

Marcello Frixione

## Scienze cognitive e semantica naturalizzata

**Abstract** - In anni recenti, alcuni filosofi del linguaggio di impostazione analitica si sono avvicinati in varia misura alle ricerche svolte in intelligenza artificiale (IA) e, più in generale, nell'ambito delle scienze cognitive. Tale accostamento è stato motivato in base al tentativo di individuare una soluzione a problemi legati al progetto di una teoria del significato per il linguaggio naturale, problemi che, per diverse ragioni, risultano difficilmente affrontabili con gli strumenti formali tipici della semantica filosofica. Tra essi, mi limito a ricordare quelle che vengono generalmente riconosciute come le difficoltà più pressanti per una teoria del significato di impostazione logica: lo sviluppo di una semantica lessicale adeguata e il trattamento dei contesti di atteggiamento proposizionale. In questa sede non intendo analizzare le motivazioni che hanno condotto a questa svolta, né mi propongo di valutare i risultati ottenuti lungo questa via. Per tali aspetti, rimando alla letteratura esistente, ad esempio (Peters e Saarinen 1982; Eco *et al.* 1986; Marconi 1992; Frixione 1994a) (per qualche cenno al problema della semantica lessicale si veda comunque il par. 4). Lo scopo di questo lavoro consiste piuttosto nel proporre qualche considerazione generale circa le conseguenze di tale tendenza rispetto alla caratterizzazione stessa della linea analitica in filosofia.

Che senza natural dimostramento  
non ho talento - di voler provare  
là dove posa e chi lo fa creare

Guido Cavalcanti

### 1. La svolta linguistica in filosofia

Gli sviluppi in senso cognitivo dello studio del significato linguistico non sono conciliabili con le posizioni antipsicologistiche predominanti nella tradizione analitica a partire da Frege, e che Michael Dummett, ad esempio, ha individuato come determinanti nella caratterizzazione stessa di questa scuola di pensiero. Secondo Dummett (1975, 1988), ciò che contraddistingue la filosofia analitica da altre tendenze filosofiche consiste nell'assunzione che una spiegazione filosofica del pensiero possa essere ottenuta esclusivamente attraverso una spiegazione filosofica del linguaggio, possa cioè essere ottenuta soltanto attraverso l'analisi dei significati linguistici. Tale *svolta linguistica* in filosofia viene fatta risalire da Dummett all'opera di Frege: "solo con Frege si è avuto finalmente un riconoscimento dell'oggetto proprio della filosofia: si è riconosciuto cioè in primo luogo che l'obiettivo della filosofia è l'analisi della struttura del *pensiero* e in secondo luogo, che lo studio del *pensiero* deve essere tenuto nettamente distinto dallo studio del processo psicologico del *pensare*; infine, si è riconosciuto che il solo metodo appropriato per l'analisi del pensiero consiste nell'analisi del *linguaggio*" (Dummett 1975, p. 458). Da questa citazione emerge il ruolo che svolge l'antipsicologismo nel caratterizzare la tradizione analitica (e che la distanzia dalla prospettiva cognitivista). Per Frege i significati non devono essere confusi con entità di tipo mentale. Essi sono oggetti di natura logica e non psicologica. Il senso di un termine o di un enunciato non è l'immagine mentale, o la rappresentazione che il parlante gli associa. Le rappresentazioni mentali infatti sono soggettive, dipendono dalla storia di ciascun individuo, dalle sue esperienze, dalle informazioni di cui dispone. Variano quindi da individuo a individuo, e, in uno stesso individuo, sono diverse in momenti diversi e in fasi diverse della sua vita. Sono inoltre indissolubilmente legate ad elementi di ordine emotivo. Infine, sono entità di tipo eminentemente privato, accessibili esclusivamente a chi ne è il portatore e non comunicabili intersoggettivamente. Viceversa i significati devono essere garanti dell'uso del linguaggio in quanto fenomeno sociale. Devono essere quindi intersoggettivamente condivisibili. Per Frege, il fatto che due individui comprendano e usino lo stesso

linguaggio può essere spiegato esclusivamente postulando il fatto che entrambi riescano ad accedere a un patrimonio comune di sensi. Da ciò per Frege segue la natura oggettiva del senso. "Il senso non costituisce invero [...] qualcosa di inscindibile dal singolo individuo, ma può formare il possesso comune di molti. Che sia così, ce lo prova l'esistenza di un patrimonio di pensieri comuni all'umanità, patrimonio di pensieri che si trasmette di generazione in generazione" (Frege 1892). O ancora: "con il termine 'pensiero' intendo non l'atto soggettivo del pensare, ma il suo contenuto oggettivo che può costituire il possesso comune di molti" (Frege 1892, p. 383 della trad. it. 1965). La conclusione è che i pensieri, i sensi, non sono né oggetti del mondo esterno, né rappresentazioni mentali, bensì sono entità che esistono in un "terzo regno" di natura platonica (Frege 1918). L'atto di pensare comporta operazioni mentali individuali, il cui fine è tuttavia quello di mettere in contatto la mente con i sensi oggettivi del terzo regno. Tali assunzioni ontologiche non sono tuttavia inscindibili dall'antipsicologismo in semantica. Dummett stesso rifiuta il platonismo fregeano rispetto ai significati, ma conserva l'antipsicologismo sulla base di una concezione sociale del significato come uso di origine wittgensteiniana.

Da quanto detto emerge l'assoluta centralità del progetto di una teoria del significato per la filosofia analitica. Con la svolta linguistica la filosofia del linguaggio viene ad assumere il ruolo che nella tradizione moderna a partire da Cartesio era riservato all'epistemologia (Dummett 1983). La filosofia del linguaggio riveste il ruolo di filosofia prima, di dottrina dei fondamenti: ancora secondo Dummett, per i filosofi analitici "la filosofia del linguaggio è il fondamento di tutta la filosofia perché è solo attraverso l'analisi del linguaggio che possiamo analizzare il pensiero" (Dummett 1975, p. 442). In questa direzione, l'analisi del linguaggio può essere vista come una via per conservare alla filosofia la sua dimensione a priori. Ernst Tugendath, ad esempio, che sostiene posizioni per molti versi simili a quelle di Dummett, afferma: "la filosofia analitica corrisponde alla concezione tradizionale della filosofia come conoscenza a priori, e interpreta l'a priori come analitico" (Tugendath 1976, p. 15 della trad. it.).

Vi sono filosofi che sono abitualmente riconosciuti interni alla tradizione analitica, le cui tesi tuttavia si discostano in qualche punto dai canoni individuati da Dummett. A questo proposito, Dummett stesso (1988) cita il caso di Evans (1982). Tugendath (nella prefazione all'edizione italiana di Tugendath 1976) cita Searle (1983) come un esempio di come "nella filosofia analitica ci si è allontanati in molti modi dal significato centrale del linguaggio" e di come il linguaggio venga talvolta considerato solo "una forma di rappresentazione fra le altre" (p. 7). In Searle, questo è motivato dalla considerazione che anche soggetti che non usano il linguaggio possono avere credenze e desideri, e che quindi esistono credenze e desideri che possono non avere natura strettamente linguistica. In base a considerazioni simili, anche Stalnaker (1984) si allontana in questo senso dall'ortodossia analitica quale è stata delineata da Dummett. In casi come questi Dummett è dell'opinione che non ci si trovi più di fronte a esempi di filosofia analitica (Dummett 1988). Nel suo "estremismo", la definizione dummettiana ha il pregio di essere estremamente precisa ed esplicita, e di rendere evidenti presupposti che sono all'opera (spesso assunti in forma implicita) in gran parte della tradizione analitica.

