

# WAT (Words As social Tools): una prospettiva socio-corporea sulla cognizione umana

Anna M. Borghi\*, Felice Cimatti°

\*Università di Bologna, Istituto di Scienze e Tecnologie della Cognizione, CNR, Roma

°Università della Calabria, Dipartimento di Studi Umanistici<sup>1</sup>

Ringraziamo preliminarmente XXX per averci chiamati in causa e per aver analizzato nel dettaglio la nostra proposta teorica. Nel commento che segue prenderemo in esame i principali punti critici individuati dall'autore, argomentando in difesa della nostra proposta.

## 1. Il problema della selezione fra le esperienze pertinenti

**Domanda:** Il primo punto critico che l'autore solleva riguarda il problema della selezione: «*quali tra le esperienze che co-occorrono all'uso di una determinata parola astratta sono quelle che realmente concorrono a costituirne il contenuto semantico?*». L'autore fa l'esempio della parola astratta "giustizia", che può essere usata in una grande varietà di contesti diversi: chi apprende ad usare questa parola come fa a selezionare, fra tutti le esperienze potenzialmente selezionabili, quelle associate a questa parola, e solo a questa? Ci chiede pertanto se ipotizziamo l'esistenza di un criterio di selezione olistico - per cui in qualche modo tutte le esperienze entrerebbero a far parte della rappresentazione collegata alla parola – oppure se riteniamo che esista un nucleo 'essenziale' di esperienze pertinenti. Abbiamo diverse risposte per questa domanda.

**Risposta 1. Elementi essenziali?** Non vediamo la ragione per cui questo problema si ponga con i concetti astratti e non per quelli concreti; la presenza di un oggetto a cui si riferisce il concetto/parola non cambia la natura del problema di quale sia il significato di una espressione linguistica. Il problema delle caratteristiche definitorie o "essenziali" dei concetti è da tempo dibattuto nell'ambito degli studi sulla categorizzazione. In generale riteniamo che non esistano e che comunque non servano caratteristiche essenziali per definire un concetto, come già da tempo evidenziato da Wittgenstein attraverso la nozione di "somiglianze di famiglia", a sua volta ripresa nell'ambito della classica teoria dei prototipi che criticava la teoria tradizionale dei concetti. Si consideri che non siamo necessariamente a favore di un modello per prototipi, basato sull'astrazione, rispetto ad un modello per esemplari, secondo cui tutti gli esemplari che incontriamo di una categoria possono essere memorizzati e riattivati, dato che entrambi sono modelli potenzialmente compatibili con una visione "embodied" – non lo è, però, a nostro avviso, una teoria essenzialista e innatista dei concetti (Barsalou et al., 2010; Borghi, 2002).

**Risposta 2. Apprendimento statistico.** Secondo l'autore «andrebbe chiarito il criterio con cui sono selezionate tali esperienze [quelle pertinenti] e, soprattutto, il processo che, a partire dall'osservazione di come una parola è usata, giunge a costruire la rappresentazione mentale (il concetto) corrispondente». Gli studi recenti sull'apprendimento statistico del linguaggio (Saffran, 2003; Smith e Yu, 2008; Smith et al., 2014) rilevano che la frequenza e la co-occorrenza di parole in situazioni diverse aiuta ad individuarne il referente, influenzando così potentemente l'acquisizione del linguaggio. Una osservazione come quella presentata dall'autore sottostima

---

<sup>1</sup> Il commento è stato scritto dai due autori a quattro mani. L'ordine dei nomi è alfabetico.

ampiamente l'informazione ecologicamente presente nella situazione di apprendimento sia dal punto di vista percettivo-motorio che sociale.

