cui prendono parte gli adulti e i bambini in età prelinguistica – come abbiamo appena visto. Ma l'apprendimento di un'espressione linguistica chiama ulteriormente in causa l'attenzione congiunta. Determinare la specifica intenzione comunicativa

di un adulto quando usa un'espressione linguistica sconosciuta nel contesto di un'attività di attenzione congiunta è tutt'altro che semplice: richiede infatti che il bambino sia in grado di comprendere i differenti ruoli svolti dal parlante e dall'ascoltatore nell'attività di attenzione congiunta come pure la specifica intenzione comunicativa dell'adulto nell'ambito di quell'attività – dopo di che, il bambino deve essere in grado di esprimere nei confronti delle altre persone la stessa intenzione comunicativa espressa in precedenza nei suoi confronti [Hobson 1993]. E, abbastanza spesso, i bambini devono fare tutto questo nel flusso di interazioni sociali nel quale adulti e bambini sono normalmente immersi nella loro vita quotidiana, senza che gli adulti interrompano quello che stanno facendo per insegnare loro questa o quella parola.

Apprendere a usare i simboli linguistici e gli altri artefatti simbolici ha molteplici conseguenze. Ovviamente essi permettono ai bambini di fare cose che altrimenti non sarebbero in grado di fare, dato che questi artefatti simbolici sono stati creati allo scopo di permettere o di facilitare certi tipi di interazioni cognitive e sociali. Ma la cosa più importante è che essi danno origine a forme di rappresentazione cognitiva completamente nuove che modificano il modo in cui i bambini concepiscono il mondo. Mentre i primati non umani e i neonati umani costruiscono rappresentazioni cognitive del proprio ambiente sulla base delle percezioni e delle propriocezioni scaturite dalla loro stessa esperienza (si tratta, fondamentalmente, di rappresentazioni sensomotorie), i bambini, quando cominciano a comunicare simbolicamente con altri agenti intenzionali, vanno oltre queste semplici rappresentazioni individuali.

Le rappresentazioni simboliche che i bambini imparano a usare nelle interazioni sociali con le altre persone sono speciali perché sono *a)* intersoggettive, nel senso che un simbolo è «condiviso» socialmente con le altre persone; e *b)* dipendenti dalla prospettiva, nel senso che ciascun simbolo rispecchia un particolare modo di concepire un certo fenomeno (la categorizzazione è un caso speciale di questo processo). Sul piano teorico, il punto essenziale è che i simboli linguistici esprimono gli innumerevoli modi di interpretare intersoggettivamente il mondo che si sono accumulati in tempi storici in una cultura, e l'apprendimento dell'uso convenzionale di questi

artefatti simbolici, e dunque l'interiorizzazione di queste interpretazioni trasforma alla radice la natura delle rappresentazioni cognitive dei bambini.

1. *L'acquisizione del linguaggio e il suo fondamento nella cognizione sociale*

Nel capitolo III abbiamo esposto una teoria dell'adattamento culturale dell'uomo fondata sulla capacità del bambino di comprendere le altre persone come agenti intenzionali, capacità che emerge tra i nove e i dodici mesi. Essa però non emerge nel vuoto, ma nell'ambito dei processi attraverso i quali il bambino entra in contatto con le altre persone e interagisce con queste in una varietà di modi. In particolare, accade a volte che le altre persone si rivolgano al bambino facendo suoni e muovendo le mani in modo divertente, e diano l'impressione di aspettarsi in cambio una risposta. Per giungere a considerare questi suoni e questi movimenti delle mani come qualcosa che abbia un significato comunicativo che potrebbe essere appreso e usato, il bambino deve capire che essi sono motivati da un tipo particolare di intenzione, cioè un'*intenzione comunicativa*.

Ma la comprensione di un'intenzione comunicativa può avere luogo solo nell'ambito di un qualche tipo di scena di attenzione congiunta che ne costituisca il fondamento sociocognitivo; inoltre, per imparare a esprimere la stessa intenzione comunicativa (attraverso gli stessi mezzi comunicativi) delle altre persone, è necessario comprendere che i partecipanti all'evento comunicativo possono scambiarsi di ruolo: io posso agire con una persona così come essa ha appena fatto con me. Al centro della teoria qui proposta vi sono dunque, nell'ordine: *a)* le scene di attenzione congiunta come fondamento sociocognitivo delle prime fasi dell'acquisizione del linguaggio; *b)* la comprensione delle intenzioni comunicative come principale processo sociocognitivo per mezzo del quale i bambini comprendono l'uso dei simboli linguistici da parte degli adulti; *c)* l'imitazione per inversione dei ruoli come principale processo di apprendimento culturale per mezzo del quale i bambini imparano a usare attivamente i simboli linguistici.

1.1. *Le scene di attenzione congiunta*

Nella plurisecolare tradizione intellettuale dell'Occidente, molti studiosi hanno descritto gli atti di riferimento linguistico nei termini di due soli elementi: il simbolo e il suo referente nel mondo percettivo. Questa concezione, però, lascia molto a desiderare. Essa è inadeguata sul piano teorico, come è stato mostrato da filosofi come Wittgenstein [1953] e Quine [1960], ed è inadeguata, sotto molti aspetti, sul piano empirico – si pensi in particolare alla sua incapacità di dar conto dell'acquisizione e dell'uso di quei simboli linguistici la cui relazione con il mondo percettivo è, nel migliore dei casi, tenue, vale a dire la maggior parte dei simboli linguistici che non sono nomi propri o nomi a livello-base, per esempio i verbi, le preposizioni, gli articoli e le congiunzioni [Tomasello e Merriman 1995]. Occorre dunque riconoscere esplicitamente, sul piano teorico, che il riferimento linguistico è un atto *sociale* nel quale una persona cerca di far sì che un'altra focalizzi la sua attenzione su qualche aspetto del mondo. E dobbiamo anche riconoscere, sul piano empirico, che il riferimento linguistico può essere compreso solo nel contesto di certe forme di interazione sociale che chiamerò *scene di attenzione congiunta* [Bruner 1983; Clark 1996; Tomasello 1988; 1992a].