## **2. Linguaggio e scienze cognitive**

Prima di procedere ad esaminare le conseguenze dell'avvicinamento tra scienze cognitive e teorie filosofiche del significato, è opportuno considerare brevemente il ruolo del linguaggio nelle scienze cognitive stesse. Molti settori della scienza cognitiva, e, in particolare, l'IA nella sua forma "classica" hanno assegnato una posizione centrale al linguaggio nell'ambito dello studio dei fenomeni cognitivi. Già in un testo che è stato considerato da molti come una sorta di "atto di nascita" dell'IA, vale a dire l'articolo "Computing machinery and intelligence" di Alan Turing (1950), il comportamento

linguistico costituisce di fatto l'unico criterio in base al quale attribuire intelligenza a un dispositivo computazionale. Secondo Turing è legittimo attribuire intelligenza ad una macchina nel caso essa riesca a superare il cosiddetto gioco dell'imitazione. Al gioco partecipano tre "attori": a) un essere umano, b) una macchina calcolatrice, c) un altro essere umano, l'interrogante. L'interrogante non può vedere a) e b), non sa chi dei due sia l'uomo, e può comunicare con loro solo in maniera indiretta (ad esempio mediante una telescrivente, o un terminale video e una tastiera). L'interrogante deve sottoporre ad a) e a b) delle domande, con lo scopo di scoprire nel più breve tempo possibile chi dei due sia l'uomo e chi la macchina; a) si comporterà in modo da agevolare c), mentre b) dovrà rispondere in maniera tale da ingannare c) il più a lungo possibile. Secondo Turing, il problema di stabilire se le macchine possano pensare può essere riformulato chiedendosi se una macchina può ingannare un uomo nel gioco dell'imitazione, o, comunque, quanto a lungo può resistergli. Si tratta di quello che in seguito verrà detto *test di Turing*, e che è stato considerato per lungo tempo un valido criterio di adeguatezza per i programmi di IA.

Benché, come ho ricordato, l'articolo di Turing venga talvolta considerato uno dei punti di partenza della ricerca in IA, tuttavia in esso sono ancora assenti molti dei presupposti teorici che saranno alla base di questa disciplina. Tutto l'articolo, e in particolare la formulazione del test di Turing, sono ancora pesantemente condizionati da una prospettiva di tipo comportamentista. Turing non avanza alcuna ipotesi circa le strutture interne necessarie perché una macchina esibisca un comportamento intelligente (se non che tale macchina deve essere analoga a un calcolatore digitale del tipo di un calcolatore di von Neumann). In seguito, con gli sviluppi successivi dell'IA, l'egemonia linguistica che il test di Turing comporta viene in un certo senso ulteriormente rafforzata. Con l'istituirsi dell'IA come ambito disciplinare specifico, l'essenza dell'intelligenza viene individuata nell'elaborazione computazionale di rappresentazioni mentali che si suppongono dotate a loro volta di una forma di tipo, in senso lato, linguistico. Vale a dire, oltre a privilegiare l'uso del linguaggio come ambito di dati rilevante nell'attribuzione dell'intelligenza, la natura stessa della mente viene concepita nei termini di elaborazione computazionale di strutture di tipo linguistico. La forma più estrema di questa concezione può essere individuata nelle posizioni di J. Fodor (1975), che riconduce le attività mentali all'elaborazione di espressioni di un *linguaggio del pensiero*. Una assunzione analoga è tuttavia formulata anche nella *knowledge representation hypothesis* di Brian Smith (1982), che costituisce forse la formulazione più esplicita dei criteri metodologici che stanno alla base della ricerca sulla rappresentazione della conoscenza nell'IA di impostazione "tradizionale". Nel programma di ricerca dell'IA il problema di *come devono essere rappresentate* le informazioni di cui un sistema intelligente dispone gioca un ruolo assolutamente centrale. Il settore di indagine che va sotto il nome di *rappresentazione della conoscenza* ha come scopo l'individuazione e lo studio di linguaggi formali adatti a questo fine. Secondo la *knowledge representation hypothesis* di Brian Smith, ogni "sistema intelligente" deve incorporare un insieme di strutture di tipo, in senso lato, linguistico, tali che, in primo luogo, queste strutture, se esaminate da un osservatore esterno al sistema, devono poter essere interpretate come la rappresentazione della conoscenza di cui il sistema dispone. Inoltre, a prescindere da tale "attribuzione di significato" data dall'esterno, tali strutture devono essere manipolabili formalmente, in modo da svolgere un ruolo causale nel determinare il comportamento del sistema.

Non sempre è chiaro, nell'ambito dei dibattiti fondazionali sull'IA e sulle scienze cognitive, come vengano utilizzati concetti come simbolo, sistema di simboli o rappresentazione. Tuttavia, l'ipotesi alla base dell'IA simbolica tradizionale, secondo cui le rappresentazioni mentali sono dotate di una struttura di tipo linguistico può essere caratterizzata nei termini seguenti. In un sistema di rappresentazioni di tipo linguistico è in generale possibile distinguere tra un insieme di *simboli primitivi*, o *atomici*, e un insieme

di *simboli complessi*. Questi ultimi vengono generati a partire dai simboli primitivi mediante opportune regole di composizione sintattica. Di norma, a partire da un insieme finito di simboli primitivi, si può generare un insieme potenzialmente illimitato di simboli complessi. Tali strutture simboliche sono rappresentazioni, nel senso che ad esse è associato un significato. Nel caso dei simboli primitivi, il significato è associato loro in maniera diretta. Nel caso dei simboli complessi, il loro significato è funzione del significato dei simboli primitivi che vi compaiono e della struttura sintattica del simbolo stesso. Vale a dire, il significato di un simbolo complesso può essere ottenuto a partire dal significato dei suoi componenti per mezzo di regole semantiche che "operano in parallelo" alle regole di composizione sintattica. In ciò consiste il cosiddetto *principio di composizionalità* del significato. Il principio di composizionalità, chiaramente ispirato alla struttura delle lingue naturali e dei linguaggi formali, viene considerato da alcuni come il tratto principale che distingue le spiegazioni della mente e della visione delle strutture mentali propria della scienza cognitiva classica (si veda ad esempio Fodor e Pylyshyn 1988). Un sistema simbolico di questo tipo comprende inoltre delle regole di inferenza (intese in senso lato) che consentono di manipolare i simboli in modo da ottenere le conseguenze desiderate. Tali regole di inferenza devono tenere conto della struttura dei simboli, ed elaborarli in accordo al loro significato. E' appunto il fatto di disporre di regole di questo tipo ciò che rende il sistema automatizzabile: la strutture simboliche possono essere elaborate in maniera formale da un dispositivo meccanico che opera in accordo alle regole sintattiche, senza che debba "conoscere" direttamente il significato dei simboli.

Tra i punti che caratterizzano l'orientamento cognitivo in teoria del significato, Marconi (1992) individua le tesi seguenti. In primo luogo, la semantica va intesa come teoria della comprensione, e la comprensione è un'attività di tipo mentale. Una spiegazione della comprensione del linguaggio deve essere fornita nei termini di quali strutture mentali vengano costruite quando viene compreso un enunciato. Ciò che garantisce la comunicazione tra soggetti diversi è la convergenza tra le loro strutture mentali, e non l'accesso a un dominio di significati oggettivi di natura extra mentale. Nei modelli sviluppati nell'ambito dell'IA simbolica e delle scienze cognitive di impostazione tradizionale le strutture mentali generate durante la comprensione vengono assimilate a espressioni simboliche di un sistema di rappresentazione del tipo sopra descritto.

Non tutti gli sviluppi dell'IA sono caratterizzati da una forte connotazione psicologica in senso stretto. Ad esempio, spesso non viene richiesta alcuna adeguatezza empirica dei modelli proposti. In ogni caso, anche nelle versioni meno orientate psicologicamente dell'IA (come nell'IA di impostazione logica - si veda ad esempio Frixione 1994a), ci si pone comunque l'obiettivo di individuare rappresentazioni finite dei significati linguistici, esigenza legata appunto agli scopi computazionali della disciplina: le rappresentazioni dei significati devono essere tali da poter essere memorizzate ed elaborate da dispositivi di calcolo dotati di risorse finite. Utilizzando una terminologia di origine chomskiana, si può dire che scopo della semantica cognitiva è elaborare modelli della *competenza semantica*. La nozione di *competenza* trae origine dalla linguistica chomskiana (ed è stata elaborata in riferimento a problemi di carattere sintattico piuttosto che di carattere semantico). Secondo Chomsky la competenza linguistica consiste in ciò che un parlante conosce della propria lingua. Si tratta tuttavia di un parlante *idealizzato*, e, in questo senso, la nozione di competenza viene contrapposta alla nozione di *esecuzione*, che concerne invece la pratica linguistica effettiva dei parlanti reali. A differenza della competenza, l'esecuzione è sottoposta a vincoli di natura extra-linguistica. "La teoria linguistica si occupa principalmente di un parlante-ascoltatore ideale, in una comunità linguistica completamente omogenea, il quale conosce perfettamente la sua lingua e non è influenzato da condizioni grammaticalmente irrilevanti quali le limitazioni di memoria, le distrazioni, i cambiamenti di attenzione e di interesse, e gli errori (casuali o caratteristici) nell'applicazione della propria conoscenza della lingua nel corso

dell'esecuzione effettiva [...]. Per considerare l'*esecuzione* linguistica effettiva, dobbiamo considerare l'interazione di vari fattori, e la competenza sottostante del parlante-ascoltatore non è che uno di essi." (Chomsky 1965, pp. 44-45 della trad. it., corsivi miei). Per quanto idealizzata, tuttavia, la competenza è sempre competenza *di un parlante*, comunque individuale e finito. L'intento di un modello computazionale della competenza semantica è quello di ipotizzare le rappresentazioni e le procedure che consentano a un parlante di esibire determinati comportamenti linguistici.

