

Un modello neurale per spiegare le asimmetrie tra produzione e comprensione linguistica

Marco Mirolli (marco.mirolli@istc.cnr.it)

Istituto di Scienze e Tecnologie della Cognizione, CNR,
Via S. Martino della Battaglia 44, 00185, Roma

Federico Cecconi (federico.cecconi@istc.cnr.it)

Istituto di Scienze e Tecnologie della Cognizione, CNR,
Via S. Martino della Battaglia 44, 00185, Roma

Domenico Parisi (domenico.parisi@istc.cnr.it)

Istituto di Scienze e Tecnologie della Cognizione, CNR,
Via S. Martino della Battaglia 44, 00185, Roma

Sommario

Evidenze empiriche di vario tipo mostrano l'esistenza di un'asimmetria tra produzione e comprensione linguistica: in generale, comprendere parole sembra essere più facile che produrle. In questo contributo proponiamo un modello neurale dei rapporti tra semantica e lessico che spiega l'asimmetria tra produzione e comprensione del linguaggio come un effetto della differenza di dimensioni tra le aree del cervello deputate al processamento dell'informazione lessicale e quelle deputate al processamento dell'informazione semantica. Il modello dimostra una maggiore difficoltà nel richiamo lessicale rispetto al richiamo semantico grazie all'assunzione che la rete semantica è costituita da più unità computazionali (neuroni) rispetto alla rete lessicale.

Abstract

Several kinds of empirical evidence show the existence of an asymmetry between linguistic production and comprehension: in general, understanding words seems to be easier than producing them. In this contribution we propose a neural model of the relationships between the semantic and the lexical systems. Our model explains the asymmetry between language comprehension and production as an effect of the difference between the dimensions of the brain areas which process semantic and lexical information. In fact, the model's performance in lexical recall is worse than in semantic recall due to the fact that the semantic network is constituted of more computational units (neurons) than the lexical network.

Introduzione

Un gran numero di evidenze empiriche di vario tipo mostrano l'esistenza di un'asimmetria tra produzione e comprensione del linguaggio: in generale, la comprensione sembra essere più facile della produzione.

Questa asimmetria si manifesta in varie forme. Innanzitutto, tale asimmetria è ben documentata in ambito evolutivo: è un dato infatti ben noto che i bambini apprendono prima a comprendere le parole che a produrle (vedi, ad esempio, Bates, Thal, Finlay & Clancy, 2002).

Un altro ambito in cui tale asimmetria è ben documentata è quello dell'invecchiamento: con l'avanzare dell'età diventa infatti sempre più difficile 'trovare le parole', mentre non sembra esserci deterioramento della

capacità di comprendere le parole sentite (vedi, ad esempio Burke & MacKay, 1997; Burke, MacKay & James, 2000).

Inoltre, la maggiore difficoltà a produrre che a comprendere sembra manifestarsi in generale, durante tutta la vita. Tutti sperimentiamo infatti il fenomeno di avere una parola 'sulla punta della lingua', e le evidenze empiriche confermano che tale fenomeno aumenta con l'età ma è presente anche in adulti non anziani (vedi, ad esempio, Brown, 1991; Heine, Ober & Shenaut, 1999).

Infine, forse tale asimmetria esiste anche nelle neuropatologie, con un maggior numero di pazienti con difficoltà di produzione di quelli con difficoltà di comprensione: ad esempio, sembra che *tutti* i tipi di afasia comportino la difficoltà a produrre nomi (anomia), mentre una disfunzione nella produzione linguistica non è necessariamente correlata con difficoltà in comprensione (vedi Bates e Goodman, 1997; Dick, Bates, Wulfeck, Utman & Dronkers, 2001).

Nonostante i fatti citati riguardino fenomeni molto diversi fra loro, il fenomeno dell'asimmetria è comune a tutti, e come tale potrebbe avere una causa comune. In questo contributo proponiamo un modello neurale dei rapporti tra semantica e lessico che spiega l'asimmetria tra produzione e comprensione linguistica -- in particolare la maggiore facilità della comprensione rispetto alla produzione -- come effetto della differente dimensione degli spazi computazionali deputati al processamento dell'informazione semantica (significati delle parole) e dell'informazione lessicale (segnali delle parole). In particolare, il modello assume che le parti del cervello deputate a processare i suoni delle parole siano più piccole, in termini di numero di neuroni, delle parti del cervello deputate a processare il significato delle parole. La nostra ipotesi è che l'asimmetria tra produzione e comprensione sia determinata da questa differenza di grandezza tra l'area semantica e quella lessicale.

Il resto del contributo è organizzato nel seguente modo. Nella prossima sezione descriviamo in generale il modello che abbiamo utilizzato per modellare i sistemi 'semantico' e 'lessicale', ossia le reti di Hopfield, e poi descriviamo i dettagli delle nostre simulazioni. Nella sezione successiva descriviamo i risultati principali delle simulazioni e, in quella conclusiva, discutiamo pregi e limiti del modello e possibili sviluppi futuri.