Le scene di attenzione congiunta sono interazioni sociali nelle quali il bambino e l'adulto prestano congiuntamente attenzione a una terza cosa, nonché all'attenzione reciproca verso quella cosa per un arco di tempo ragionevolmente esteso. I termini che sono stati usati sono quelli di *interazione di attenzione congiunta*, *episodio di attenzione congiunta*, *impegno di attenzione congiunta*, e *format di attenzione congiunta*. Sto introducendo qui un termine nuovo, ancorché affine, allo scopo di dare il giusto rilievo a due punti essenziali, a volte trascurati nelle trattazioni precedenti di questo fenomeno generale.

Il primo punto riguarda gli elementi inclusi nelle scene di attenzione congiunta. Da un lato, le scene di attenzione congiunta non sono eventi percettivi: esse includono solo una sottoparte del mondo percettivo del bambino. D'altro lato, le scene di attenzione congiunta non sono neanche eventi linguistici: vi è in esse più di quel che è esplicitamente indicato da qualsivoglia insieme di simboli linguistici. Le scene di attenzione

congiunta si trovano perciò in una sorta di terra di mezzo – una importantissima terra di mezzo fatta di una realtà socialmente condivisa – tra il mondo percettivo, più ampio, e il mondo linguistico, più ristretto. Il secondo punto che va sottolineato è il fatto che la comprensione di una scena di attenzione congiunta da parte del bambino include come parte integrante il bambino stesso e il suo ruolo nell'interazione considerati dallo stesso punto di vista «esterno» dal quale sono considerati l'altra persona e l'oggetto, di modo che vi sia un unico *format* rappresentazionale – cosa che è di cruciale importanza per il processo di acquisizione dei simboli linguistici.

Queste due caratteristiche delle scene di attenzione congiunta possono essere illustrate con un esempio. Immaginiamo un bambino che sta giocando nella sua camera con un giocattolo; egli, nello stesso tempo, percepisce molte delle altre cose che si trovano nella stanza. Un adulto entra nella stanza e si mette a giocare con il bambino. La scena di attenzione congiunta viene a essere costituita da quegli oggetti e quelle attività su cui il bambino sa essere concentrata l'attenzione sua e dell'adulto, e su cui entrambi sanno essere concentrata la loro attenzione: non sarebbe attenzione congiunta se la loro attenzione fosse accidentalmente concentrata sulla stessa cosa ma ciascuno fosse inconsapevole dell'altro [Tomasello 1995a].

In questa situazione, cose come il tappeto, il divano e il pannolino non sono parte della scena di attenzione congiunta – anche se il bambino, a livello individuale, può percepirla senza discontinuità significative – perché non sono parte di «ciò che stiamo facendo». D'altra parte, se l'adulto entra nella stanza con un nuovo pannolino e si accinge a cambiare il bambino sul tappeto, allora la scena di attenzione congiunta è completamente differente. In questo caso, gli elementi cruciali sono il pannolino, le salviette detergenti e, forse, il tappeto – ma non i giocattoli, poiché «noi» non abbiamo scopi riferiti ai giocattoli. Il punto è che le scene di attenzione congiunta sono definite intenzionalmente; esse, cioè, traggono la loro identità e la loro coerenza dal fatto che il bambino e l'adulto comprendono «ciò che stiamo facendo» nei termini delle attività dirette a un obiettivo nelle quali siamo impegnati. In un caso stiamo giocando con un giocattolo, e dunque di ciò che stiamo

facendo sono parte certi oggetti e certe attività; nell'altro caso, invece, stiamo cambiando un pannolino, il che chiama in causa, dal punto di vista della nostra attenzione congiunta, un insieme di oggetti e di attività del tutto differenti. In una qualunque scena di attenzione congiunta siamo dunque mutuamente interessati solo a un sottoinsieme di tutte le cose che sarebbe possibile percepire in quella situazione.

Ma la scena di attenzione congiunta non coincide con la scena referenziale simbolizzata esplicitamente in un'espressione linguistica; la scena di attenzione congiunta fornisce semplicemente la cornice intersoggettiva nella quale il processo di simbolizzazione ha luogo. Illustrando con un altro esempio i principi generali in gioco, immaginiamo un turista americano che visita l'Ungheria e si trova in una stazione ferroviaria. All'improvviso gli si avvicina un parlante nativo che gli rivolge la parola in ungherese. È molto improbabile che in questa situazione il turista americano apprenda l'uso convenzionale di qualsivoglia parola o espressione della lingua ungherese. Ma immaginiamo che l'americano si trovi allo sportello della biglietteria, dove c'è un bigliettaio che parla ungherese, e cerchi di acquistare un biglietto. Stavolta il turista può apprendere un po' di parole e di espressioni della lingua ungherese perché ciascuno dei due, in questo contesto, comprende gli scopi dell'altro in relazione all'orario dei treni, all'acquisto dei biglietti, al cambio di valuta, e così via – scopi espressi direttamente tramite l'esecuzione di azioni dotate di significato e comprese in modo indipendente, come le azioni di dare o prendere un biglietto o del denaro.

In una situazione siffatta, la chiave dell'apprendimento del linguaggio potrebbe essere l'uso di una nuova parola o una nuova espressione da parte del parlante nativo in un contesto che rendesse chiara la ragione di quell'enunciato in quella circostanza – per esempio, nell'atto di allungare la mano per prendere le banconote che il turista gli porge o nel dargli il biglietto o il resto. In questi casi, l'apprendimento potrebbe essere fondato su un'inferenza di questo tipo: se quell'espressione sconosciuta significasse X, allora essa sarebbe *pertinente* rispetto allo scopo del bigliettaio in questa scena di attenzione congiunta [Sperber e Wilson 1986; Nelson 1996]. La scena referenziale simbolizzata nel linguaggio riguarda perciò solo

un sottoinsieme di quel che accade nelle interazioni intenzionali comprese nella scena di attenzione congiunta.

