

# SVILUPPO COGNITIVO E NAMING EXPLOSION, IL CONTRIBUTO DEI MODELLI COMPUTAZIONALI NELLO STUDIO DELLE TAPPE FONDAMENTALI DELL'ACQUISIZIONE DEL LINGUAGGIO

Giuseppe Città - langravio@gmail.com

Dipartimento di Scienze Cognitive, della Formazione e degli Studi Culturali, Università di Messina

## Abstract

The aim of this article is to focus the use of computational models in the study of language acquisition. Specifically, we will refer to a kind of model – neural networks – whose primary focus is to simulate the computations that occur in the brain. These models, moreover, assume the biological plausibility and the psychological plausibility as their hallmarks. After examining some central issues that affect the process of language learning in the specific phase of the vocabulary spurt (the multimodality of human cognition and the trajectory of cognitive and linguistic development), we will show, through the analysis of two computational approaches, how these issues are computationally addressed.

## Keywords

Neural networks, multimodality of cognition, naming explosion, shape of change, naming errors.

## 1. Strumenti ed oggetto d'analisi: modelli computazionali e acquisizione del lessico

L'utilizzo dei modelli computazionali per lo studio di un fenomeno complesso quale è l'acquisizione del linguaggio implica numerose riflessioni relative alla validità di tali modelli soprattutto in riferimento alla loro plausibilità biologica e psicologica.

In letteratura si possono distinguere due macro-classi di modelli computazionali. All'interno della prima classe si possono inserire tutti quei modelli basati sull'intelligenza artificiale simbolica dei quali uno scopo primario di realizzazione è lo studio di determinati fenomeni cognitivi prescindendo dai processi cerebrali coinvolti.

Nella seconda classe, invece, si possono inserire tutti quei modelli che assumono tra i loro fondamenti teorici la centralità della simulazione delle computazioni che avvengono nel cervello. Un loro fine realizzativo primario è generare comportamenti simili a quelli osservati nella realtà gettando luce, così, su come il cervello riesce ad implementare i vari compiti cognitivi.

Lo scopo di questo lavoro è problematizzare l'utilizzo delle reti neurali, appartenenti, alla seconda tipologia di modelli sopra menzionata, nei termini di strumenti d'analisi del fenomeno dell'apprendimento del linguaggio, in relazione a delle problematiche specifiche quali la traiettoria dello sviluppo cognitivo-linguistico e la natura multimodale dell'input con cui il bambino viene in contatto in determinate fasi del processo d'ontogenesi.

Come punto d'avvio si utilizzerà la seguente definizione: *un modello è la rappresentazione compatta di un fenomeno che può generare un comportamento paragonabile a quello sotto esame nel sistema reale di riferimento. In più è un mezzo per rendere più evidenti le caratteristiche chiave di un oggetto d'analisi* (Nyamapfene, 2009: 36). Inoltre, un modello, incrementando la chiarezza della teoria che affianca, offre la possibilità di formulare e verificare nuove ipotesi.

Generalmente lo si trova all'interno di una teoria, all'interno cioè di un modo ben preciso di guardare e concepire un determinato fenomeno. Lo si può concepire, come detto, come ciò che articola con maggior precisione di dettaglio una teoria specifica, cosa che una descrizione semplicemente verbale potrebbe non essere in grado di fare. Esso offre l'opportunità di esplorare aspetti di un fenomeno che può non essere facilmente testato nel mondo reale; lo si può pensare come un vero e proprio generatore di ipotesi: suggerisce nuovi modi di comprendere un fenomeno, anche se la validità di un'ipotesi dipenderà, in ultima istanza, dai test empirici effettuati nel mondo reale.

Allo scopo di chiarire ulteriormente il ruolo giocato da tali di-

spositivi il primo passo da compiere è esplicitare i loro oggetti d'analisi: il fenomeno dell'acquisizione del linguaggio e in particolare i processi che interessano l'acquisizione del lessico.

L'approccio che si prediligerà è quello portato avanti da quegli studi che pongono il lessico come elemento cognitivo fondamentale (Elman, 2005). Particolarmente illuminante è una metafora operativa (Elman, 2004) secondo cui è molto più fruttuoso concepire le parole come operatori piuttosto che come operandi. Esse hanno una funzione attiva che si viene a configurare come la base di uno studio del rapporto che intercorre tra la struttura del lessico stesso e i processi di organizzazione della conoscenza. Un tal modo di pensare alle parole di una lingua, richiamando una pregnante similitudine, consente uno studio di esse non tanto come particelle linguistiche che danno vita a impalcature grammaticali, ma quanto come elementi portanti che sono già di per sé ricche entità grammaticali (Elman, 2009: 2).