**Risposta 3. Non solo statistica: la guida degli altri.** Riteniamo che alcuni aspetti della situazione ecologica di apprendimento presentino – durante l'ontogenesi linguistica - una forte salienza percettiva, al di là della frequenza con cui si presentano. Ad esempio, la nostra proposta prevede che per l'acquisizione delle parole astratte abbia un ruolo fondamentale la presenza di altri – tipicamente adulti - linguisticamente competenti, che 'filtrano' l'informazione ambientale per presentarla al bambino/a, aiutandolo a focalizzare l'attenzione su quello che effettivamente conta per apprendere quella parola. Ovviamente riteniamo continuo anche altre fonti autorevoli, come illustrazioni, testi scritti etc., anche se in misura minore rispetto all'interazione sociale diretta. Una proposta come la nostra prevede ovviamente che il bambino/a attivamente ricerchi i significati e si aspetti la collaborazione altrui, non subendo passivamente gli input che gli arrivano dagli altri. Esistono quindi due meccanismi fondamentali, uno bottom-up e uno top-down, che nella nostra ipotesi non sono in alcun modo in contrasto l'uno con l'altro. Si veda a questo proposito quanto riportato nel libro di Borghi e Binkofski (2014: 90), in cui si propone che i due meccanismi operino sinergicamente:

In fact, we argue that a similar associative mechanism might work for acquisition of both concrete and abstract words. Perceptive salience, embodied attention and bodily actions would contribute to the success of learning concrete words through associations, eventually aided by parents: children move in their environment, and through selective attention they get closer and focus on interesting objects, encoding also the similarities and the relationship between them. However, for acquiring abstract word an associative mechanism based on words to referents mapping might be more time consuming and difficult to apply: given the sparse variety of referents abstract words have, children might need a lot of evidence to form a category as, for example, "freedom". This does not mean that an embodied mechanism of referent searching and of selective attention is not present. However, due to the higher difficulty in selecting the perceptual referent, both the social and the linguistic input might be particularly relevant for abstract word learning. As shown in the acquisition studies, the linguistic input becomes particularly precious since the referents of abstract words are more sparse than those of concrete ones. In addition, words might be associated not only to perceptually salient referents but also to other words, to facilitate learning. This is possible also because, as shown by literature on age of acquisition, abstract words are acquired rather late compared to concrete ones, when the vocabulary burst has already occurred. Thus, an associative learning mechanism is present both in associating words and referents and words to other words. But this is not the whole story: in keeping with an embodied perspective, learning of these associations has an impact not only on representation but on the body as well, since different effectors, the mouth and the hand, are activated.

## 2. Rappresentazioni mentali

**Domanda.** L'autore – che sembra seguire l'impostazione delle scienze cognitive classiche, incentrata sulla nozione di "rappresentazione mentale" - si chiede se la WAT, con la sua insistenza sulla esperienza corporea, abbia qualcosa da dire: *«sull'esistenza o meno di rappresentazioni mentali costitutive del contenuto semantico delle parole astratte: gli autori ritengono che tali rappresentazioni esistano oppure intendono propendere per una teoria eliminativista, difendendo una posizione wittgensteiniana dell'uso»?*

**Risposta:** La nostra proposta non analizza a fondo il problema delle rappresentazioni e della loro natura (simboliche o non simboliche, ad esempio). Quello che ipotizziamo, alla luce delle più recenti prove sperimentali, è che i processi di acquisizione delle parole influenzino profondamente e costituiscano la rappresentazione neurale dei concetti. Abbiamo infatti trovato tramite lavori

condotti con risonanza magnetica funzionale (es. Sakreida et al., 2013) che, mentre i concetti concreti sono rappresentati primariamente in aree sensorimotorie, quelli astratti attivano sia le aree sensorimotorie che quelle linguistiche, in linea con la nostra proposta socio-corporea della cognizione. Ulteriori evidenze mostrano anche che il tipo di rappresentazione neurale ha un'influenza sul comportamento: i concetti concreti attivano primariamente la mano, quelli astratti primariamente la bocca, cioè l'insieme dell'apparato fonatorio (es. Borghi et al., 2011; Scorolli et al., 2011, 2012), mentre lo sguardo, legato alla focalizzazione dell'attenzione, è rilevante per entrambi.