Il secondo punto cruciale relativo alle scene di attenzione congiunta è che, dal punto di vista del bambino, esse pongono sullo stesso piano concettuale tutti e tre gli elementi della scena: l'oggetto di attenzione congiunta, l'adulto, e il bambino stesso. Fino a oggi, l'inclusione del bambino non è mai stata sottolineata esplicitamente (da me o, per quanto ne so, da altri), e in effetti l'attenzione congiunta è caratterizzata a volte come la coordinazione dell'attenzione del bambino tra due sole entità: l'oggetto e l'adulto. Ma, come abbiamo visto nel capitolo III, quando il bambino comincia a controllare l'attenzione degli adulti verso le entità esterne, l'entità esterna è a volte il bambino stesso, il quale comincia così a controllare l'attenzione degli adulti nei suoi confronti e a vedere se stesso, per così dire, dal di fuori. Il bambino comprende da questo punto di vista esterno anche il ruolo dell'adulto, cosicché, in definitiva, è come se stesse osservando l'intera scena dall'alto, con se stesso come un attore tra gli altri. All'opposto, le altre specie di primati e i bambini di sei mesi considerano l'interazione sociale da un punto di vista «interno», nel quale gli altri partecipanti appaiono in un *format* (esterocezione in terza persona), e «io» in un *format* differente (propriocezione in prima persona, vedi Barresi e Moore [1996]). Nella teoria delle immagini mentali viene fatta una distinzione simile tra le immagini mentali dal punto di vista soggettivo (per esempio, vedo una palla che viene scagliata via dal mio piede) e le immagini da un punto di vista esterno (per esempio, vedo me stesso – tutto il mio corpo – che dà un calcio alla palla, e lo vedo da un punto di vista esterno, proprio come quando vedo un'altra persona che dà un calcio alla palla).

È difficile sopravvalutare l'importanza di questo modo di comprendere le scene di attenzione congiunta. Per fungere da *format* per l'acquisizione del linguaggio, la scena di attenzione congiunta deve essere compresa dal bambino in modo tale che i ruoli dei partecipanti non siano fissi ma, in un certo senso, intercambiabili [Bruner 1983]. Ciò permette al bambino, come vedremo più avanti, di assumere il ruolo dell'adulto e usare una nuova parola per orientare l'attenzione dell'adulto nello stesso modo in cui l'adulto ne aveva appena usata una per orientare la

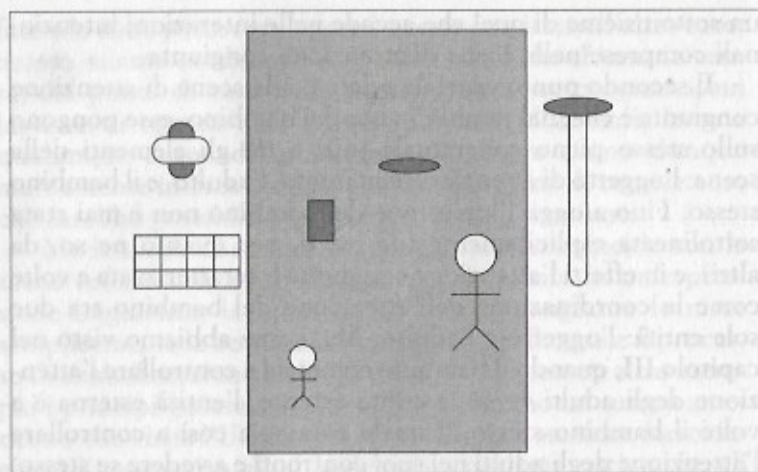


FIG. 4.1. Una scena di attenzione congiunta comprendente il bambino (il Sé), l'adulto e due oggetti di attenzione congiunta, con tre oggetti percepiti ma estranei alla scena stessa.

sua – che è ciò che chiamerò imitazione per inversione dei ruoli. Per il momento, mi limiterò semplicemente a un'ipotetica scena di attenzione congiunta considerata dal punto di vista del bambino, come nella figura 4.1. Due punti vanno ancora una volta sottolineati: *a*) nell'ambito della scena percettiva, la scena di attenzione congiunta è focalizzata su un sottoinsieme di entità e di attività; analogamente, nell'ambito della scena di attenzione congiunta, la scena referenziale è focalizzata su un sottoinsieme di entità e di attività; e *b*) il bambino considera se stesso come uno dei partecipanti alla scena, sullo stesso piano dell'adulto e dell'oggetto di attenzione congiunta.

1.2. La comprensione delle intenzioni comunicative

Immaginiamo ora che un adulto rivolga la parola a un bambino troppo piccolo per comprendere o partecipare a una scena di attenzione congiunta, e a maggior ragione per comprendere il linguaggio. Per un bambino così piccolo l'adulto sta solo emettendo dei suoni. A volte, naturalmente, egli può

imparare ad associare uno di questi suoni con un evento percettivo, proprio come il cane di casa può comprendere che il suono *si mangia* preannuncia l'arrivo del cibo. Ma questo non è linguaggio. Per il bambino il suono diventa linguaggio quando e solo quando egli comprende che l'adulto sta emettendo quel suono affinché il bambino stesso presti attenzione a qualcosa. Questa comprensione non è una conclusione scontata, ma un'importante conquista. Essa richiede la comprensione delle altre persone come agenti intenzionali, come abbiamo visto nel capitolo III; richiede la partecipazione a una scena di attenzione congiunta, come abbiamo chiarito sopra; e richiede anche la comprensione di un particolare tipo di atto intenzionale in una scena di attenzione congiunta, cioè un atto comunicativo che esprima un'intenzione comunicativa.