Lo sviluppo in senso cognitivo di una teoria del significato risulta dunque incompatibile con gli assunti antipsicologisti della tradizione analitica. Tuttavia un approccio al significato basato sulle posizioni delle scienze cognitive classiche, sebbene rifiuti la caratterizzazione extrapsicologica dei significati, condivide la natura puramente linguistica del pensiero. Sebbene il significato venga in un certo senso trasferito "all'interno della mente", è la mente stessa ad essere concepita in termini linguistici<sup>[1]</sup>. Resta quindi in un certo senso valido il postulato dell'imprescindibilità del linguaggio nella caratterizzazione della conoscenza e del pensiero.

### 3. Oltre il linguaggio

Durante l'ultimo decennio l'evoluzione di una parte considerevole delle scienze cognitive è andata nella direzione dello studio di fenomeni e dell'elaborazione di modelli in cui l'egemonia degli aspetti linguistici nell'ambito dello studio della mente è stato posto per molti versi in discussione. Sono stati proposti da più parti modelli della mente ispirati più direttamente alle basi biologiche del pensiero e del comportamento (per una lettura esplicitamente "anti linguistica" di queste tendenze si veda ad esempio Parisi 1992). Le ragioni di tale evoluzione sono molteplici. Nelle scienze cognitive, la prospettiva simbolico-linguistica classica ha condizionato il dominio dei fenomeni presi in considerazione. Ha comportato che venisse posta l'enfasi su certe forme di conoscenza e su certi tipi di fenomeni cognitivi a scapito di altri. Sono stati di fatto privilegiati i fenomeni cognitivi di "alto livello" (come il ragionamento logico e il linguaggio naturale), mentre aspetti di "basso livello" come percezione e aspetti motori sono stati emarginati dal filone centrale della ricerca in IA. L'enfasi sugli aspetti linguistici dei fenomeni cognitivi ha reso problematico prendere in considerazione sistemi intelligenti che non adoperano forme di comunicazione di tipo linguistico, come gli animali o i bambini piccoli. In generale, risultava problematico prendere in considerazione tipi di conoscenza che non presupponessero una forma di rappresentazione logico-linguistica. Esistono poi problemi più specifici posti dai modelli di tipo simbolico. Ad esempio, i sistemi simbolici tollerano male gli errori, nel senso che se i loro *input* si discostano anche di poco da quelli previsti da chi ha progettato il sistema, le prestazioni crollano drammaticamente. Questo contrasta con la capacità che hanno gli esseri umani e i sistemi intelligenti reali di "arrangiarsi" anche in condizioni tutt'altro che ottimali. Inoltre sono poco resistenti ai danni: ogni danneggiamento locale di un sistema simbolico comporta conseguenze globali disastrose. Vi sono poi i problemi dell'apprendimento e dell'evoluzione di un sistema intelligente, e delle sue interrelazioni con l'ambiente circostante, che restano per molti versi problematici per la scienza cognitiva simbolica. Infine, nell'approccio linguistico classico le capacità cognitive più elevate vengono in un certo senso assunte come date, senza che vengano indagate le condizioni che ne consentono l'emergere da forme di organizzazione mentale e biologica più semplici.

Queste ed analoghe considerazioni hanno fatto sì che, come si è detto, le scienze cognitive superassero in parte l'egemonia linguistico-simbolica dell'IA tradizionale, elaborando paradigmi che postulano altri tipi di rappresentazioni mentali, o che negano esplicitamente la centralità del concetto di rappresentazione per una teoria della mente. Il fenomeno più noto in questa prospettiva è costituito certamente dall'emergere del

*connessionismo* (Rumelhart *et al.* 1986; Smolensky 1988). Il paradigma connessionista si basa su modelli computazionali, le cosiddette *reti neurali*, caratterizzati da un alto grado di parallelismo ed ispirati, in senso lato, alle proprietà del sistema nervoso. Una rete neurale è costituita da un insieme di *unità*, collegate tra loro tramite una fitta rete di *connessioni*. In base all'analogia neurale, le unità corrispondono ai neuroni, e le connessioni sono l'equivalente delle sinapsi. Ciascuna unità costituisce un singolo dispositivo di calcolo che esegue forme molto semplici di calcolo, ed opera in parallelo con tutte le altre unità della rete. La potenza computazionale del sistema deriva appunto dal parallelismo e dal grande numero delle unità e delle connessioni.

Va precisato che, di norma, i modelli connessionisti non devono essere intesi come descrizioni adeguate dal punto di vista anatomico e neurofisiologico della struttura del sistema nervoso. Ad esempio, le analogie tra i neuroni reali e le unità di una rete neurale non vanno al di là di alcune caratteristiche estremamente astratte e generiche. La struttura del sistema nervoso svolge di solito soltanto un ruolo ispiratore nei confronti della struttura dei modelli connessionisti. Sono evidenti tuttavia le differenze con l'ispirazione simbolico-linguistica dei modelli classici: si tratta comunque di una visione "dal basso" della mente e del pensiero, che assume come dato di partenza la struttura del substrato neurologico.

Il connessionismo condivide con l'IA e con l'approccio classico alle scienze cognitive una concezione della mente di tipo *rappresentazionale*. Il connessionismo riconosce cioè al concetto di rappresentazione mentale un ruolo esplicativo nell'ambito delle generalizzazioni di una teoria delle attività mentali. Di norma tuttavia le rappresentazioni nei sistemi connessionisti hanno una natura molto diversa dalle rappresentazioni dell'IA classica. Ciò è particolarmente evidente nei modelli connessionisti di tipo subsimbolico e distribuito, quali sono stati teorizzati ad esempio da Paul Smolensky (1988). Nei modelli subsimbolici le rappresentazioni mentali vengono ricondotte a distribuzioni di valori di attivazione sulle unità di una rete neurale. Si tratta anche in questo caso di una caratteristica ispirata alle proprietà del sistema nervoso: questo è il modo in cui si ipotizza che la conoscenza sia rappresentata nel cervello, distribuita su *pattern* di attivazione di grandi insiemi di neuroni. Ogni singola unità di una rete connessionista (un "subsimbolo" appunto) non costituisce a sua volta una rappresentazione, ma è soltanto un "ingrediente" che consente alle rappresentazioni di emergere a un livello di organizzazione superiore. Secondo Smolensky, le strutture simboliche dell'IA e delle scienze cognitive tradizionali (come ad esempio sistemi a regole di produzione, schemi, *frame* e così via) non devono essere intese come entità primitive e non ulteriormente analizzabili. Esse vanno considerate piuttosto come effetti emergenti, o descrizioni approssimate, di fenomeni che sono descrivibili in maniera esatta soltanto al livello dei subsimboli. Secondo una metafora utilizzata dallo stesso Smolensky, i modelli dell'IA classica sono paragonabili a descrizioni di fenomeni fisici effettuate a livello macroscopico. Si tratta di descrizioni approssimate che valgono solo in "condizioni ideali". Se si vuole ottenere una descrizione completa ed esatta dei fenomeni cognitivi si deve scendere a un livello di analisi inferiore, utilizzando appunto gli strumenti offerti dal connessionismo subsimbolico. Se dunque si ha ancora a che fare con una concezione rappresentazionale della mente, in essa tuttavia il concetto di rappresentazione non viene assunto come primitivo, ma viene ricondotto a meccanismi cognitivi sottostanti di tipo non rappresentazionale.