Il modello

Modelliamo il sistema linguistico come l'accoppiamento (collegamento reciproco) tra due reti auto-associative, o reti di Hopfield (Hopfield, 1982; Rolls & Treves, 1998): la rete semantica e la rete lessicale.

Le reti di Hopfield

Una rete di Hopfield è una rete neurale costituita da un unico gruppo di neuroni reciprocamente interconnessi. Ogni neurone è connesso con tutti gli altri, tranne che con se stesso, e le connessioni sono simmetriche: il peso della connessione che va dal nodo i al nodo j è uguale al peso della connessione che va dal nodo j al nodo i . La simmetria delle connessioni è garantita dalla regola di apprendimento che viene usata per addestrare la rete, che è una semplice regola di Hebb (Hebb, 1949). Informalmente, la regola di Hebb consiste nel fatto che se una connessione unisce due nodi aventi lo stesso valore di attivazione (1 e 1, oppure -1 e -1), il peso della connessione viene aumentato, mentre se unisce due nodi aventi valore opposto (1 e -1, oppure -1 e 1), il peso viene diminuito. Formalmente, la regola di Hebb è espressa dalla seguente formula:

$$\Delta w_{ij} = r a_i a_j$$

dove Δw_{ij} è la modifica del peso dall'unità i all'unità j , a_x è l'attività dell'unità x e r è una costante, chiamata il tasso di apprendimento.

La rete apprende a memorizzare un certo numero di pattern di attivazione, che consistono in vettori bipolari (composti da 1 o -1) di N valori, dove N è il numero di nodi della rete. Nel nostro caso, i pattern di attivazione rappresentano la rappresentazione interna del significato di una parola per quanto riguarda la rete semantica e la rappresentazione interna del suono della parola per quanto riguarda la rete lessicale.

L'apprendimento dei pesi avviene nel seguente modo. All'inizio i pesi delle connessioni della rete sono tutti a zero. Poi, per ciascun pattern di attivazione, i nodi della rete vengono attivati ai valori bipolari di quel pattern ed i pesi vengono modificati tramite la regola di Hebb.

Una volta addestrate, le reti di Hopfield hanno un'interessante proprietà. Se presentiamo un qualunque pattern in ingresso e lasciamo poi la rete libera di ricalcolare lo stato di attivazione di ciascun nodo in base ai pesi delle connessioni che il nodo riceve e all'attivazione degli altri nodi, dopo un certo numero di cicli (in genere inferiore a 15), la rete converge su uno stato stabile, chiamato attrattore, che non cambia più.

Lo stato delle attivazioni dei neuroni di una rete di Hopfield può essere concepito come una superficie in uno spazio multidimensionale, con una dimensione che rappresenta l'energia della rete (Figura 1). Ciascun punto della superficie è un possibile stato della rete e i punti di minimo (le valli) sono gli attrattori: dato un qualsiasi pattern la rete tenderà a scendere lungo il gradiente di energia fino a raggiungere una valle, dove, a meno di interferenze esterne, si fermerà indefinitamente.

L'apprendimento fa sì che i pattern su cui una rete viene addestrata diventino gli attrattori della rete. Pertanto, se

presentiamo alla rete uno dei pattern appresi, la rete lo manterrà indefinitamente. E se presentiamo un pattern parziale, ossia un pattern in cui una certa percentuale di nodi è non attiva (attivazione a 0), dopo un certo numero di cicli di attivazione dei nodi la rete sarà in grado di ricostruire perfettamente il pattern appreso, che viene poi mantenuto costante nei cicli successivi.

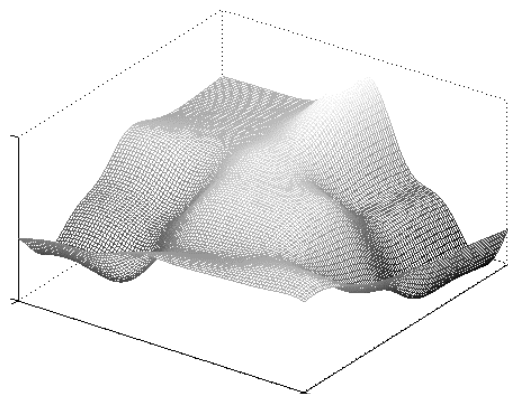


Figura 1: Ipotetico spazio di attivazione di una rete di Hopfield a due dimensioni. La terza dimensione (asse verticale) rappresenta l'energia del punto che indica lo stato corrispondente della rete. Le valli rappresentano gli stati stabili verso cui l'attività della rete converge, chiamati attrattori.