La cosa può essere chiarita alla luce di quel che accade quando uno sperimentatore cerca di comunicare con scimmie antropomorfe e con bambini di due anni per mezzo di segni che siano loro completamente nuovi. Tomasello, Call e Gluckman [1997] hanno fatto esattamente questo, suggerendo in vari modi a degli scimpanzé e a dei bambini di età compresa tra i due e i tre anni in quale di tre contenitori fosse racchiusa una ricompensa. Gli sperimentatori usavano tre tipi di suggerimenti: *a*) indicare il contenitore giusto; *b*) mettere un pezzetto di legno sopra il contenitore giusto; e *c*) mostrare una riproduzione esatta del contenitore giusto. I bambini già conoscevano il gesto di indicare, ma non l'uso dei pezzi di legno e della riproduzione come segni comunicativi. Nondimeno, essi usavano perfettamente questi nuovi segni per trovare la ricompensa. Invece, nessuna delle scimmie riusciva nel compito con i segni comunicativi che non conoscesse già prima dell'esperimento. Una spiegazione di questi risultati è che le scimmie non erano in grado di capire che lo sperimentatore aveva delle intenzioni riguardanti i loro stati attentivi. Le scimmie consideravano i tentativi di comunicazione dell'uomo alla stregua di indizi discriminativi che, come tutti gli altri indizi discriminativi, dovevano essere laboriosamente appresi attraverso ripetute esperienze. I bambini, invece, consideravano i tentativi di comunicazione come un'espressione dell'intenzione dell'adulto di influenzare appropriatamente la loro attenzione nella situazione in cui si trovavano.

In altri termini, i bambini comprendevano, almeno in parte, le intenzioni comunicative dello sperimentatore. Le interpretazioni e le spiegazioni della natura delle intenzioni comunicative hanno una complessa storia filosofica [per un'utile rassegna, vedi Levinson 1983], ma io seguirò piuttosto le orme di Clark [1996], che affronta alcuni degli stessi temi da un punto di vista più psicologico. Secondo l'analisi qui proposta, per comprendere la tua intenzione comunicativa io devo comprendere che:

Tu intendi che [io condivida l'attenzione verso (X)].

È opinione generale, da Grice [1975] in poi, che la comprensione di un'intenzione comunicativa debba avere questa struttura interna. Per esempio, se una persona arriva e mi fa sedere su una sedia, io riconosco la sua intenzione che io sieda, mentre se mi dice «Siedi», io riconosco la sua intenzione che io presti attenzione alla sua richiesta di sedermi. Questa analisi mostra abbastanza chiaramente che la comprensione di un'intenzione comunicativa è un caso speciale della comprensione di un'intenzione; si tratta di comprendere l'intenzione di un'altra persona riguardo al mio stato attentivo. Questa è una forma di comprensione chiaramente più complessa rispetto alla mera comprensione dell'intenzione di un'altra persona. Per comprendere che l'intenzione di un'altra persona è dare un calcio a una palla, devo semplicemente determinare il suo scopo in relazione alla palla. Ma per comprendere che cosa intende un'altra persona quando emette il suono «Palla!» rivolgendosi a me, devo determinare il suo scopo in relazione ai miei stati intenzionali/attentivi verso una certa entità terza.

Questa teoria discende direttamente dalla mia analisi del modo in cui i bambini comprendono gli altri come agenti intenzionali e comprendono e vedono il Sé come un agente intenzionale che partecipa a scene di attenzione congiunta al pari di altri agenti intenzionali. Da questo punto di vista, solo un bambino in grado di tenere sotto controllo gli stati intenzionali degli altri nei suoi confronti – in particolare, nei confronti dei suoi stati intenzionali – è in grado di comprendere un'intenzione comunicativa. Se proviamo a illustrare la cosa con un diagramma, differenziando il caso del bambino dal

caso dello scimpanzé che non è in grado di comprendere le intenzioni comunicative, otteniamo qualcosa di simile alla figura 4.2. La figura 4.2a illustra l'esperienza dello scimpanzé quando vede un altro individuo nell'atto di alzare il braccio. Lo scimpanzé anzitutto vede il braccio alzato, poi si aspetta un certo seguito (sulla base della sua passata esperienza di situazioni simili). La figura 4.2b illustra l'esperienza del bambino quando comprende con successo il tentativo linguistico dell'adulto di far sì che egli presti attenzione a un'entità esterna. Il primo quadro raffigura il bambino che, dall'esterno, vede se

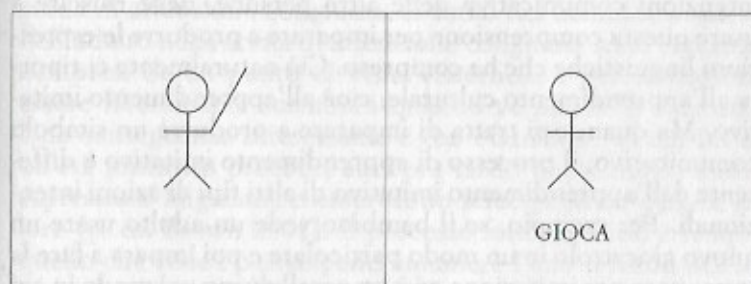


FIG. 4.2a. Quel che gli scimpanzé comprendono quando percepiscono e interpretano un segnale gestuale: prima vedono l'altro individuo che fa un gesto, poi immaginano quel che farà dopo. Il Sé non è rappresentato.