Se il connessionismo condivide almeno in parte con l'IA classica una concezione rappresentazionale dei fenomeni mentali, recentemente sono state elaborate posizioni più radicali, che assumono un atteggiamento di tipo *eliminativista* nei confronti delle rappresentazioni mentali (la cui controparte filosofica può essere identificata nelle posizioni di filosofi della mente eliminativisti come Paul Churchland). In base a queste posizioni, viene negato al concetto stesso di rappresentazione mentale qualsiasi ruolo in una spiegazione scientifica della mente o nella simulazione di attività intelligenti. Un

esempio in questa direzione è costituito dalle posizioni del *planning reattivo* e dalla cosiddetta *nuova robotica*, il cui rappresentante più noto è Rodney Brooks (si veda ad esempio Brooks 1991). Secondo Brooks i tentativi di costruire macchine intelligenti in grado di operare interagendo con il mondo reale basandosi sulle tecniche tradizionali dell'IA sono destinati all'insuccesso. Nel settore della robotica, l'IA ha posto l'enfasi sulla messa a punto di formalismi di rappresentazione per la pianificazione e per il ragionamento su eventi e azioni. Tuttavia, molte cose devono già funzionare perfettamente in un organismo, sia esso naturale o artificiale, prima che abbia senso tentare di fargli effettuare inferenze complesse dal punto di vista logico. Se si guarda all'evoluzione naturale degli esseri viventi, è facile constatare che sono stati necessari tempi enormemente lunghi perché si sviluppassero specie animali in grado di percepire il mondo circostante e di muoversi con sicurezza nel loro ambiente. A partire da queste, le forme più elevate di organizzazione mentale, sino alla razionalità e all'uso del linguaggio, si sono evolute in tempi relativamente brevi. Quindi, l'essenza dell'intelligenza non risiederebbe tanto nella possibilità di disporre e di elaborare rappresentazioni e strutture mentali complesse, quanto piuttosto nella capacità di interagire con la complessità del mondo esterno. Il paradigma per la simulazione dell'intelligenza diventa dunque per Brooks lo studio delle capacità di organismi relativamente semplici e primitivi come gli insetti.

Oltre alla nuova robotica, sono stati proposti altri tipi di modelli cognitivi di ispirazione biologica, e orientati più allo studio delle capacità di sistemi intelligenti inferiori e agli aspetti evolutivi dell'intelligenza che alla razionalità umana. Tra essi, ad esempio, gli algoritmi genetici (Goldberg 1989) o i modelli basati sugli automi cellulari (Toffoli e Margolis 1987). In generale, per una rassegna sui modelli cognitivi basati su vari tipi di sistemi dinamici complessi si veda (Serra e Zanarini 1990). Queste linee di ricerca si intersecano con quell'ambito di studio noto come *vita artificiale*, in cui ci si propone la sintesi di comportamenti biologici per mezzo di strumenti di tipo computazionale (e in cui l'intelligenza e i fenomeni mentali sono visti appunto come un particolare tipo di fenomeno biologico - si veda ad esempio *Sistemi intelligenti* 1992).

Vale la pena ricordare che l'egemonia del linguaggio nell'ambito delle scienze cognitive era stata posta in discussione già prima dell'emergere del paradigma connessionista nel corso dello scorso decennio. In primo luogo, il connessionismo stesso è erede diretto della cibernetica degli anni cinquanta e sessanta. Inoltre, il predominio del linguaggio come modello della mente era stato posto in discussione nel corso degli anni settanta in ambiti quali il dibattito sul ruolo cognitivo delle immagini mentali e sulla loro irriducibilità a rappresentazioni di tipo linguistico (si veda ad esempio Block 1981a). Anche i modelli basati su immagini mentali suscitano oggi un rinnovato interesse in IA (si veda ad esempio *Computational Intelligence* 1993).

#### **4. Calcolare il riferimento**

Le linee di ricerca cui ho fatto cenno nel paragrafo precedente sono in forte contrasto con le assunzioni dummettiane (e fregeane) sulla natura intrinsecamente linguistica del pensiero: non soltanto lo studio del pensiero viene intrinsecamente correlato allo studio del processo psicologico del pensare, ma viene posta in discussione la natura linguistica del pensiero stesso. Vengono riconosciute forme di pensiero (ossia, attività mentali che possono essere ricondotte all'attività di pensare, o che di questa attività sono un prerequisito immediato) che prescindono da forme di rappresentazione linguistica, o persino da forme di rappresentazione di qualsiasi tipo. In questa prospettiva, il ruolo del linguaggio nell'ambito della mente e dei fenomeni cognitivi viene fortemente ridimensionato. Inoltre, le attività mentali di tipo linguistico e il linguaggio stesso

vengono ricondotte e radicate a strutture cognitive e ad attività mentali di tipo prelinguistico.

Nella prospettiva dello sviluppo di una teoria del significato, è interessante notare come il progetto di una semantica, una volta scelta una prospettiva di tipo cognitivo, non possa prescindere, almeno per alcuni aspetti, dal prendere in considerazione i meccanismi non linguistici all'opera nella mente. Questo avviene tipicamente nel caso del lessico. Il problema della semantica lessicale può essere ricondotto a due aspetti distinti, che chiamerò *aspetto inferenziale* e *aspetto referenziale* (questa distinzione è stata introdotta da Marconi a proposito della competenza lessicale - si veda ad esempio Marconi 1989). La componente *inferenziale* del lessico concerne quel complesso insieme di connessioni che esprimono relazioni di tipo non logico che sussistono tra le parole di una lingua. Esempi di relazioni di questo genere sono che i cani sono mammiferi, che gli uccelli di solito volano, che un padre è un essere umano di sesso maschile il quale ha dei figli che sono, a loro volta, esseri umani, che se qualcuno corre allora (di solito) si sposta, e così via. Tali informazioni vengono abitualmente utilizzate per eseguire inferenze, e fanno parte, a tutti gli effetti, del significato lessicale: "se uno non sa che, per mangiare, normalmente, si deve aprire la bocca diciamo che non sa cosa *significa* mangiare" (Marconi 1989, p. 79). L'*aspetto referenziale* del lessico è relativo alle relazioni che sussistono tra parole e oggetti, eventi e situazioni nel mondo. Ad esempio, fa parte della competenza referenziale associata a parole come "cane" o "correre" saper dire se un certo animale è un cane, o saperlo distinguere da un gatto, saper riconoscere qualcuno che corre, e così via. La competenza referenziale è ciò che consente, si potrebbe dire, di "calcolare" i valori della funzione interpretazione per i simboli primitivi non logici di un linguaggio, e che consente quindi, in ultima istanza, di individuare le condizioni di verità effettive di un enunciato.

Sia l'*aspetto inferenziale* che l'*aspetto referenziale* del lessico sono, in diversa misura, problematici per le teorie semantiche di impostazione logico-filosofica. Il primo in quanto gli strumenti logici tradizionali (quali i postulati di significato classici) risultano insufficienti a caratterizzare tutti gli aspetti inferenzialmente rilevanti del lessico; il secondo in quanto nelle semantiche di tipo logico l'interpretazione degli elementi del lessico (almeno di quelli che corrispondono a simboli primitivi non logici) viene semplicemente assunta come data, e non viene fornito alcun modo perché possa essere calcolata. Sono esattamente questi limiti che hanno comportato, come si diceva, un avvicinamento tra teorie filosofiche del significato linguistico e scienze cognitive. Per quanto riguarda gli aspetti inferenziali del lessico, tra i fattori che hanno motivato tale avvicinamento vi è ad esempio l'esigenza di caratterizzare i concetti lessicali non soltanto nei termini di condizioni necessarie o sufficienti (che è quanto consentono i tradizionali postulati di significato), ma anche nei termini di tratti tipici. Per un verso, le caratteristiche tipiche che contribuiscono al significato dei concetti lessicali non sono certamente riconducibili a informazioni di tipo analitico. Di conseguenza, risulta praticamente impossibile separare in maniera definitiva ciò che costituisce il significato di un termine dalle conoscenze di tipo fattuale che lo concernono, e soprattutto risulta impossibile praticare tale distinzione a prescindere dalla competenza di un parlante specifico. Le rappresentazioni cognitive dei parlanti diventano quindi parte integrante del significato (si veda ancora Marconi 1989). Inoltre, l'IA e le scienze cognitive di impostazione simbolica hanno messo a disposizione formalismi, ispirati a considerazioni e a dati di ordine psicologico, che consentono di rappresentare i concetti lessicali nei termini di tratti prototipici anziché di condizioni necessarie e/o sufficienti. Si possono ricordare i *frame* (Minsky 1975) e molti sistemi a rete semantica (si vedano ad esempio Sowa 1991 e gli articoli in Brachman e Levesque 1985), oppure, sul versante logico, le logiche non monotone, ossia formalismi logici che consentono di rappresentare generalizzazioni che ammettono eccezioni (Brewka 1991).