Che le parole siano operatori suggerisce che esse sono molto più che semplici oggetti; suggerisce un abbandono della concezione di un lessico-lista in cui le parole sono degli elementi su cui operare attraverso un gruppo di regole; suggerisce l'assunzione di una rappresentazione più complessa, articolata e flessibile nella quale l'informazione lessicale sia sintattica che semantica gioca una partita fondamentale nella strutturazione e nella costruzione on-line delle proposizioni (Elman, 1995, Jackendoff, 2007). Le parole, secondo la prospettiva che si sta sostenendo, potrebbero essere di per sé gli elementi fondativi dai quali epifenomenicamente emergerebbe la grammatica (Elman, 2009: 2); elementi che solo un'osservazione che rimanga in superficie restituirebbe come costituenti semplici del linguaggio celando una loro natura predicativa interna molto più complessa (Lenci, 2006).

Attribuire alle parole un tale ruolo, di conseguenza, significa concepirle alla stessa stregua degli stimoli sensoriali, come componenti, cioè, che agiscono direttamente sugli stati mentali (Elman, 2004).

Una prospettiva questa che si fa carico di ricercare nei processi cognitivi che interessano la strutturazione del lessico, delle costanti, delle regolarità tra categorie lessicali, proprietà percettive degli oggetti del mondo e proprietà dei vari elementi che compongono il linguaggio (Colunga & Smith, 2008), ed è proprio questo il terreno entro il quale si registra uno dei contributi maggiormente fruttuosi fornito dai modelli computazionali dell'acquisizione del linguaggio, in particolare da quelli che hanno come fulcro d'analisi il processo di acquisizione: i modelli connessionisti.

## 2. Nuove ipotesi da 'vecchi' problemi

Questo genere di modelli ha contribuito nel corso del tempo a gettar luce su nuclei centrali del processo di acquisizione del linguaggio e, più in generale, dello sviluppo cognitivo di cui l'apprendimento della lingua è tappa fondamentale. Ora, proprio di questi 'vecchi' nuclei teorici sarà opportuno occuparsi per fornire al lettore gli strumenti necessari per una comprensione chiara e lineare delle ipotesi che emergeranno nei paragrafi successivi (Barsalou *et al.*, 2007; Elman, 2005, 2006; Rogers *et al.*, 2004).

Verranno presi in esame:

- la multimodalità della cognizione umana;
- la traiettoria dello sviluppo linguistico-cognitivo, la cosiddetta *shape of change*, e le correlate fasi critiche quali l'improvviso incremento quantitativo delle performance linguistiche e la decadenza delle stesse.

### 2.1 La natura multiforme della cognizione

Un assunto fondamentale da cui è opportuno trarre avvio è la tesi secondo cui la cognizione sia un fenomeno composito emergente da un insieme articolato di processi connessi tra loro e al mondo (Smith, 2005), un fenomeno, cioè, che emerge dalla stretta relazione tra i diversi sistemi cognitivi, quali, ad esempio, la percezione, l'azione, il linguaggio, l'apprendimento, la memoria, attraverso i quali interagiamo con la realtà. Descrivere la cognizione in tal modo significa ancorarla a delle caratteristiche che vanno a costituire ciò che viene riconosciuto come *la magia della cognizione umana*, espressione, questa, che sottolinea profondamente la capacità di costruire rappresentazioni e comportamenti che favoriscono il raggiungimento di obiettivi e la coordinazione sociale riferendosi tanto alla duttilità quanto alla solidità dell'apprendimento nei vari contesti (Barsalou *et al.*, 2007). Si fa riferimento ad alcuni tratti che sarà opportuno scandire più chiaramente marcando il fatto che, quando si parla di cognizione, non ci si riferisce ad un insieme di contenuti mentali immobili legati alla realtà, contenuti che forniscano eventualmente una spiegazione del fenomeno in termini di strutture cognitive rigide e astratte messe alla prova in contesti differenti e in modo reiterato. Per comprenderne in pieno la vera natura, piuttosto, la si deve legare alla comprensione e allo studio dei contributi che hanno origine da quei sistemi che sono generalmente e a torto classificati come sistemi non cognitivi. Ciò comporta che si metta da parte un approccio alla problematica che leghi la cognizione esclusivamente alla sfera del pensiero. Una visione che seziona la vita mentale in gradini separati quali, ad esempio, pensiero, sensibilità e azione, e conferisce alla fase del pensiero il ruolo di componente cognitiva, si concentra, di conseguenza, sulla stabilità della cognizione e tende a darne una spiegazione attraverso la formulazione di contenuti mentali (i concetti, le categorie) organizzati in rappresentazioni simboliche fisse e cristallizzate. Secondo questa prospettiva ognuno di noi avrebbe degli schemi fissi di vario genere (temporali, causali, logici) relativi alla struttura del mondo che però sarebbero nettamente separati da processi in tempo reale quali il percepire, l'agire e il ricordare. Imparare a pensare adeguatamente al fenomeno della cognizione equivale a pensare ogni singolo processo cognitivo in stretta coordinazione con gli altri. È necessario, cioè, non perdere di vista la relazione di co-azione tra percezione, linguaggio, azione, memoria che insieme ad altri elementi fondamentali quali, ad esempio, le emozioni e l'attenzione si configurano come i connotati peculiari di un elemento stratificato e multiforme.