Riassumendo in modo brutale, ma non lontano dal succo della questione: le parole concrete la capiamo con la mano, quelle astratte con la lingua e le labbra. In entrambi i casi abbiamo a che fare con esperienze corporee, le prime più legate alle azioni dirette del corpo sul mondo, le seconde ai "gesti vocali", ossia alle azioni fatte di parole. A questo proposito la nostra proposta non è necessariamente eliminativista; non sostiene cioè che non esistano le rappresentazioni, sostiene però che sono sempre incarnate. Quanto alla nozione di "uso" di Wittgenstein, in realtà non è chiaro in che senso possa essere considerato un eliminativista *ante litteram*. Wittgenstein introduce la nozione di *uso* proprio per non rimanere ingabbiato né nel mentalismo né nel comportamentismo (cfr. Cimatti, 2004; 2008). Questa nozione serve a Wittgenstein per spostare il punto di vista sul linguaggio, che così diventa un modo di agire nel mondo.

### 3. Regola e normatività

**Domanda.** Un altro aspetto della WAT su cui l'autore si concentra è quello della normatività, su cui finora il nostro lavoro si è effettivamente meno concentrato. L'autore pone questa domanda: «che cosa intendono gli autori quando parlano di corretto uso delle parole e di "seguire una regola sociale"? Le parole sono, infatti, da loro descritte come strumenti sociali, il cui uso può essere corretto o scorretto, il che dipende dal fatto che l'individuo nell'utilizzarle segua o meno determinate norme sociali. Ora, in che modo tali norme sono apprese e interiorizzate dal parlante? E come influisce ciò sul processo di formazione delle rappresentazioni mentali?».

**Risposta.** Ancora una volta la nostra risposta si basa sull'ispirazione ecologica alla base di ogni teoria "embodied". L'acquisizione di norme avviene nel contesto sociale in cui i bambini si muovono e vivono. A nostro avviso non esiste qualcosa come l'"essenza" della norma: la norma è un comportamento che viene interiorizzato a partire da relazioni sociali con altri autorevoli, ad esempio gli adulti, in particolare i genitori. Un esempio è quello in cui i genitori correggono il parlato e/o gli ipercorrettismi dei bambini, come "puliscio", "bevere", oppure l'uso dell'articolo "la" per riferirsi al panda la prima volta che se ne sente il nome abbinato all'animale. La "norma" è il controllo sociale esercitato sul bambino/a, che viene più o meno esplicitamente "guidato" verso la forma linguistica "corretta" nel suo ambiente di sviluppo.

Come mostrato nella letteratura sull'apprendimento socio-pragmatico delle parole (v. la rassegna in Borghi e Binkofski, 2014), esistono vincoli che portano i bambini a tener conto del punto di vista degli adulti; la presenza dell'altro (che per noi è un altro corpo) costituisce dunque il principale criterio di rilevanza. Nel corso dello sviluppo, un altro fattore importante è l'acquisizione della scrittura che esercita un potente ruolo normativo e normalizzante. Nell'insieme questi meccanismi sociali esercitano un ruolo "normativo" sul comportamento linguistico in sviluppo del bambino. In questo modo la norma è letteralmente incarnata nel corpo degli altri umani con cui il bambino/a ha a che fare durante l'ontogenesi linguistica.

### 4. Sintassi e teorie incarnate

**Domanda.** Secondo l'autore «la sintassi ha elementi non riducibili al sistema sensomotorio». Il livello sintattico, cioè, sarebbe indipendente da quello sensomotorio, ciò che – secondo l'autore – dovrebbe rendere necessario ammettere la presenza, nell'apparato cognitivo umano, di un livello di rappresentazionale amodale e in qualche modo disincarnato.