Dal punto di vista inferenziale quindi l'IA tradizionale di impostazione logico-linguistica ha fornito certamente strumenti importanti per affrontare il problema del lessico (si veda Frixione 1994a per una rassegna). Diversa è la situazione per gli aspetti referenziali. Se infatti la prospettiva linguistica dell'IA classica permette di affrontare importanti aspetti della competenza inferenziale, rispetto alla competenza referenziale l'uso di modelli di tipo logico-linguistico equivale semplicemente a spostare il problema. Una volta che il significato degli elementi del lessico viene ricondotto a rappresentazioni mentali di tipo linguistico, resta da risolvere il problema dell'ancoramento (*grounding*) di queste ultime alla realtà extra mentale (nel senso di Harnad 1990). Resta cioè il problema di caratterizzare il modo in cui è dato il riferimento delle rappresentazioni mentali stesse, problema che non sembra essere affrontabile nel quadro di una spiegazione esclusivamente simbolico-linguistica della mente.

In una prospettiva di tipo cognitivista, il problema della competenza referenziale può essere visto come il problema di stabilire "come è calcolata" dal punto di vista cognitivo la funzione interpretazione per i termini del lessico (o, analogamente, per i simboli primitivi del "linguaggio del pensiero"). In un certo senso, quindi, la funzione interpretazione deve essere computabile (parziale), e deve essere possibile sapere "come è fatta", cioè deve essere possibile identificare una procedura che la calcola. Ma, poiché si tratta di una funzione che mette in relazione strutture linguistiche di rappresentazione con oggetti del mondo e viceversa, tale procedura non può essere definita esclusivamente nei termini di altre strutture linguistiche, poiché questo causerebbe un regresso all'infinito. Se si vuole spiegare in che modo le parole (di una lingua naturale o del "linguaggio del pensiero") si riferiscono a oggetti, eventi o fenomeni extra linguistici, non si può pensare di utilizzare esclusivamente altre entità che, a loro volta, si suppongono dotate di significato (Frixione *et al.* 1989, Frixione e Spinelli 1992). La funzione interpretazione deve quindi essere calcolata, almeno per alcuni dei termini primitivi, da qualche dispositivo computazionale di tipo non linguistico-simbolico. Poiché per spiegare la competenza referenziale il livello linguistico-simbolico deve essere abbandonato, un livello subsimbolico (del tipo ad esempio di quello teorizzato da Smolensky) potrebbe essere un buon candidato per modellare quella attività non simbolica della mente che fa sì che i simboli si riferiscano a entità extra mentali. Non è in linea di principio necessario che la computazione subsimbolica debba essere in qualche senso di tipo "neurale", o ispirata alla struttura anatomica del sistema nervoso. Né, a maggior ragione, è necessario ipotizzare che gli strumenti computazionali necessari debbano andare in qualche modo oltre i confini della teoria della computabilità effettiva. Si tratta esclusivamente di negare che l'ipotesi di una struttura linguistica della mente sia sufficiente per affrontare il problema degli aspetti referenziali del lessico. Diventano rilevanti in questa prospettiva (almeno per vaste classi di termini) i risultati delle ricerche sui meccanismi della percezione, e sulle sue basi neurofisiologiche. Diventano cioè rilevanti quegli aspetti psicologici "di basso livello" che in qualche senso "precedono" (sia ontogeneticamente che filogeneticamente) l'emergere del linguaggio, e che sono presumibilmente in parte indipendenti da esso.

Quanto detto non va inteso nel senso della proposta di una concezione "verificazionista" della semantica, in cui il significato dei termini venga ricondotto a criteri di tipo osservativo per la loro applicabilità. E' evidente che solo una parte dei termini del linguaggio possono essere ancorati al loro riferimento tramite i meccanismi percettivi. E' altresì ovvio che i meccanismi del riferimento in un sistema simbolico come il linguaggio naturale costituiscono un fenomeno estremamente complesso. Usando un'immagine wittgensteiniana, si può dire che le parole utilizzano molti tipi diversi di "uncini" per ancorarsi alla realtà (e la stessa cosa vale, ovviamente, anche per le rappresentazioni mentali). La competenza referenziale si basa su differenti modalità (sensorie, motorie, eccetera) variamente correlate tra loro, e dipende in larga misura da componenti di natura

culturale e quindi, in parte, simbolica. Tuttavia, il ricorso a un livello di attività cognitiva di tipo non linguistico è inevitabile. La competenza referenziale è legata alle attività motorie e sensoriali non solo per la ragione ovvia che molti referenti sono percepiti come oggetti dei sensi, ma anche perché l'interazione (simbolica e non simbolica) con altri membri della comunità dei parlanti è fondamentale sia nel apprendere l'uso di un sistema di simboli, che nell'uso referenzialmente adeguato di molti termini astratti. L'idea che un sistema intelligente possa usare simboli per riferirsi, ad esempio, a numeri, leggi o sentimenti (qualunque cosa "riferirsi" possa significare in casi del genere) è inconcepibile senza prendere in considerazione qualche forma di interazione non simbolica con il mondo e con altri sistemi intelligenti. Lo stesso test di Turing è insufficiente, oltre che come criterio generale per l'attribuzione di intelligenza a un sistema, anche come criterio per l'attribuzione della competenza linguistica, qualora si prenda in considerazione anche la competenza referenziale (ed è decisamente poco adeguato ritenere linguisticamente competente un sistema che non sia in grado di porre in relazione il linguaggio con il mondo).

## 5. Stati mentali e rappresentazioni private

Come abbiamo visto, gli sviluppi in senso cognitivo delle teorie filosofiche del significato sono in contrasto con gli assunti antipsicologisti alla base della filosofia del linguaggio di Frege. E' rilevante quindi esaminare in che misura le posizioni dello psicologismo "cognitivo" contemporaneo si possano considerare un superamento dell'antipsicologismo fregeano. Le posizioni di Frege relative allo psicologismo in semantica sono strettamente correlate alle sue posizioni rispetto al problema dei fondamenti delle matematiche, e al rifiuto di una fondazione psicologista delle matematiche (dell'assunzione cioè che gli enti matematici possano essere identificati con oggetti di tipo mentale). Seguendo Thomason (1978), prenderò dunque in considerazione alcune delle motivazioni fregeane contro una fondazione di tipo psicologico della matematica. Nei *Grundlagen* (1884) Frege mette in evidenza alcune conseguenze inaccettabili derivanti dal fatto di assumere che gli oggetti matematici siano definiti come entità di natura mentale. Ad esempio, afferma Frege, se una dimostrazione fosse un processo mentale, nel caratterizzare una dimostrazione del teorema di Pitagora si dovrebbe tener conto del contenuto di fosforo del cervello di chi la esegue. Inoltre, una posizione di tipo psicologista comporterebbe una definizione del tutto inadeguata, ad esempio, del concetto stesso di numero. Ciascuno ha idee differenti a proposito dei numeri. Quindi, se i numeri fossero rappresentazioni mentali, dovrebbero esserci numeri diversi per ciascun essere umano: ad esempio, si sarebbe costretti ad ammettere che esiste un numero due diverso per ciascuno di noi. Da un altro punto di vista, però, ci sarebbero troppo pochi numeri. Infatti ogni essere umano, nel corso della sua esistenza, può comunque pensarne soltanto un numero finito. Queste considerazioni sono molto simili a quelle che Frege adopera per motivare la natura extrapsicologica ed oggettiva dei significati linguistici (si veda il paragrafo 1). Ad esempio, le stesse difficoltà poste dal teorema di Pitagora si applicano all'inferenza dei nessi di conseguenza logica tra enunciati. Che a menti diverse corrispondano rappresentazioni dei numeri diverse è analogo al problema della convergenza delle rappresentazioni mentali che parlanti diversi associano a certe espressioni linguistiche (ad esempio, alle costanti logiche). Infine, il fatto che una mente possa rappresentarsi solo un insieme finito di numeri pone difficoltà analoghe alla spiegazione delle capacità generative delle lingue naturali: si deve poter caratterizzare il significato di un insieme potenzialmente infinito di enunciati, anche di quegli enunciati che non sono mai stati pronunciati o pensati da nessuno. Tuttavia, afferma Thomason, questo tipo di argomentazioni si può considerare superato dalle posizioni della moderna psicologia cognitiva, e dalle teorie psicologiche di impostazione