### 2.2 La fisionomia dello sviluppo cognitivo

All'interno di un contesto che fornisce una raffigurazione dinamica del fenomeno, pensare l'evoluzione di un determinato processo cognitivo come un percorso lineare che partendo da un minimo quantitativo approda ad un massimo quantitativo significa abbracciare un'immagine della conoscenza poco fedele al reale.

In realtà si è di fronte ad un fenomeno ben più complesso che si configura come tutt'altro che statico e che è strettamente legato al mondo (Smith, 2005) sia nello sviluppo dei suoi specifici processi che nella sua generale evoluzione. E capire lo sviluppo cognitivo richiede la comprensione di come i processi che legano il nostro sistema cognitivo al mondo fisico si modificano attraverso l'interazione con il mondo stesso. Per dirla in altri termini, i meccanismi di apprendimento che supportano l'evoluzione del sistema cognitivo sono profondamente sensibili, nei vari domini, compreso quello linguistico, alle variazioni e agli schemi di co-variazione presenti nell'ambiente (Rogers *et al.*, 2004). Facendo riferimento agli schemi di sviluppo linguistico-cognitivo, quindi, una traiettoria di sviluppo non-lineare appare molto più adeguata per descrivere i processi relativi all'acquisizione di una lingua, processi che ci si troverebbe in profondo difetto nel leggere come semplici eventi incrementali e progressivi. Tale traiettoria sembra procedere, difatti, alternando lunghi periodi in cui sembra mutare ben poco a brevi fasi esplosive in cui si registra un cambiamento significativo sia in termini di qualità che in termini di quantità. Si tratta di un percorso evolutivo a cui generalmente ci si riferisce con l'espressione 'traiettoria di sviluppo ad U' (o ad N) la quale ha un ruolo cruciale nella formulazione di ipotesi teoriche circa i meccanismi coinvolti nei processi cognitivi<sup>6</sup>.

#### 2.2.1 Il naming explosion: fisionomia dell'evoluzione di un fenomeno linguistico

Un esempio empirico di quanto descritto lo si può rintracciare nella specifica traiettoria ad U relativa al più ristretto dominio del processo di apprendimento del linguaggio nella fase dell'esplosione del vocabolario e degli errori linguistici (*naming errors*) ad esso connessi. Puntando la lente su uno di quei momenti critici che rappresentano i due versanti opposti del percorso evolutivo<sup>7</sup> delle prestazioni in un dominio circoscritto, il dominio linguistico, si potrà comprendere con maggior chiarezza come si configura la parabola fin qui discussa.

La fase critica in questione è la cosiddetta esplosione del lessico (*vocabulary spurt*) che è da rintracciare in senso ampio nella fascia d'età che va dai 18 ai 30 mesi e il cui picco quantitativo si assesta mediamente nel periodo compreso tra i 18 e i 21 mesi. Si tratta di quello stadio in cui i bambini attraversano un repentino passaggio che interessa il loro vocabolario sia nella produzione che nella comprensione. Generalmente l'incremento del vocabolario ad un primissimo stadio, che è da rintracciare nei 4-8 mesi precedenti al *vocabulary spurt* e che solitamente viene indicato con l'espressione *one-word stage*, è molto lento e a livello quantitativo la vera e propria esplosione si verifica generalmente quando tale vocabolario raggiunge un numero compreso tra le cinquanta e le cento parole. Ora, nel momento in cui si verifica una veloce crescita della massa del vocabolario specularmente, in contemporanea al raggiungimento dell'acme del processo esplosivo, in contesti in cui il bambino si trova a dover denominare oggetti del mondo con dei termini che già fanno parte del suo lessico, ha luogo un cambiamento nelle abilità del bambino nel recuperare i nomi dalla memoria. In altri termini, avviene un incremento temporaneo degli errori di denominazione nel momento in cui il vocabolario produttivo viene, nella fase esplosiva, inondato da nuove numerose componenti lessicali. Man mano che il bambino continua ad usare gli elementi lessicali origine degli errori, la frequenza e la quantità degli stessi comincia a diminuire e le performance corrette cominciano a stabilizzarsi<sup>8</sup>. Gli errori di denominazione, dunque, sono da localizzare in una fetta