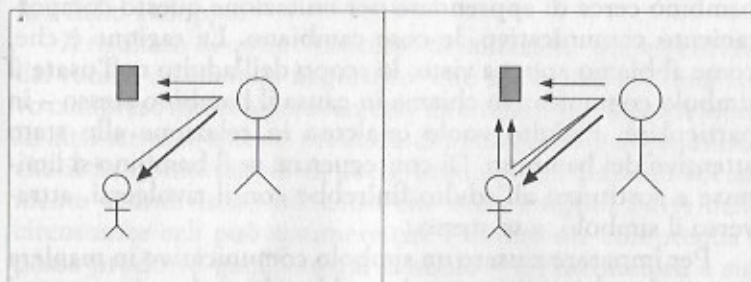


FIG. 4.2b. Quel che i bambini comprendono quando percepiscono e interpretano un simbolo linguistico: prima comprendono che l'altra persona vuole che essi condividano l'attenzione verso qualcosa, poi immaginano in che cosa potrebbe consistere tale condivisione. Condividere significa che entrambi prestano attenzione sia al referente, sia l'uno all'attenzione dell'altro verso il referente. Il Sé è concettualizzato nello stesso modo in cui è concettualizzata l'altra persona.

stesso partecipare all'interazione nella quale l'adulto sta cercando di far sì che il bambino stesso presti attenzione a X; il secondo quadro raffigura il bambino che risponde in modo appropriato alla richiesta dell'adulto e condivide con lui l'attenzione verso X (entrambi i partecipanti prestano attenzione sia all'oggetto sia l'uno all'attenzione dell'altro verso l'oggetto).

1.3. *Imitazione per inversione dei ruoli e intersoggettività*

Una volta che il bambino sia in grado di comprendere le intenzioni comunicative delle altre persone, deve riuscire a usare questa comprensione per imparare a produrre le espressioni linguistiche che ha compreso. Ciò naturalmente ci riporta all'apprendimento culturale, cioè all'apprendimento imitativo. Ma quando si tratta di imparare a produrre un simbolo comunicativo, il processo di apprendimento imitativo è differente dall'apprendimento imitativo di altri tipi di azioni intenzionali. Per esempio, se il bambino vede un adulto usare un nuovo giocattolo in un modo particolare e poi impara a fare la stessa cosa per imitazione, vi è un parallelismo nel modo in cui l'adulto e il bambino usano il giocattolo: il bambino si limita a sostituirsi all'adulto. Ma quando un adulto si rivolge al bambino attraverso un nuovo simbolo comunicativo con l'intenzione che il bambino presti attenzione a quel giocattolo, e il bambino cerca di apprendere per imitazione questo comportamento comunicativo, le cose cambiano. La ragione è che, come abbiamo appena visto, lo scopo dell'adulto nell'usare il simbolo comunicativo chiama in causa il bambino stesso – in particolare, l'adulto vuole qualcosa in relazione allo stato attentivo del bambino. Di conseguenza, se il bambino si limitasse a sostituirsi all'adulto finirebbe con il rivolgersi, attraverso il simbolo, a se stesso.

Per imparare a usare un simbolo comunicativo in maniera convenzionalmente appropriata, il bambino deve impegnarsi in quella che ho chiamato *imitazione per inversione dei ruoli* [Tomasello 2001]. Il bambino deve, cioè, imparare a usare un simbolo nei confronti dell'adulto nello stesso modo in cui l'adulto lo ha usato nei suoi confronti. Questo evidentemente è un processo di apprendimento imitativo nel quale il bambi-

no si conforma all'adulto quanto allo scopo e al mezzo per raggiungere tale scopo; tuttavia, in questo caso, il bambino non deve soltanto sostituirsi all'adulto come attore (che è quel che accade in tutte le forme di apprendimento culturale), ma anche sostituire l'adulto a se stesso come bersaglio dell'atto intenzionale (cioè, lo scopo non deve essere il proprio stato attentivo ma, al posto di questo, lo stato attentivo dell'adulto). Se osserviamo nuovamente il secondo quadro della figura 4.2b, vediamo che l'inversione dei ruoli all'opera in questo tipo di apprendimento imitativo è la conseguenza diretta dell'assunzione di un punto di vista esterno nei confronti della scena di attenzione congiunta. Il ruolo del bambino e quello dell'adulto nella scena di attenzione congiunta sono entrambi compresi da un punto di vista «esterno», e perciò possono essere liberamente scambiati quando ve ne sia la necessità. Una conseguenza interessante è che i bambini (alcuni fin da un'età piuttosto precoce, altri più tardi) apprendono nuove espressioni linguistiche osservando terze persone parlare tra loro [p. es. Brown 2001]. Il processo fondamentale è sempre quello che vede i partecipanti assumere l'uno il ruolo dell'altro; solo che in questo caso il bambino non è uno dei partecipanti allo scambio linguistico originale. Questo tipo di apprendimento del linguaggio non è stato studiato abbastanza in dettaglio da permetterci di dire come i bambini riescano nell'impresa, o se essa crei loro particolari difficoltà nelle prime fasi dello sviluppo.

Il risultato di questo processo di imitazione per inversione dei ruoli è un *simbolo linguistico*: uno strumento comunicativo compreso intersoggettivamente da entrambi gli interlocutori. In altri termini, questo processo di apprendimento garantisce che il bambino sappia di avere acquisito un simbolo socialmente «condiviso», nel senso che nella maggior parte delle circostanze egli può assumere che l'ascoltatore comprenda e possa produrre quello stesso simbolo – e l'ascoltatore a sua volta sa che entrambi possono comprendere e produrre il simbolo. Il processo di comprensione dei segnali comunicativi – per esempio, negli scimpanzé e in certa comunicazione gestuale dei bambini in età prelinguistica – è molto differente perché ciascuno dei partecipanti all'interazione comprende il proprio ruolo solo da un punto di vista interno. Ma anche nel

caso dei gesti non linguistici, se il processo di apprendimento implica la comprensione delle intenzioni comunicative e l'imitazione per inversione dei ruoli entro una scena di attenzione congiunta, il risultato sarà un simbolo comunicativo. Così, se un bambino indica qualcosa a qualcuno avendo appreso il gesto di indicare attraverso l'imitazione degli adulti che in precedenza gli avevano rivolto quel gesto, per ciò stesso il suo indicare sarà simbolico – vedi anche i primi gesti simbolici dei bambini, dall'agitare la mano in segno di saluto allo sbattere le braccia come un uccello, di cui si sono occupati Acredolo e Goodwyn [1988].