computazionale. Tali teorie si basano infatti sulla distinzione tra competenza ed esecuzione. La competenza viene caratterizzata nei termini di stati computazionali astratti, definiti funzionalmente. Ad esempio, in base alle posizioni funzionaliste sulla natura degli stati mentali (posizioni che sono condivise da gran parte degli approcci cognitivisti classici e che, in qualche misura, sono compatibili anche con molti degli approcci connessionisti), si assume che sia possibile caratterizzare la competenza in maniera del tutto indipendente dalle proprietà fisico-chimiche e anatomiche del sistema nervoso. Quindi, in questa prospettiva, la quantità di fosforo presente nel cervello di chi esegue una certa inferenza non è rilevante nella caratterizzazione dell'inferenza stessa. Ciò risponde alla prima obiezione. Per quanto riguarda il fatto che esseri umani diversi hanno idee diverse sui numeri (o sul significato di certe parole), tali differenze possono essere limitate all'ambito dell'esecuzione, e non riguardare una teoria della competenza. Se consideriamo la competenza come caratterizzabile nei termini di una macchina astratta, ad esempio una macchina di Turing, allora, data la rappresentazione di un numero in due realizzazioni fisiche diverse della stessa macchina, si può caratterizzare in modo esatto la loro equivalenza in termini funzionali. Rispetto al problema della finitezza, infine, una teoria della competenza formulata nei termini di dispositivi di calcolo che, sebbene finiti, siano astratti e idealizzati (come appunto le macchine di Turing) consente di ipotizzare, ad esempio, una memoria potenzialmente espandibile indefinitamente, in maniera che (in un lasso comunque finito di tempo) possa essere generato qualsiasi numero computabile<sup>[2]</sup>

A prescindere dalla praticabilità del funzionalismo classico (le cui posizioni sono state in qualche misura poste in discussione in molti degli sviluppi più recenti delle scienze cognitive - si veda ad esempio il Post Scriptum critico di Edelman 1992), qualsiasi teoria della mente deve impegnarsi ad individuare criteri di identità tra stati mentali a livello di tipi, i quali consentano di stabilire a quali condizioni due stati mentali (anche di soggetti diversi) appartengano alla stessa classe. I criteri del funzionalismo classico sono solo un tipo particolare di criteri di questo genere. In generale, il fatto di disporre di una teoria della mente consente di escludere appunto una totale incommensurabilità tra stati mentali di individui differenti. Ciò non toglie ovviamente che l'attribuzione di uno stato mentale a un dato soggetto cognitivo, e la categorizzazione stessa degli stati mentali, siano comunque operazioni altamente ipotetiche, e continuamente soggette a possibilità di revisione. Il punto è che gli stati mentali sono comunque costrutti teorici di una teoria della mente, alla stregua dei costrutti teorici utilizzati in qualsiasi disciplina scientifica. Come tali, sono sottodeterminati dalla base empirica della teoria stessa. In questa prospettiva, attribuire un ruolo agli stati mentali dei parlanti nell'ambito di una teoria del significato non equivale tuttavia a ridurre a un miracolo la possibilità della comunicazione linguistica. La possibilità che parlanti diversi assegnino "lo stesso significato" a un'espressione non è incompatibile con la natura (almeno parzialmente) mentale dei fenomeni semantici qualora si accettino alcune assunzioni che sono ampiamente condivise dalle teorie psicologiche contemporanee.

Vale la pena di notare come, nella tradizione analitica, è la stessa possibilità di introdurre costrutti teorici in una teoria del significato che è stata talvolta posta in discussione. Le posizioni antipsicologistiche, soprattutto sotto l'influenza del neoempirismo logico, hanno spesso assunto una connotazione marcatamente comportamentista e verificazionista. Una posizione di questo tipo emerge ad esempio in quanto afferma Cresswell a proposito dello *status* degli argomenti delle espressioni di atteggiamento proposizionale. A proposito della teoria dei significati strutturati proposta da Cresswell, BŠuerle e Cresswell (1988) sostengono che un significato strutturato deve essere "una 'entità pubblica' nel senso che la sua struttura dipende dalla struttura dell'enunciato usato nell'attribuzione della credenza" (p. 499). Queste affermazioni sono motivate in base a considerazioni di ascendenza fregeana: "I significati non possono essere entità private.

Due persone possono avere la stessa credenza, e la credenza che essi hanno in comune deve essere qualcosa cui devono poter fare pubblico riferimento" (ivi, p. 498). Già Fodor (1978), a proposito della teoria dell'isomorfismo intensionale di Carnap, da cui Cresswell prese le mosse, aveva notato la matrice comportamentista alla base di questo tipo di argomenti. Lo scopo dell'analisi semantica degli enunciati di atteggiamento proposizionale viene visto in ultima istanza nell'individuazione di criteri per l'attribuzione degli atteggiamenti stessi, criteri che devono essere intersoggettivamente accessibili e accettabili in una prospettiva verificazionista, in quanto non possono far riferimento a entità di tipo non osservativo (ad esempio di tipo mentale). E' opportuno però osservare che l'attribuzione, ad esempio, di una credenza a un dato soggetto è sempre un'operazione a carattere ipotetico, sottodeterminata dal comportamento (linguistico o di altro tipo) del soggetto stesso. Non è quindi ingiustificato riferirsi a costrutti teorici e ad entità non "osservabili" (come appunto le rappresentazioni o gli stati mentali).

Un'argomentazione contro lo psicologismo in semantica diversa da quella basata sul carattere privato degli stati mentali ma in qualche misura imparentata ad essa si trova in (Dummett 1988 - cap. 12). Dummett argomenta che concetti e pensieri (nell'accezione fregeana) non possono essere identificati con contenuti di coscienza, in quanto i contenuti di coscienza non hanno una struttura abbastanza ricca e non possono esibire le proprietà necessarie per svolgere le funzioni di significati linguistici. Dummett caratterizza uno stato di coscienza associato a un'espressione linguistica come ciò che "viene in mente" a un parlante quando, poniamo, ascolta o proferisce una certa parola o un certo enunciato. Ora, avverte Dummett, "il capire una parola come esprime un certo concetto non può essere descritto dicendo che la parola fa venire in mente al parlante il concetto che le è associato, poiché non esiste nulla di simile al venire in mente di un concetto: una melodia, un nome, una scena o un profumo ricordati possono venire alla mente, ma un concetto non è il genere di cose di cui si può sensatamente dire che vengono in mente" (Dummett 1998, pp. 105-6 della trad. it.). E ancora, se anche possedere un concetto fosse riconducibile ad essere in un certo stato di coscienza, al "venire in mente" di un'immagine mentale, tuttavia "avremmo egualmente bisogno di chiarire in che cosa consiste *applicare* quel concetto" (ivi, p. 106). A proposito di chi possiede un concetto, infatti, "la sua capacità di applicare il concetto *costituisce* il possesso di quel concetto: non fa differenza se, quando lo applica, gli viene in mente qualcosa [...] qualunque fosse la cosa che potrebbe venirgli in mente, non sarebbe tuttavia un concetto" (ivi). Considerazioni simili vengono avanzate a proposito dei pensieri: "Un pensiero può essere afferrato solo come qualcosa di complesso: avere un pensiero è un modo di attivare capacità che coinvolgono la comprensione di ciascuno dei concetti che quel pensiero comporta. Non è solo che [...] nessuno può avere, ad esempio, il pensiero 'questa rosa ha un profumo dolce' senza sapere, all'incirca, che cosa sono le rose o senza essere in grado di avere altri pensieri sulle rose o senza essere conscio di altre cose con un profumo dolce o di qualcosa che non lo ha. Il punto è, piuttosto, che non possiamo apprendere alcun pensiero senza apprenderne la complessità, senza, ad esempio, concepire noi stessi come pensanti alla rosa, o come pensanti della rosa qualcosa che può essere vero di alcune rose e falso di altre" (ivi, pp. 107-108). E' evidente che il contenuto di coscienza che accompagna un enunciato non gode di queste caratteristiche. Da ciò consegue, secondo Dummett, la necessità di estromettere i significati dalla mente, secondo i principi posti a fondamento della tradizione analitica.