6. In Rogers *et al.* (2004) si mette in atto una simulazione che illustra come un sistema di apprendimento generale è in grado di produrre una traiettoria di sviluppo ad U man mano che esso progredisce in un compito di categorizzazione di un dato numero di oggetti (animali e piante) sulla base delle loro proprietà fisiche specifiche.

7. Quello ascendente con l'incremento quantitativo e qualitativo della performance e quello discendente con il deterioramento quantitativo e qualitativo della stessa.

8. Secondo versante della ormai famosa traiettoria ad U.

temporale ristretta, in un punto in cui i bambini trovandosi con un vocabolario quantitativamente numeroso e in rapida crescita si trovano per la prima volta a dover nominare oggetti in scansioni temporali tra loro molto vicine e a ritmi più veloci. Sulla base di analisi ben precise (Gershkoff-Stowe & Smith, 1997) proprio il ruolo centrale che gioca il tempo (tentativi di nominare oggetti sempre più vicini temporalmente – quindi ritmi di performance più sostenuti) determina la natura specifica dell'errore che empiricamente appare come l'errata attribuzione di un'etichetta linguistica ad un determinato oggetto (*labeling*) ma che, in realtà, è strettamente legato al recupero della parola esatta (*retrieval error*) per un determinato oggetto in un particolare contesto. Errore, quindi, non relativo alla conoscenza di una singola parola ma relativo all'accesso alla parola corretta nel lessico mentale e che con molta probabilità nasce da interferenze linguistiche provocate da parole recuperate immediatamente prima nel tempo (*repetition effects*), o provocate da parole simili, fonologicamente o semanticamente, a quella da recuperare<sup>9</sup>.

### 3. Due originali generatori d'ipotesi

Esposte alcune delle problematiche fondamentali relative al processo di apprendimento della lingua, è giunto il momento di mostrare come queste stesse vengono affrontate computazionalmente traendo spunto dalla realizzazione di due modelli di apprendimento linguistico. Il primo è relativo a quella fase transitoria in cui il bambino passa dall'utilizzo delle espressioni monoparola all'utilizzo delle espressioni bi-parola<sup>10</sup>. Il secondo, invece, è relativo all'acquisizione delle prime parole e al ruolo centrale che hanno in questo compito la percezione visiva, la percezione uditiva e le aree cerebrali ad esse dedicate. Dei due modelli si prenderanno in esame principalmente la natura dell'input linguistico a cui essi sono esposti e la natura delle loro prestazioni in seguito alle fasi di addestramento, in modo tale da evidenziare all'interno dei processi di simulazione l'emergenza dei fenomeni fin qui discussi o di fenomeni ad essi connessi.

#### 3.1 Una multirete multimodale

Il modello in questione – una multirete neurale multimodale non supervisionata<sup>11</sup> (Nyamapfene & Ahmad, 2007; Nyamapfene 2009 e 2011) – simula quella transizione linguistica, proprio a ridosso dell'esplosione del vocabolario, che vede il bambino passare dalla fase ad una parola (*one-word stage*) alla fase successiva a due parole. Esso, inoltre, ha come obiettivo primario generare comportamenti specifici usando processi simili a quelli che si verificano nel cervello<sup>12</sup>. L'assunto teorico primario è la multimodalità del processo cognitivo che la rete deve simulare, dove con il termine "multimodalità" qui ci si riferisce precisamente a quel processo che consente di stabilire delle connessioni tra varie modalità e fonti di informazione agevolando la possibilità di passare da un modalità all'altra (Nyamapfene, 2011: 1614). È

9 Per uno studio dettagliato sulle varie tipologie di interferenze linguistiche cfr. Gershkoff-Stowe & Smith (1997: 52-54)

10 In questa sede ci si limiterà a prendere in considerazione solamente quella parte del modello dedicata alla simulazione della fase monoparola.