**Risposta.** In realtà non abbiamo mai sostenuto che parlare consista soltanto nell'emissione di suoni che 'stanno' per esperienze corporee; la nostra proposta non riduce assolutamente l'esperienza linguistica e lessicale al solo sistema fonatorio, acustico e motorio, al limite individua le possibili relazioni tra questi aspetti e il significato delle parole, che si fonda sul sistema sensorimotorio. Il punto generale dell'Autore però è che «l'esperienza linguistica lessicale non può essere ridotta al solo *sistema fonatorio, acustico e motorio*, ma coinvolge anche altri sistemi come quello *sintattico, morfologico e fonologico*». Preliminarmente facciamo notare che le categorie sintattiche e morfologiche non nascono già come sintattiche e morfologiche, bensì invece come lessicali, e quindi pienamente incarnate (cfr. Heine e Kuteva 2002). All'inizio, anche della amodale sintassi, c'è sempre la affatto *modale* esperienza corporea; la sintassi è una specializzazione e formalizzazione dei gesti corporei.

Più in particolare, rispetto agli aspetti sintattici e morfologici (a parte il fatto che non capiamo il perché della collocazione di quelli fonologici insieme a questi ultimi), è vero che almeno inizialmente la letteratura sull'*embodiment* si è concentrata sulla comprensione del linguaggio. Tuttavia, esistono numerosi studi recenti che mostrano che comprensione e produzione sono due facce della stessa medaglia. Quanto alla sintassi concordiamo con chi afferma che le evidenze non sono ancora conclusive (es. Tettamanti e Moro, 2012 su sintassi e sistema dei neuroni specchio). Tuttavia, esistono prove ormai classiche che mostrano che le *affordance* percettive vincolano l'organizzazione sintattica delle frasi e ne influenzano la comprensione (es. Glenberg e Robertson, 2000). Più recentemente, numerosi lavori hanno mostrato che le reti neurali relative alla grammatica si sono sviluppate in aree deputate ai circuiti relativi ad azioni coordinate e complesse (es. Pulvermüller e Fadiga, 2010), evidenziando così il legame tra aree motorie e aree linguistiche connesse alla grammatica. Per quanto riguarda la fonologia, esiste una letteratura consolidata e ripresa di recente sulle relazioni tra aspetti fonologici e di significato (es. sinestesia, iconismo fonologico)(es. Flumini et al. 2014), oltre che riflessioni a proposito del probabile ruolo evolucionistico del gesto vocale rispetto al linguaggio (es. Gentilucci e Corballis, 2006).

## 5. Il problema delle rappresentazioni amodali

**Domanda.** La critica centrale dell'articolo dell'autore è che la proposta WAT non riesce a dare conto dell'insieme del funzionamento cognitivo umano: «anche se tali esperimenti mostrano che vi è coinvolgimento di rappresentazioni modali associate ai concetti astratti, tuttavia, ciò non dimostra che il contenuto di questi sia esaurito dalle suddette rappresentazioni. Nulla vieta, infatti, che oltre a rappresentazioni sensomotorie essi possano contenere *anche* simboli/rappresentazioni amodali».

**Risposta.** Crediamo che in primo luogo andrebbe chiarito che cosa si intende per modale ed amodale. Se con rappresentazioni amodali si intende il fatto che l'informazione sensorimotoria, modale, confluisce e viene tradotta in simboli astratti, amodali e preposizionali, non riteniamo che al momento esistano evidenze che ne dimostrino l'esistenza. Al punto in cui è arrivata la ricerca empirica, riteniamo che la prospettiva "embodied" che non prevede l'esistenza di simboli amodali sia quella più 'economica' da un punto di vista teorico, e più plausibile biologicamente; al contrario crediamo che la tesi più difficile da giustificare – evolutivamente e teoricamente – sia proprio quella delle scienze cognitive classiche, che presuppone l'esistenza di simboli amodali. Sta a chi sostiene che accanto ai simboli modali esistono anche simboli amodali dimostrarne l'esistenza. Se ci fossero chiare prove dell'esistenza di simboli amodali, siamo pronti a rivedere la nostra posizione.