Ma il punto è che le rappresentazioni, e, più in generale, gli stati mentali come vengono concepiti dalle scienze cognitive non sono identificabili con segmenti del flusso di coscienza. E' certamente vero che gli stati di coscienza associati dai parlanti alla parola "rosa" non possono essere un candidato plausibile a spiegare che cosa "rosa" significhi. Ma, nelle scienze cognitive, la rappresentazione mentale del concetto "rosa" (sia essa modellabile con un *frame*, o con un insieme di postulati di significato non monotoni, o

come *pattern* di attivazione su una rete neurale, o in base a qualche altro tipo di modello) non è certamente identificabile con uno stato qualitativo presente alla coscienza del parlante. E' certamente vero che possedere un concetto comporta saperlo applicare. Ma la rappresentazione mentale di un concetto deve proprio rendere conto di come il concetto stesso venga applicato, sia in relazione ad altri concetti che in relazione alla realtà extra-mentale, e questo sia in rapporto al comportamento linguistico che ad altre forme di comportamento. Questo vale, a maggior ragione, per la rappresentazione mentale di un enunciato. Nella scienza cognitiva classica viene ipotizzato che la rappresentazione di un enunciato sia dotata di una struttura compositiva in grado di rendere conto della struttura dell'enunciato stesso. In ogni caso, in qualsiasi approccio alle scienze cognitive, i costrutti mentali in qualche senso associati a un dato enunciato devono essere tali da consentire di spiegare le sue relazioni sistematiche con altri enunciati e con il comportamento (linguistico e non linguistico) del parlante.

Tutto questo non ha necessariamente a che fare con ciò di cui il soggetto è cosciente quando adopera certe espressioni linguistiche. I meccanismi mentali alla base dell'uso dei concetti non possono certamente essere pienamente accessibili al soggetto cognitivo. In questo senso, la riduzione degli stati mentali a stati di coscienza è frutto di una visione inadeguata della mente, viziata da ciò che potremmo definire una sorta di "verificazionismo solipsistico". E' come se si trattasse di una estensione, solipsistica appunto, del pregiudizio empiristico: del mentale si può concepire solo ciò che si può esperire pubblicamente (il comportamento), o, al più, ciò che si può esperire in maniera privata (gli stati qualitativi di coscienza). Ma le strutture e i processi mentali ipotizzati dalle scienze cognitive non sono qualcosa che percepiamo o conosciamo privatamente. Esse, caso mai, sono all'opera nel nostro percepire e nel nostro conoscere, e, in un certo senso, lo "costituiscono". Non sono, quindi, né pubbliche né private, se non nel senso in cui è "privato" il funzionamento del mio apparato digerente, o il metabolismo delle cellule che formano i tessuti del mio corpo, o il DNA contenuto nel loro nucleo<sup>[3]</sup>.

## 6. Verso una semantica naturalizzata

L'evoluzione della semantica filosofica in rapporto con le scienze cognitive è in parte analoga, oltre che per molti aspetti inestricabilmente connessa, ad alcuni sviluppi recenti dell'epistemologia e della filosofia delle scienze quali vengono delineati ad esempio in (Kitcher 1992). Nella prospettiva della svolta linguistica in filosofia, il pensiero di Frege è stato interpretato come una destituzione dell'epistemologia dal ruolo di filosofia prima che essa aveva rivestito nella filosofia moderna a partire da Cartesio (Dummett 1973). Con la svolta di Frege tale ruolo fondazionale ed egemone è stato assunto dalla filosofia del linguaggio. Nella filosofia di ispirazione fregeana l'epistemologia (e con essa la teoria della conoscenza scientifica) è stata subordinata alla teoria filosofica del significato, e vincolata ai presupposti che sono alla base della tradizione analitica quale è stata delineata da Dummett (tra cui, appunto, l'antipsicologismo). In anni recenti tuttavia buona parte della ricerca epistemologica si è emancipata dalla tutela analitica (almeno, nella sua forma "ortodossa") e si è sviluppata in accordo a varie linee di evoluzione, il cui fattore comune può essere individuato in un ritorno al naturalismo epistemologico (dove il termine fa chiaro riferimento alla *epistemologia naturalizzata* quineana - Quine 1969). Kitcher pone in luce due elementi come caratteristici di questo ritorno all'epistemologia naturalistica. Il primo di essi consiste nel rifiuto dell'antipsicologismo in epistemologia e in filosofia della scienza. Il secondo consiste nell'abbandono della concezione della ricerca filosofica come impresa effettuabile a priori.

Entrambi questi fattori sono presenti anche nell'evoluzione delle teorie del significato sotto l'influsso delle scienze cognitive. Del superamento dell'antipsicologismo in semantica ho già detto nelle pagine precedenti. Il rifiuto della concezione della filosofia

come pratica a priori è in certo qual modo un corollario del rifiuto dell'antipsicologismo, se si fa riferimento alla psicologia in quanto scienza empirica. Kitcher individua due diversi gradi di coinvolgimento psicologico di una teoria filosofica. Il primo, più debole, si limita al riferimento a concetti di tipo mentale nell'elaborazione della teoria. Il grado più forte di coinvolgimento psicologico si ha invece qualora vengano impiegati specifici risultati della psicologia in quanto scienza empirica. E' chiaro che in questo caso che la teoria filosofica viene a poggiare su conoscenze ottenute a posteriori. Ed è chiaro altresì che gli influssi delle scienze cognitive sulla filosofia del linguaggio vanno, almeno in parte, in questa seconda direzione. Si noti, di passaggio, che i vari approcci alle scienze cognitive comportano, in un certo senso, gradi diversi di coinvolgimento naturalistico. La scienza cognitiva classica, con la sua enfasi sugli aspetti linguistico-simbolici della mente e con il suo disinteresse per le basi biologiche del pensiero, è stata accusata talvolta di assumere un atteggiamento razionalistico e poco orientato empiricamente, e di dare un'immagine della mente eccessivamente disincarnata e separata dal mondo naturale (si vedano ancora Edelman 1992 e Parisi 1992). Tuttavia, già nel richiamo ai risultati delle scienze cognitive classiche, è evidente l'allontanamento dell'ipotesi che una teoria semantica possa nascere da un'analisi puramente a priori dei significati linguistici. D'altra parte, si è visto come il problema di costruire modelli dei meccanismi referenziali del linguaggio sembri chiamare in gioco proprio gli aspetti più "biologici" della percezione e del comportamento.

In ogni caso, qualora si adotti una prospettiva di tipo cognitivo nello studio filosofico del significato, viene a cadere la possibilità che tale studio possa essere condotto a priori. Cade quindi, di conseguenza, il ruolo di filosofia prima per la filosofia del linguaggio, quale è stato rivendicato, ad esempio, oltre che da Dummett, da Tugendath (1976). E cade la possibilità che la teoria del significato possa svolgere un ruolo fondazionale nei confronti di tutta la filosofia e del resto del sapere. In base a tutto quanto si è detto, questo ovviamente non implica che il ruolo di filosofia prima venga assunto da qualche forma di teoria epistemologica o di psicologia filosofica. Né, tantomeno, un ritorno all'evidenza introspettiva quale forma di evidenza indubitabile. Il richiamo al mentale avviene sempre attraverso il richiamo a una teoria della mente, ad un sapere quindi intriso di assunti teorici, e sempre, in linea di principio, rivedibile. Ne consegue dunque che, in questa prospettiva, non può esistere una filosofia prima, né un criterio di demarcazione netto tra filosofia e scienze empiriche, e deve essere abbandonata ogni pretesa fondazionale della filosofia nei confronti delle scienze.

\* Questo lavoro è stato parzialmente finanziato dal Contratto quinquennale C.N.R. - I.I.A.S.S.

#### Note

1 Non è un caso che il principio stesso di composizionalità del significato (che, come ho detto, costituisce uno dei tratti caratteristici delle rappresentazioni mentali nella tradizione cognitiva classica) sia stato individuato proprio da Frege come una delle caratteristiche centrali dei linguaggi sia naturali che artificiali, e sia stato considerato, nella tradizione fregeana, un aspetto di cui ogni teoria del significato è tenuta a render conto.