È interessante notare che l'intersoggettività inerente ai simboli socialmente condivisi, ma non ai segnali unidirezionali, è alla base di molte delle «implicature» pragmatiche studiate da Grice [1975], che riguardano la mia aspettativa che l'interlocutore usi i mezzi di espressione convenzionali – che (entrambi sappiamo che) egli conosce – e non altri più lenti o più indiretti. Ciò accade, per esempio, quando un bambino si rende conto che un nuovo simbolo è usato per indicare un nuovo aspetto di una situazione perché, se l'adulto avesse inteso riferirsi ad aspetti della situazione di cui qualcuno aveva già parlato, avrebbe usato un simbolo noto (questo comportamento, che permette al bambino di apprendere rapidamente il significato delle parole, è detto «mappatura rapida»; Carey [1978]).

Riassumendo, per apprendere l'uso convenzionale di un simbolo linguistico compreso intersoggettivamente il bambino deve:

- comprendere gli altri come agenti intenzionali;
- partecipare a una scena di attenzione congiunta che definisca lo sfondo sociale e cognitivo degli atti di comunicazione simbolica, e in particolare linguistica;
- comprendere non soltanto le intenzioni, ma le intenzioni comunicative nelle quali qualcuno intende che il bambino stesso presti attenzione a qualcosa nella scena di attenzione congiunta;
- scambiare il proprio ruolo con quello dell'adulto nel processo di apprendimento culturale e così fare nei confronti dell'adulto ciò che questi ha fatto nei suoi confronti – è questo processo che crea le convenzioni o i simboli comunicativi compresi intersoggettivamente.

Apprendere in questo modo i simboli linguistici permette ai bambini di cominciare a trarre profitto dalle abilità e dalle conoscenze sociali di ogni genere preesistenti nella loro comunità locale e nel complesso della loro cultura. Ma c'è dell'altro. Quel che rende i simboli linguistici effettivamente unici dal punto di vista cognitivo è il fatto che ciascun simbolo esprime un particolare punto di vista su un'entità o un evento: questo oggetto è nello stesso tempo una rosa, un fiore e un regalo. La natura prospettica dei simboli linguistici moltiplica indefinitamente il grado di specificità con cui essi possono essere usati per manipolare l'attenzione degli altri, e questo fatto ha profonde implicazioni per la natura delle rappresentazioni cognitive, come vedremo più avanti. Ma nel presente contesto – al centro del quale è il modo in cui i bambini apprendono nuovi simboli linguistici – esso crea un problema. Il problema è che questa grande specificità richiede che il bambino non solo comprenda che l'adulto vuole qualcosa in relazione al suo stato attentivo, ma identifichi anche il particolare bersaglio cui l'adulto vuole che egli si rivolga in una particolare scena di attenzione congiunta.

2. *L'acquisizione del linguaggio e il suo fondamento nell'interazione sociale*

Ora che il bambino è munito di vari tipi di abilità socio-cognitive, e dando per scontato che posseda le abilità generali dei primati in fatto di percezione, memoria, categorizzazione, e così via, resta la questione di come queste abilità siano usate per apprendere i simboli linguistici. Il problema – formulato originariamente da Wittgenstein [1953] e poi approfondito da Quine [1960] – deriva dalla natura prospettica dei simboli linguistici (sebbene questi filosofi non formulassero la questione in questi termini). A causa della natura prospettica dei simboli linguistici, non vi sono procedure algoritmiche per determinare la particolare intenzione comunicativa di una persona in una particolare occasione. Quando un adulto mostra una palla e dice *dax*, come può il bambino sapere se l'adulto si stia riferendo proprio a quell'entità, al suo colore, a qualche insieme di entità più grande (come quello dei giocat-

tolì), all'atto di mostrare qualcosa, o a una qualunque di un'infinità di altre cose? Vi è chi ha cercato di risolvere questo problema assumendo che quando l'acquisizione del linguaggio ha inizio il bambino disponga di certi «vincoli» sull'apprendimento delle parole che, in modo automatico, lo indirizzerebbero verso ciò cui il parlante intendeva riferirsi [per esempio, Markman 1989; 1992; Gleitman 1990].

Crede poco alle soluzioni che fanno appello a questa sorta di «armonia prestabilita», e sono incline piuttosto a un approccio basato sulla capacità sociopragmatica dei bambini di comprendere le intenzioni comunicative degli adulti nel loro contesto [Tomasello 1992a; 1995c; 2001]. Così, almeno parte della mia soluzione del problema di Wittgenstein sta nella comprensione da parte del bambino delle intenzioni comunicative degli adulti all'interno di una scena di attenzione congiunta dotata di significato (quella che Wittgenstein chiamava una «forma di vita»), una comprensione che è indipendente da qualsivoglia comprensione dell'espressione linguistica da apprendere, sebbene possa ricollegarsi, naturalmente, alla comprensione di altre espressioni linguistiche nella stessa situazione. Ciò avviene in modi spesso molto sottili e complessi, dato che i bambini devono identificare le intenzioni comunicative degli adulti nel flusso ininterrotto del discorso e dell'interazione sociale. Vi è poi l'altra parte della soluzione, che nasce là dove nasce il problema stesso. La natura prospettica dei simboli linguistici fa sì che in molti casi i significati di questi simboli siano in contrasto gli uni con gli altri – in un certo senso, essi sono definiti gli uni relativamente agli altri, come *comprare* e *vendere*, *credito* e *debito* – e questo aiuta i bambini a discernere significati sottilmente differenti, specialmente dopo che abbiano apprese alcune parole-base.