11 Cfr. Nyamapfene (2007; 2009; 2011).

12 Si prendono le distanze da quegli approcci secondo cui i processi multimodali sono processati nel cervello da diversi moduli specifici legati tra di loro. Difatti, uno degli obiettivi primari del modello in questione è la fedeltà ai processi cerebrali mantenendosi nella sfera della plausibilità psicologica e biologica, in riferimento ai più recenti risultati delle tecniche di neuroimaging dai quali emerge il fatto che gli input di diversa natura convergono nello stesso gruppo di rappresentazioni che, quindi, vengono ad essere rappresentazioni amodali (modal-nonspecific). E ciò equivale a dire che si tenta di simulare un sistema di rappresentazione concettuale che in partenza sembra abbastanza indifferenziato. (Nyamapfene, 2011:1615)

questo, difatti, un tratto primario dei processi cognitivi che per natura sono multimodali e ciò è evidenziato dall'abilità che ha il cervello di integrare informazioni provenienti da input di diversa natura. Secondo tale assunto, e secondo quanto si è sostenuto nelle sezioni precedenti, l'acquisizione del linguaggio bisogna vederla, precisamente, come un processo multimodale in cui i vari elementi sono intrinsecamente connessi.

In riferimento a ciò la rete codifica le prime espressioni del bambino come unità multimodali complesse a tre dimensioni rappresentanti l'entità percettiva di riferimento, l'intenzione comunicativa, l'espressione fonologica.

#### 3.1.1 L'input complesso della rete e la sua traiettoria d'apprendimento

A livello di architettura il modello in esame è una multi-rete costituita da tre reti neurali in connessione reciproca di cui la prima simula il comportamento linguistico nella fase ad una parola, la seconda simula il comportamento linguistico nella fase a due parole e la terza mette in connessione reciproca le due precedenti<sup>13</sup>.

Nella simulazione della fase ad una parola la rete dedicata a tale compito viene addestrata, come anticipato, su un gruppo di dati trimodale comprendente ciascuno (1) l'entità percettiva a cui un'espressione a parola singola si riferisce, (2) la relazione concettuale che si inferisce dal contesto comunicativo da cui l'espressione è stata estratta, (3) l'espressione vera e propria.

(1) Per quanto riguarda le entità viene usato uno schema di codifica costituito da otto categorie base: (i) persona, (ii) animale, (iii) veicolo, (iv) arredamento, (v) indumento, (vi) cibo, (vii) parte del corpo, (viii) luogo. Ogni categoria possiede delle caratteristiche che distinguono le entità che ne fanno parte. Alcune di queste caratteristiche sono specifiche di una sola categoria, mentre altre si spalmano su diverse categorie. Nel caso della categoria 'animale', ad esempio, viene codificato se l'entità in questione è munita di corna o no, se possiede una pelliccia folta oppure no, ecc. La rete, durante la fase di addestramento, impara a discriminare la presenza-assenza di caratteristiche in generale, la presenza-assenza di una caratteristica specifica, la presenza-assenza di caratteristiche non specifiche (Nyamapfene, 2011: 1625).

(2) Relativamente alle relazioni concettuali si segue l'assunto secondo cui in questa fase linguistica il bambino impara a rappresentare le regolarità nel suo ambiente in termini di relazioni tra persone, oggetti ed eventi del mondo<sup>14</sup> (Bloom, 1973). Ciò equivale, secondo questa prospettiva, al tentativo del bambino di esprimere una propria intenzione comunicativa, dei tratti descrittivi di un evento, dei tratti peculiari di uno stato fisico, per mezzo di una parola che, quindi, rappresenta la realizzazione linguistica di una relazione concettuale determinata. La rete durante la fase di addestramento impara a discriminare la presenza-assenza di una specifica intenzione comunicativa<sup>15</sup>, la presenza-assenza di tratti descrittivi di un evento<sup>16</sup>, la presenza-assenza di tratti peculiari di uno stato fisico (Nyamapfene, 2011: 1628).

(3) Infine, in relazione all'espressione reale, alla parola vera e propria, lo schema di codifica dell'input della rete rappresenta ogni parola tenendo conto di tre peculiarità fondamentali. La prima riguarda le consonanti le quali vengono distinte in base alla modalità di articolazione (le caratteristiche fonetiche composte dai tratti acustici che costituiscono il suono)<sup>17</sup> e in base alle proprietà articolatorie (caratteristiche relative al modo in cui il tratto

13. Per un'analisi dettagliata delle diverse tipologie di reti utilizzate cfr. Nyamapfene, 2011.

14. Su un oggetto si esercita un'azione, può esistere, cessare di esistere o ricorrere in diversi eventi, apparire e sparire.

15. Commentare, nominare, esprimere una posizione nello spazio, esprimere un possesso, formulare una richiesta, esprimere un rifiuto.