Se con rappresentazioni amodali intendiamo il confluire di informazione di origine sensorimotoria relativa a diverse modalità in aree neuronali meno direttamente legate a specifiche esperienze, come le zone di convergenza alla Damasio (1989), non vediamo incompatibilità con la nostra proposta, purché sia mantenuto il nesso indistricabile tra informazione sensorimotoria e conoscenza. C'è invece incompatibilità tra la nostra prospettiva e teorie amodali come quella del “semantic hub”, proposte inizialmente per spiegare la demenza semantica, secondo cui esiste una zona singola, indipendente dal dominio, in cui tutte le diverse modalità sono integrate (Patterson et al., 2007). Questa teoria, che si basa su evidenze messe di recente in discussione da alcuni autori (si vedano ad esempio Gainotti, 2012; Simmons e Martin, 2009), è chiaramente in conflitto con una visione embodied e grounded dei concetti in quanto considera la percezione soltanto come un canale d'accesso, di rilevanza marginale, alla conoscenza astratta.

## 6. Il caso della negazione

**Domanda.** Secondo l'autore una sfida particolarmente difficile per la WAT, e più in generale per ogni teoria “embodied”, è quella posta dalla negazione: «sembra, infatti, pressoché impossibile descrivere in termini di rappresentazioni modali – quindi di simulazioni mentali – la comprensione e la rielaborazione cognitiva di enunciati come: “Non è vero che Mario non dice che Chiara non è andata a casa senza la bicicletta”».

**Risposta.** In primo luogo non ci sembra che questa frase costituisca un buon esempio di negazione, che rifletta il modo in cui usiamo quotidianamente la lingua. In generale, pensiamo sia imprescindibile studiare le categorie naturali e il modo in cui usiamo realmente la lingua, e meno cruciale soffermarci su esempi che potrebbero essere qualificati come “outlier”. In effetti un enunciato del genere è di difficile comprensione a meno che non lo si possa leggere, appoggiandosi quindi ad un supporto *esterno* come la scrittura. L'esempio scelto dall'autore, in realtà, mostra come le forme linguistiche più complesse dipendono dalla scrittura, e quindi dalla vista, cioè dagli *occhi*, e che probabilmente sarebbero impossibili se il nostro pensiero non fosse stato potentemente modificato dalla scrittura. Ciò costituisce una ulteriore conferma del carattere sempre incarnato della comprensione linguistica (Goody, 1980).

Al di là di questo, esistono numerosi studi sia comportamentali che neurali che mostrano come le teorie embodied siano in grado di spiegare la negazione, simulando le situazioni che vengono negate (es. Tettamanti et al., 2008; Kaup et al., 2013).

**Domanda.** A proposito dei ruoli tematici, l'autore osserva: «ritengo sia molto difficile rendere conto in termini simulativisti di determinate informazioni semantiche, quali i ruoli tematici».

**Risposta.** In realtà proprio nelle prime proposte di teorie “embodied” (si veda ad esempio Barsalou, 1999) si faceva riferimento a modelli basati su *frame*, in cui i diversi ruoli tematici sono riempiti di contenuti modali. I simboli percettivi, secondo Barsalou (1999), erano appunto integrati in frame. In questo senso ci sembra che i ruoli tematici che possano essere integrati in frame non costituiscano affatto un problema per la WAT.

## 7. Conclusioni

Ringraziamo l'autore dell'articolo per averci consentito di entrare nel merito della proposta WAT. Crediamo di avere risposto a tutti gli interrogativi posti, anche a quelli non esplicitamente affrontati o approfonditi nella nostra teoria (ad esempio il problema della rappresentazione, quello della sintassi etc.). Nei casi di temi non specificamente affrontati nella nostra teoria, abbiamo fornito risposte preliminari; si tratta spesso di temi che andrebbero approfonditi anche attraverso un'attenta

analisi delle evidenze empiriche, che in questa sede non abbiamo svolto. Ci permettiamo un'ultima osservazione: la teoria WAT, in realtà, sostiene una posizione molto semplice e, crediamo, del tutto plausibile; per studiare la conoscenza umana si deve partire dal corpo umano e dal suo ambiente, che è fatto di oggetti, animati e inanimati, e altri corpi umani. All'inizio c'erano i corpi, e non erano soli.