2.1. L'attenzione congiunta e il primo linguaggio

Bruner [1975; 1983] è stato il primo studioso dell'apprendimento del linguaggio ad affrontare il problema di Wittgenstein e a proporre una soluzione. In linea con l'orientamento generale di Wittgenstein, Bruner sosteneva che il bambino acquisisce l'uso convenzionale di un simbolo linguistico imparando a

partecipare a un *format* interattivo (forma di vita, scena di attenzione congiunta) che egli comprende indipendentemente dal linguaggio, di modo che il linguaggio dell'adulto possa poggiare su una base di esperienze condivise il cui significato sociale è già noto al bambino.

Un elemento chiave di questo processo è ovviamente il fatto che il bambino sia in grado di comprendere gli adulti come esseri intenzionali e possa condividere con loro l'attenzione in particolari contesti. Ma un altro elemento è il mondo sociale esterno nel quale il bambino vive e che gli preesiste. Per acquisire il linguaggio, il bambino deve vivere in un mondo nel quale siano incluse attività sociali strutturate che egli sia già in grado di comprendere, come il nostro ipotetico turista in Ungheria comprendeva il processo di comprare i biglietti e quello di viaggiare in treno. Spesso, nel caso dei bambini, questo richiede che la stessa attività generale abbia luogo su base regolare o di routine, in modo che i bambini possano giungere a capire come essa si svolge e come funzionano i vari ruoli sociali al suo interno. E naturalmente, nel caso dell'acquisizione del linguaggio, l'adulto deve usare un nuovo simbolo linguistico in un modo di cui il bambino possa comprendere la pertinenza rispetto a quell'attività condivisa (come il primo ungherese incontrato dal turista non faceva). In generale, un bambino che nascesse in un mondo nel quale gli eventi non si ripetono mai, lo stesso oggetto non appare mai due volte, gli adulti non dicono mai la stessa cosa nello stesso contesto, difficilmente potrebbe acquisire una lingua naturale – quali che fossero le sue capacità cognitive.

Una varietà di studi ha mostrato che una volta compiuti i primi passi nell'acquisizione del linguaggio i bambini apprendono più facilmente nuove parole nelle scene di attenzione congiunta che sono socialmente condivise con altre persone, e in particolare in quelle che ricorrono spesso nella loro esperienza quotidiana, come fare il bagno, mangiare, cambiare il pannolino, leggere libri e viaggiare in auto. Queste attività sono per molti aspetti analoghe allo scenario «comprare i biglietti alla stazione», nel senso che il bambino comprende quelli che sono i suoi scopi e gli scopi dell'adulto in quella situazione, il che gli permette di inferire la pertinenza rispetto a quegli scopi di ciò che l'adulto dice, dal che il bambino può

inferire quale precisamente debba essere l'oggetto della sua attenzione. Per esempio, Tomasello e Todd [1983] hanno osservato che i bambini che tra i dodici e i diciotto mesi trascorrevano più tempo in attività di attenzione congiunta con le madri, a diciotto mesi avevano vocabolari più ampi [vedi anche Smith, Adamson e Bakeman 1988; Tomasello, Mannle e Kruger 1986]. Per quel che riguarda l'uso del linguaggio da parte degli adulti in queste scene di attenzione congiunta, Tomasello e Farrar [1986] hanno trovato conferma sia correlazionale sia sperimentale all'ipotesi che le madri che, parlando, cercavano di seguire l'attenzione del bambino (parlando, per esempio, di un oggetto su cui l'interesse e l'attenzione del bambino erano già focalizzati) avessero bambini con un vocabolario più ampio rispetto alle madri che cercavano piuttosto di orientare l'attenzione del bambino verso qualcosa di nuovo [vedi anche Akhtar, Dunham e Dunham 1991; Dunham, Dunham e Curwin 1993].

Particolarmente importanti sono forse i risultati di Carpenter, Nagell e Tomasello [1998], secondo i quali relazioni analoghe valgono ancor più precocemente, in effetti già quando i bambini cominciano appena ad apprendere e a usare il linguaggio. Essi hanno osservato che i bambini che a dodici mesi trascorrevano una maggiore quantità di tempo in attività di attenzione congiunta con la madre avevano risultati migliori nella comprensione e nella produzione del linguaggio alla stessa età e nei mesi immediatamente successivi. Essi hanno inoltre osservato che se la madre seguiva, sottolineandolo con parole, il fuoco attentivo del bambino quando questi aveva dodici mesi, nei mesi immediatamente successivi il bambino mostrava di possedere un vocabolario di comprensione più ampio (mentre le relazioni con la produzione del linguaggio venivano alla luce un po' più tardi).

Quando queste due variabili – la quantità di tempo in cui il bambino era impegnato in attività di attenzione congiunta e la tendenza della madre a «tener dietro» al fuoco attentivo del bambino nell'uso referenziale del linguaggio – erano usate assieme in un sistema di equazioni di regressione, era possibile predire più della metà della varianza nella comprensione e nella produzione del linguaggio rilevata in diversi momenti tra i dodici e i quindici mesi, con ciascuna variabile che dava

conto di una significativa quantità di varianza. Diversi aspetti dello sviluppo cognitivo non sociale dei bambini – soprattutto quelli riguardanti la conoscenza degli oggetti e la conoscenza spaziale – emergevano indipendentemente dal linguaggio e dalle altre attività di attenzione congiunta, il che dimostra come la correlazione tra le attività di attenzione congiunta e il linguaggio non fosse soltanto il risultato di un processo di sviluppo più generale.

Il risultato di questo studio – che conferma i risultati correlazionali e sperimentali di studi analoghi con bambini leggermente più grandi – è che l'emergere della capacità dei bambini di impegnarsi assieme agli adulti in attività di attenzione congiunta non mediate linguisticamente attorno all'anno di età è strettamente legato all'emergere delle loro abilità linguistiche [Rollins e Snow 1999, per alcuni risultati analoghi relativi all'attenzione congiunta e alla comparsa delle abilità sintattiche]. Questo risultato è importante perché dimostra che la indubbia corrispondenza cronologica tra le abilità di attenzione congiunta e il linguaggio – ambedue emergono nei mesi che precedono e seguono il primo anno di vita, con le abilità non linguistiche di attenzione congiunta un po' più precoci – non è una coincidenza.