16. Ricorrenza, esistenza, non-esistenza, sparizione, cessazione di un evento.

17. «Nasal, stop, fricative, approximant and lateral» (Nyamapfene, 2011: 1629).



vocale, la bocca, la lingua e gli organi associati creano il suono)<sup>18</sup>. La seconda riguarda le vocali le quali sono distinte in base all'altezza<sup>19</sup> e alla posizione della lingua<sup>20</sup>(Ladefoged, 1982). La terza, infine, è relativa allo status del fonema che rende chiaro se un determinato fonema è una vocale, una consonante sorda o una consonante sonora (Li & MacWhinney, 2002).

Ricevuto un tale input la rete dedicata alla simulazione della fase mono-parola, a seguito di un ben preciso addestramento, mostra un comportamento che sembra simulare fedelmente lo sviluppo del vocabolario produttivo nel bambino, andando incontro, per qualsiasi valore del ritmo di apprendimento, ad un periodo iniziale ad alta percentuale di errore che va scemando man mano che l'addestramento prosegue. Essa segue, coerentemente con quanto precedentemente esposto, una traiettoria evolutiva ad U e mostra una simulazione psicologicamente plausibile (a) dell'abilità di padroneggiare l'uso delle espressioni mono-parola e (b) una simulazione psicologicamente plausibile del deterioramento della capacità di recupero della parola corretta, entrambi fenomeni che si potrebbero indicare come limite interno e limite esterno dell'esplosione del lessico.

### 3.2 Una multirete che vede e sente

Il secondo modello che verrà preso in esame è una rete che simula i processi corticali di auto-organizzazione che si dispiegano lungo la fase dell'apprendimento delle prime parole (Plebe *et al.*, 2010). La peculiarità primaria del modello in questione la si deve rintracciare nel tentativo di simulare il più fedelmente possibile il contributo che alcune aree della corteccia forniscono nell'emergenza delle prime parole. Tale modello, quindi, evidenzia la centralità dei meccanismi interni del cervello sottolineando la plasticità cerebrale come una delle proprietà fondamentali dei processi cognitivi in generale e dell'apprendimento del linguaggio in particolare. La realizzazione della rete nasce dalla connessione di due differenti approcci di cui uno adatto a simulare le mappe fisicamente esistenti nella corteccia e l'altro adatto ad astrarre funzioni di livello superiore che non sono processate in aree definite del cervello ma che sono distribuite con molta probabilità in diverse zone (Plebe *et al.*, 2010: 219).

L'architettura del modello consta di una collezione di reti<sup>21</sup> organizzate in modo da costituire una versione semplificata dei percorsi visivi e uditivi della corteccia cerebrale. I percorsi principali sono due: uno relativo al processo visivo (LGN)<sup>22</sup>, l'altro relativo al percorso uditivo (MGN)<sup>23</sup>. Per quanto riguarda il percorso visivo i processi ad esso relativi confluiscono in una mappa corticale che simula l'area che per prima, nell'uomo, viene coinvolta nel riconoscimento degli oggetti (LOC)<sup>24</sup>. In riferimento, invece, al percorso uditivo i processi ad esso relativi confluiscono in una mappa (STS)<sup>25</sup> che simula i processi della principale area cerebrale sensibile ai suoni vocali. Le due mappe gerarchicamente superiori (LOC, STS) sono quelle che nel modello forniscono rispettivamente la rappresentazione delle caratteristiche visive di un oggetto e la rappresentazione delle caratteristiche della parola. Queste due rappresentazioni, in più, sono connesse in una map-

pa astratta (ACM)<sup>26</sup> che è l'unica componente non riconducibile ad una precisa area del cervello essendo un'astrazione di processi che interessano diverse aree cerebrali (Plebe *et al.*, 2010: 221).

#### 3.2.1 L'input bimodale nell'addestramento del modello

Nell'addestramento della multirete viene seguita una procedura che riproduce, in versione semplificata, le varie tappe dell'ontogenesi del linguaggio che conducono all'apprendimento delle prime parole: (1) prenatale, (2) prelinguistica e (2) linguistica vera e propria.

(1) Durante la fase pre-natale gli stimoli/input presentati alle mappe che costituiscono il percorso visivo (V0,V1,V2) sono macchie casuali che riproducono onde di attività spontanea e che giocano un ruolo essenziale nello sviluppo precoce del sistema visivo. Sul versante del percorso uditivo l'input consta di catene di onde sonore (Plebe *et al.*, 2010: 222).