2 Si noti che Thomason (1979) non si propone di superare l'antipsicologismo in semantica, bensì di sostituire gli usuali argomenti fregeani con nuovi argomenti contro lo psicologismo. Tali argomenti si fondano sulla difficoltà di individuare formalizzazioni coerenti della conoscenza che siano finitamente assiomaticizzabili ma al tempo stesso abbastanza ricche dal punto di vista espressivo da consentire la rappresentazione dei significati linguistici. Infatti, vi sono risultati in base ai quali ogni assiomatizzazione finita al primo ordine basata sulla logica classica, e abbastanza espressiva da formalizzare il significato di una lingua naturale, risulta inconsistente. Tuttavia questo, più che costituire un vero e proprio argomento contro la possibilità del progetto di una semantica cognitiva, sembra piuttosto costituire un problema che la ricerca nel settore deve superare (si veda ad esempio Frixione 1994b).

3 Anche secondo Cussins (1987), la visione criticata da Frege, in base alla quale la comprensione di un'espressione linguistica è questione di indovinare gli stati introspettivi privati di chi la ha proferita, è solo una fra le forme possibili di psicologismo, legata ad una concezione della psicologia come introspezione che non è certamente

quella più attuale. Fra le possibili alternative, Cussins cita la concezione (attribuita a Block 1981b) secondo cui la psicologia è lo studio metodologicamente solipsistico dell'elaborazione interna dell'informazione, o la concezione di Evans (1982), che mette l'enfasi sulle capacità individuali di discriminazione e di riconoscimento.

### Bibliografia

- M.J. B uerle e M.J. Cresswell (1988). Propositional attitudes. In D. Gabbay e F. Guentner (a cura di), *Handbook of Philosophical Logic*, vol. IV, Reidel, Dordrecht e Boston.
- N. Block (1981a). *Imagery*. MIT Press, Cambridge (Ma).
- N. Block (1981b). Psychologism and behaviorism. *The Philosophical Review*, 90(1), pp. 5-43.
- R. Brachman, H. Levesque (a cura di) (1985). *Readings in Knowledge Representation*. Morgan Kaufman, Los Altos (CA).
- G. Brewka (1991). *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge University Press, Cambridge, UK.
- R.A. Brooks (1991). Intelligence without representation. *Artificial Intelligence*, 4, 1991.
- N. Chomsky (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass. Trad. it. in N. Chomsky, *La grammatica generativa trasformazionale. Saggi linguistici*, II, Boringhieri, Torino, 1970.
- Computational Intelligence* (1993). Numero speciale su *The Imagery Debate Revisited*. 9(4), novembre 1993.
- A. Cussins (1987). Varieties of psychologism. *Synthese*, 70, pp. 123-154.
- M. Dummett (1973). *Frege. Philosophy of Language*. Duckworth, Londra. Trad. it. parziale: *Filosofia del linguaggio. Saggio su Frege*, Marietti, Genova.
- M. Dummett (1975). Can analytical philosophy be systematic, and ought it to be? In M. Dummett, *Truth and Other Enigmas*, 1978, Duckworth, London. Trad. it. in M. Dummett, *Verit  ed altri enigmi*, Il Saggiatore, Milano.
- M. Dummett (1988). *Urspr nge der analytischen Philosophie*. Suhrkamp Verlag, Frankfurt. Trad. it.: *Alle origini della filosofia analitica*, Il Mulino, Bologna, 1990.
- U. Eco, M. Santambrogio, P. Violi (a cura di) (1986). *Meanings and Mental Representations*. Versus, 44/45, maggio-dicembre 1986.
- G. Edelman (1992). *Bright Air, Brilliant Fire. On the Matter of Mind*. Basic Books, New York. Trad. it. *Sulla materia della mente*, Adelphi, Milano, 1993.
- G. Evans (1982). *The Varieties of Reference*. Clarendon Press, Oxford.
- J. Fodor (1975). *The Language of Thought*. The Harvester Press.
- J. Fodor (1978). Propositional attitudes. *The Monist*, 61, 501-523.
- J. Fodor, Z. Pylyshyn (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition*, 28, 3-71.
- G. Frege (1884). *Die Grundlagen der Arithmetik*. K sbner, Breslavia. Trad. it. in G. Frege, *Logica e aritmetica*, a cura di C. Mangione, Boringhieri, Torino, 1977.
- G. Frege (1892).  ber Sinn und Bedeutung. *Zeitschrift f r Philosophie und philosophische Kritik*, 100, 25-50. Trad. it. in A. Bonomi (a cura di), *La struttura logica del linguaggio*, Milano, Bompiani, 1973.
- G. Frege (1918). Der Gedanke. Eine logische Untersuchung. *Beitr ge zur Philosophie des deutschen Idealismus*, 1, 58-77. Trad. it. in G. Frege, *Ricerche logiche*, a cura di M. Di Francesco, Guerini, Milano.
- M. Frixione (1994a). *Logica, significato e intelligenza artificiale*. Franco Angeli, Milano.
- M. Frixione (1994b). Intelligenza artificiale e semantica: il problema degli atteggiamenti proposizionali. Apparir  sugli *Atti del convegno AICA*, Palermo, settembre 1994.
- M. Frixione, S. Gaglio, G. Spinelli (1989). Symbols and subsymbols for representing knowledge: A catalogue raisonn . *Atti del 11th International Joint Conference on Artificial Intelligence*, Detroit, USA.
- M. Frixione, G. Spinelli (1992). Connectionism and functionalism: The importance of being a subsymbolist. *Journal of Experimental and Theoretical Artificial Intelligence*, 1(1992), 3-17.
- D.E. Goldberg (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading (Mass.).
- S. Harnad (1990). The symbol grounding problem. *Physica D*, 42(1-3).
- P. Kitcher (1992). The naturalists return. *The Philosophical Review*, 101(1), 53-114.
- D. Marconi (1989). Rappresentare il significato lessicale. In R. Viale (a cura di), *Mente umana e mente artificiale*, Feltrinelli, Milano.
- D. Marconi (1992). Semantica cognitiva. In Santambrogio (1992).
- M. Minsky (1975). A framework for representing knowledge. In P. Winston (a cura di), *The Psychology of Computer Vision*, McGraw Hill, New York. Una versione parziale   compresa in Brachman e Levesque (1985).
- D. Parisi (1992). Uscire dal linguaggio. *Il Mulino*, 61(344), 994-1000.
- S. Peters, E. Saarinen (a cura di) (1982). *Processes, Beliefs and Questions*. Reidel, Dordrecht e Boston.
- W. Quine (1969). Epistemology naturalized. In *Ontological Relativity and Other Essays*, Columbia University Press, New York.

- D.E. Rumelhart, J.L. McClelland e il gruppo di ricerca PDP (1986). *Parallel Distributed Processes. Exploration in the Microstructure of Cognition*. 2 vol. MIT Press, Cambridge, Mass. Trad. it.: *La microstruttura dei processi cognitivi. Elaborazione distribuita in parallelo*, Il Mulino, Bologna, 1991.
- M. Santambrogio (a cura di) (1992). *Introduzione alla filosofia analitica del linguaggio*, Laterza, Bari.
- R. Serra e G. Zanarini (1990). *Complex Systems and Cognitive Processes*. Springer, Berlino.
- Sistemi intelligenti* (1992). Numero speciale sulla vita artificiale. 4(2), agosto 1992.
- J.R. Searle (1983). *Intentionality: an Essay in the Philosophy of Mind*. Cambridge University Press, Cambridge, UK. Trad. it.: *Dell'intenzionalità*, Bompiani, Milano, 1985.
- B. Smith (1982). Prologue to 'Reflection and Semantics in a Procedural Language'. In Brachman e Levesque (1985), 31-40.
- P. Smolensky (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11. Trad. it.: *Il connessionismo tra simboli e neuroni*, Marietti, Genova, 1992.
- V. Somenzi, R. Cordeschi (a cura di) (1986). *La filosofia degli automi*. Boringhieri, Torino.
- J. Sowa (a cura di) (1991). *Principles of Semantic Networks*. Morgan Kaufmann, San Mateo, CA.
- R.C. Stalnaker (1984). *Inquiry*. MIT Press, Boston, MA.
- R. Thomason (1979). Some limitations to the psychological orientation in semantic theory. Manoscritto inedito.
- T. Toffoli e N. Margolis (1987). *Cellular Automata Machines*. MIT Press, Cambridge (Mass.).
- E. Tugendath (1976). *Vorlesungen zur Einföhrung in die sprachanalytische Philosophie*. Suhrkamp, Francoforte. Trad. it. parziale: *Introduzione alla filosofia analitica*, Marietti, Genova.
- A. Turing (1950). Computing machinery and intelligence. *Mind*, 59, 433-460. Trad. it. in V. Somenzi, R. Cordeschi (a cura di) (1986).