Il problema che questo risultato costituisce per le teorie dell'acquisizione del primo linguaggio che trascurano le dimensioni sociali del processo è serio e immediato. Per le teorie che sono focalizzate soprattutto sulle dimensioni cognitive dell'apprendimento delle parole [per esempio, Markman 1989] o sui processi di apprendimento associativo implicati [Smith 1995], il problema è spiegare perché l'acquisizione del linguaggio segua immediatamente l'emergere delle abilità di attenzione congiunta. E qualunque risposta che faccia appello a processi cognitivi o di apprendimento di tipo non sociale – per esempio, quella secondo cui i bambini a un anno di età cominciano a essere capaci di categorizzare o di apprendere nuovi tipi di cose in generale – deve poi spiegare perché il linguaggio, nel momento in cui emerge, è correlato con abilità non linguistiche di cognizione e di interazione sociale. Che io sappia, l'unica teoria dello sviluppo iniziale dell'apprendimento delle parole e dell'acquisizione del linguaggio che sia in grado di dar conto di questi risultati è la *teoria sociopragmatica* di

Bruner [1983], Nelson [1985] e Tomasello [1992a; 1995c; 2001].

È interessante notare che nello studio di Carpenter, Nagell e Tomasello [1998] la relazione tra l'uso del linguaggio da parte della madre per seguire il fuoco attento del bambino e l'apprendimento del linguaggio da parte del bambino diventa più debole a mano a mano che i bambini diventano più grandi. Questo è un risultato interessante, poiché il fatto che le madri usino il linguaggio per seguire il fuoco attento del bambino potrebbe essere una sorta di *scaffolding* per l'acquisizione del primo linguaggio, nel senso che nelle prime fasi dell'apprendimento del linguaggio esso potrebbe aiutare i bambini a identificare le intenzioni comunicative dell'adulto, mentre questo tipo di *scaffolding* non sarebbe più necessario quando i bambini crescono e imparano a individuare le intenzioni comunicative anche nelle interazioni linguistiche più complesse. In effetti, a partire almeno dai diciotto mesi d'età, i bambini si dimostrano straordinariamente abili nell'identificare le intenzioni comunicative degli adulti in un'ampia varietà di contesti interattivi che non sono di per sé adattati alle loro esigenze.

2.2. L'apprendimento delle parole nel flusso dell'interazione sociale

Nella classe media occidentale è pratica comune che un adulto esibisca o indichi un oggetto e nello stesso tempo ne dica il nome al bambino. La natura sociale di questo processo è evidente: il bambino è chiamato a capire su quale aspetto della situazione l'adulto vuole che egli focalizzi la sua attenzione. A dispetto degli intricati risvolti di questo processo, che Wittgenstein e Quine hanno analizzato, esso è purtuttavia *relativamente* semplice perché il seguire visivamente la direzione dello sguardo e il gesto di indicare sono comportamenti fondamentali per il bambino. È tuttavia noto che in molte culture gli adulti non fanno questo gioco di denominazione con i bambini [Brown 2001]. Inoltre, anche nella classe media occidentale non accade spesso che gli adulti facciano questo gioco di denominazione usando parole diverse dalle etichette

degli oggetti. I verbi, per esempio, sono usati di solito per regolare o prevenire il comportamento dei bambini, e non per denominare le azioni in quanto tali; in effetti, suonerebbe strano se un adulto si rivolgesse a un bambino esclamando: «Guarda, questo è un esempio di mettere (o dare, o prendere)» [Tomasello e Kruger 1992]. I bambini si imbattono invece in molti verbi quando l'adulto dirige il loro comportamento, come in: «Metti via i tuoi giocattoli», detto indicando la scatola dei giocattoli. È evidente che in tali casi gli indizi sociopragmatici che potrebbero segnalare al bambino il referente inteso dall'adulto (cioè, l'azione di mettere) sono molto più sottili, complessi e variegati che non nel contesto ostensivo di denominazione degli oggetti, e in effetti essi possono cambiare profondamente da una situazione all'altra: l'adulto chiede al bambino di mangiare la minestra dirigendo il cucchiaino verso la bocca del bambino stesso, chiede che il bambino gli dia qualcosa tendendo la mano e chiede che i giocattoli siano messi via indicando la destinazione desiderata. Non vi è dunque alcun «gioco di denominazione» standardizzato per i verbi come ve ne è uno per le etichette degli oggetti (anche se non per tutti i bambini) [Tomasello 1995c]. E la situazione si complica ancora se prendiamo in considerazione altre classi di parole, come le preposizioni [Tomasello 1987].

In alcuni recenti studi è stato dimostrato sperimentalmente che i bambini possono apprendere nuove parole in una varietà di situazioni complesse di interazione sociale. Essi apprendono nuove parole non solo quando gli adulti interrompono quello che stanno facendo e dicono a loro beneficio il nome degli oggetti, ma anche nel flusso ininterrotto dell'interazione sociale, nella quale sia il bambino sia l'adulto sono impegnati a perseguire i propri scopi. In nessuno di questi casi il bambino può contare sul fatto che l'adulto segua il fuoco preconstituito della sua attenzione; semmai dev'essere il bambino ad adattarsi al fuoco attento dell'adulto. Per esempio, Baldwin [1991; 1993] ha provato a insegnare nuove parole a bambini di diciannove mesi in due situazioni. Nella prima situazione, l'adulto seguiva il fuoco dell'attenzione del bambino e, come in altri studi, i bambini apprendevano le nuove parole piuttosto bene – meglio che in ogni altra condizione, in effetti. Ma i bambini riuscivano ad apprendere le