### **Riferimenti bibliografici**

- Barsalou L.W. (1999), Perceptions of perceptual symbols. *Behavioral and brain sciences*, 22, pp. 637-660.
- Barsalou, L.W., Wilson, C.D., Hasenkamp, W., Mesquita, B., Feldman Barrett, L., Smith, E. (2010). On the vices of nominalization and the virtues of contextualizing. In *The mind in context*, pp. 334-360.
- Borghgi, A.M. (2002). Concetti e azione. In Borghgi, A.M. e Iachini, T. (a cura di). *Scienze della mente*. Bologna: Il Mulino, pp. 203-222.
- Borghgi, A.M., Flumini, A., Cimatti, F., Marocco, D. Scorolli, C. (2011). Manipulating objects and telling words: A study on concrete and abstract words acquisition. *Frontiers in Psychology*, 2:15. doi: 10.3389/fpsyg.2011.00015.
- Borghgi, A.M., Binkofski, F. (2014). *Words as Social Tools: An Embodied View on Abstract Concepts*. New York: Springer.
- Cimatti, F. (2004). *Il senso della mente. Per una critica del cognitivismo*. Torino: Bollati Boringhieri.
- Cimatti, F. (2008). Quanto è cognitiva la scienza cognitiva? Una nota sul rapporto fra psicologia e scienze della mente. *Giornale italiano di psicologia*, xxxv, 3, pp. 715-723.
- Damasio, A.R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation*, 1(1), pp. 123-132.
- Flumini, A., Ranzini, A., Borghgi, A.M. (2014). Nomina sunt consequentia rerum—Sound—shape correspondences with every-day objects figures. *Journal of memory and language*, 76, pp.47-60.
- Gainotti, G. (2012). The format of conceptual representations disrupted in semantic dementia: a position paper. *Cortex*, 48(5), pp. 521-529.
- Gentilucci, M., Corballis, M.C. (2006). From manual gesture to speech: a gradual transition. *Neuroscience & Biobehavioral Reviews*, 30, pp. 949-960.
- Glenberg, A. M., Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of memory and language*, 43, pp. 379-401.
- Goody, J. (1980). Thought and writing. In E. Gellner (ed.) *Soviet and Western anthropology*. Londra: Duckworth, pp. 119-133.

- Heine, B., Kuteva, T. (2002). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Kaup, B., Zwaan, R. A., Lüdtke, J. (2013). The experiential view of language comprehension: How is negation represented?.
- Patterson, K., Nestor, P. J., Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), pp. 976-987.
- Pulvermüller, F., Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11, pp. 351-360.
- Saffran, J.R. (2003). Statistical language learning mechanisms and constraints. *Current directions in psychological science*, 12(4), pp. 110-114.
- Scorolli, C., Binkofski, F., Buccino, G., Nicoletti, R., Riggio, L., Borghi, A.M. (2011). Abstract and concrete sentences, embodiment, and languages. *Frontiers in Psychology*, 2, 227. doi: 10.3389/fpsyg.2011.00227.
- Scorolli, C., Jacquet, P., Binkofski, F., Nicoletti, R., Tessari, A., Borghi, A.M. (2012). Abstract and concrete phrases processing differently modulates cortico-spinal excitability. *Brain Research*, 1488, 60-71. doi: 10.1016/j.brainres.2012.10.004.
- Simmons, W., Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, 15(05), pp. 645-649.
- Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), pp.1558-1568.
- Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. *Trends in cognitive sciences*, 18(5), pp. 251-258.
- Tettamanti, M., Manenti, R., Della Rosa, P. A., Falini, A., Perani, D., Cappa, S. F., & Moro, A. (2008). Negation in the brain: modulating action representations. *Neuroimage*, 43(2), pp. 358-367.
- Tettamanti, M., Moro, A. (2012). Can syntax appear in a mirror (system)?. *Cortex*, 48(7), 923-935.