(2) Per quanto riguarda la fase pre-linguistica gli input relativi al percorso visivo, diretti alla mappa che simula il LOC, sono costituiti da immagini di oggetti ritratti da otto prospettive differenti<sup>27</sup> con lo scopo di modellare il più realisticamente possibile le diverse facoltà coinvolte nel processo di riconoscimento. Omologamente la mappa che simula MGN e tutte le altre mappe del percorso uditivo ricevono come input parole di lunghezza compresa tra i tre e i dieci caratteri convertite in onde sonore. In questa fase, al termine dell'addestramento, accade che ogni mappa evolve nelle proprie specifiche funzioni<sup>28</sup> tra le quali una delle più importanti è l'invarianza visiva. Una proprietà mostrata dalla mappa che simula il LOC, riscontrata anche nel LOC umano, e che consiste nella capacità che ha un neurone di rispondere a delle caratteristiche specifiche di un oggetto nonostante i mutamenti dello stesso dovuti ai differenti punti di vista.

(3) La fase linguistica, infine, coinvolge direttamente la mappa delle astrazioni ad alto livello (ACM) e comprende la simulazione del processo di denominazione ostensiva intenzionale di un oggetto e la simulazione del associazione casuale di schemi sonori a scene naturali. L'input è formato (a) da 100 oggetti raggruppati in 38 categorie e dai rispettivi nomi convertiti in onde sonore e (b) da scene visive<sup>29</sup> e suoni<sup>30</sup> non connessi tra loro.

Nelle prime fasi d'addestramento il modello viene esposto molto più spesso a coincidenze casuali tra suoni e scene e meno alla denominazione intenzionale degli oggetti. Man mano che ci si avvicina alle fasi finali, invece, la pratica si inverte privilegiando un input linguistico in cui il rapporto oggetto-etichetta linguistica non è casuale.

Alla fine del processo di addestramento<sup>31</sup>, la multi-rete mostra una buona capacità di operare delle generalizzazioni, di estendere un nome a un nuovo oggetto, già sulla base di brevi esposizioni ad un solo esemplare di una nuova categoria, evento questo che si configura come una buona simulazione del fenomeno del *fast mapping*<sup>32</sup> riscontrabile nel processo reale di ontogenesi e tutto interno al fenomeno dell'esplosione del lessico.

## 4. Le prossime sfide computazionali

Nonostante la soluzione della multirete consenta di formulare nuove ipotesi e di suggerire nuove soluzioni relative a problemi centrali come quelli messi in campo nelle sezioni precedenti, la strada che si pone davanti ai modelli connessionisti è in salita e

18. «Bilabial, labio-dental, dental, alveolar, palatoalveolar, palatal, velar and glottal» (Nyamapfene, 2011: 1629).

19. «High, mid-high, mid, mid-low and low» (Nyamapfene, 2011: 1629).

20. «Front, central and back» (Nyamapfene, 2011: 1629).

21. LISSOM (*Laterally Interconnected Synergetically Self-Organizing Map*) ad eccezione della mappa che simula i processi di astrazione che è una SOM (*Self-Organizing Map*).

22. Nucleo genicolato laterale (LGN).

23. Nucleo genicolato mediale (MGN)

24. Complesso occipitale laterale (LOC). I processi che confluiscono nel LOC viaggiano lungo due flussi: la componente acromatica processata dall'area visiva primaria (V1) e dall'area visiva secondaria (V2) e le due componenti spettrali processate dall'area occipitale ventrale (VO) (cfr. Plebe *et al.*, 2010: 220).

25. Solco temporale superiore (STS).

26. Mappa categoriale astratta (ACM).

27. Estratti dal database COIL-100 (Murase & Nayar, 1995).

28. Per un resoconto dettagliato dell'evoluzione delle funzioni delle diverse aree cerebrali del percorso visivo e uditivo cfr. Plebe *et al.* (2010: 222).

29. Immagini di paesaggi e fiori.

30. Suoni naturali.

31. Cfr. Plebe *et al.* (2010:222-224).

32. «It has been shown that children [...] can correctly and consistently map a novel word to a novel object in the presence of other familiar object.» (Alishahi, 2011:26).

piena di difficoltà. Pur essendo i modelli connessionisti, in definitiva, modelli dell'apprendimento, bisogna considerare il fatto che essi sono comunque delle semplificazioni relative a fenomeni molto complessi. Difatti, al di là della semplificazione offerta da un modello, lo sviluppo cognitivo e linguistico di ogni bambino si presenta come un processo che si dispiega su diversi livelli e a diversi gradi di complessità. A livello temporale, ad esempio, esso sembra un processo di integrazione costante in cui ci sono componenti che si dispiegano lungo scale temporali differenti. Ci sono componenti che si dispiegano lungo l'asse dell'immediatezza che orientano il processo nell'immediato e che possono essere identificate con l'input percettivo immediato – l'oggetto e le sue proprietà sensibili tra le quali facciamo rientrare anche la percezione del suo nome. Asse temporale immediato che si dispiega nel dominio della consuetudine, della reiterazione e della sensibilità al contesto che di per sé necessita e, nello stesso tempo, contribuisce a strutturare la storia a lungo termine di chi apprende facendo emergere regolarità di relazione tra le parole, le proprietà degli oggetti e le organizzazioni categoriali che sono acquisite su una scala temporale molto meno frenetica. Spostandosi su un altro livello, invece, bisogna tenere in considerazione anche il fatto che i bambini avendo pulsioni, desideri e attenzioni nel processo di crescita recitano un ruolo di attori e non sono mere entità passive che assorbono semplicemente stimoli esterni più o meno complessi.

Bene, i modelli computazionali, quelli connessionisti in particolare, sapranno farsi carico di questi problemi?

Sapranno accettare queste sfide?

#### **Bibliografia:**

- Alishahi, A. (2011). *Computational Modeling of Human Language Acquisition*. Morgan & Claypool Publishers.
- Barsalou, L.W., Breazeal, C., & Smith, L.B. (2007). Cognition as Coordinated Non-Cognition. *Cognitive Processes*, 8, 79-91.
- Bloom, L. (1973). *One Word at a Time: The Use of Single-Word Utterances before Syntax*. The Hague, Mouton.
- Colunga, E., & Smith, L.B. (2008). Knowledge Embedded in Process: The Self-Organization of Skilled Noun Learning. *Developmental Science*, 11, 2, 195-203.
- Elman, J. (2009). On the Meaning of Words and Dinosaur Bones: Lexical Knowledge without a Lexicon. *Cognitive Science*, 33, 1-36.
- Elman, J. (2006). Computational Approaches to Language Acquisition. *Encyclopedia of Language & Linguistics*, 2, 726-732.
- Elman, J. (2005). Connectionist Models of Cognitive Development: Where Next. *Trends in Cognitive Science*, 9, 3, 111-117.
- Elman, J. (2004). An Alternative View of Mental Lexicon. *Trends in Cognitive Science*, 7, 301-306.
- Elman, J. (1995). *Language as a Dynamical System*. In Port, R., & van Gelder, T. (eds.). *Mind as Motion*. Cambridge, Mass.: MIT Press.
- Gershkoff-Stowe, L., & Smith, L.B. (1997). A Curvilinear Trend in Naming Errors as a Function of Early Vocabulary Growth. *Cognitive Psychology*, 34, 37-71.
- Jackendoff, R. (2007). A Parallel Architecture Perspective on Language Processing. *Brain Research*, 1146, 2-22.
- Ladefoged, P. (1982). *A Course in Phonetics*. Harcourt Brace Jovanovich: New York.
- Lenci, A. (2006). The Lexicon and the Boundaries of Compositionality. In Tuomo, A., & Ahti-Veikko, P., (eds.). *Truth and Games. Essays in Honor of Gabriel Sandu*. Helsinki: Helsinki University Press, (pp. 119-138).
- Li, P., & MacWhinney, B. (2002). PatPho: A Phonological Pattern Generator for Neural Network. *Behavior Research Methods, Instruments, and Computers*, 34, 408-415.
- Nayar, S., & Murase, H. (1995). Visual Learning and Recognition

of 3-D Object Coherence. *International Journal of Computer Vision*, 14, 5-24.

- Nyamapfene, A. (2011). Towards Understanding Child Language Acquisition: An Unsupervised Multimodal Neural Network Approach. *Journal of Information Science and Engineering*, 27, 1613-1639.
- Nyamapfene, A. (2009). Computational Investigation of Early Child Language Acquisition Using Multimodal Neural Networks: A Review of Three Models. *Artificial Intelligence Review*, 31, 35-44.
- Nyamapfene, A., & Ahmad, K. (2007). A Multimodal Model of Child Language Acquisition at the One-Word Stage. *Proceedings of International Joint Conference on Neural Networks*, Orlando, Florida, USA, 1-6.
- Plebe, A., Mazzone, M., & De la Cruz, V. (2010). First Words Learning: A Cortical Model. *Cognitive Computation*, 2, 217-229.
- Rogers, T. T., Rakison, D. H., & McClelland, J. L., (2004). U-Shaped Curves in Development: A PDP Approach. *Journal of Cognition and Development*, 1, 5, 137-145.
- Samuelson, L.K., & Smith, L.B. (2000). Grounding Development in Cognitive Processes. *Child Development*, 71, 1, 98-106.
- Smith, L. B. (2005). Cognition as a Dynamic System: Principles from Embodiment. *Developmental Review*, 25, 278-298.