La scelta di modellare i sistemi semantico e lessicale tramite reti di Hopfield è dovuta al fatto che tali reti rappresentano degli ottimi modelli del funzionamento di singole aree del cervello (Rolls and Treves, 1998). Oltre a possedere tutte quelle proprietà biomimetiche che anche le altre reti neurali hanno, quali la robustezza al rumore, il degradamento graduale e la capacità di completare i pattern incompleti, le reti di Hopfield hanno almeno altre due proprietà che le rendono biologicamente plausibili. Innanzitutto, della regola di apprendimento hebbiana è stato trovato il corrispettivo neurale nei fenomeni di potenziamento e depotenziamento a lungo termine delle sinapsi tra neuroni (Kelso, Ganong & Brown, 1986; Stanton, & Sejnowski, 1989). Inoltre, la ricorsività delle connessioni, ossia il fatto che neuroni di uno stesso gruppo siano reciprocamente connessi tra loro, rappresenta un'importante caratteristica anatomica del cervello, e fa sì che la rete dimostri una dinamica temporale che permette di dar conto di importanti fenomeni reali quali il priming o, più in generale, le differenze nei tempi di reazione ad uno stimolo (vedi, ad esempio, Masson, 1995; Sharkey & Sharkey, 1992).

La rete semantica e la rete lessicale

Il nostro modello consiste in due reti di Hopfield, che rappresentano, rispettivamente, il sistema semantico e quello lessicale. Il funzionamento delle due reti è del tutto simile. L'unica differenza tra le due reti sta nelle dimensioni. La rete semantica è infatti costituita da 2500 nodi, mentre la rete lessicale è costituita da 500 nodi. Le due reti sono accoppiate, ossia reciprocamente connesse

(Figura 2). Ma, mentre i nodi che appartengono ad una stessa rete sono tutti reciprocamente connessi tra loro, i nodi che appartengono alle due diverse reti hanno connettività molto bassa (tra lo 0.02 e lo 0.001, vedi sotto).

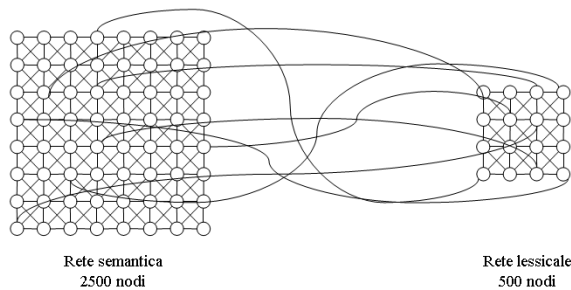


Figura 2: Rappresentazione schematica del modello. I cerchi rappresentano nodi (neuroni), mentre le linee rappresentano connessioni tra nodi. All'interno delle due sotto-reti i neuroni sono completamente connessi (tutti con tutti), mentre la connettività tra le due reti è molto bassa. Vedi testo per i dettagli.

Ciascuna sotto-rete è addestrata a memorizzare 50 pattern diversi: la rete semantica apprende 50 pattern (vettori bipolari) di 2500 valori, ciascuno rappresentante un 'significato', mentre la rete lessicale apprende 50 pattern di 500 valori, ciascuno rappresentante una 'parola'. Tutti i pattern sono generati casualmente, con una probabilità uniforme, per ciascun nodo, di essere nello stato attivo (1) o inattivo (-1).

Ad ogni parola è associato un significato, e la rete complessiva, formata dalle due sotto-reti semantica e lessicale, apprende tale associazione nel seguente modo. Per ciascuna parola, il pattern corrispondente è presentato alla rete lessicale mentre simultaneamente il pattern corrispondente al significato della parola viene presentato alla rete semantica, e i pesi che collegano i nodi delle due sotto-reti vengono appresi, in modo del tutto simile ai pesi interni delle due sotto-reti, tramite la regola di Hebb descritta sopra.

Risultati

Nel nostro modello la produzione linguistica consiste nell'attivare il pattern nella rete lessicale (suono) che corrisponde al pattern che è presente nella rete semantica (significato). Viceversa, la comprensione linguistica consiste nell'attivare il pattern nella rete semantica (significato) che corrisponde al pattern presente nella rete lessicale (suono).

Abbiamo fatto due test, uno di richiamo lessicale e uno di richiamo semantico. Nel test di richiamo lessicale (produzione linguistica), per tutti i pattern che rappresentano il significato di una parola, il pattern viene immesso nella sotto-rete semantica, mentre tutti i nodi della rete lessicale vengono attivati a 0. La rete complessiva viene quindi fatta 'rilassare', ossia l'attivazione dei nodi viene aggiornata finché la rete non ha raggiunto un attrattore. A questo punto verificiamo se il pattern che si è attivato nella rete lessicale è quello del suono della parola corrispondente a quel significato. Il test

di richiamo semantico (comprensione linguistica) avviene in maniera inversa a quello di richiamo lessicale: dato un suono immesso nella sotto-rete lessicale, guardiamo se nella sotto-rete semantica viene attivato il significato corretto.

Per prima cosa abbiamo verificato se, in presenza di un'adeguata connettività tra le due sotto-reti, il nostro modello ha acquisito sia la capacità di comprendere il significato di una parola, sia la capacità di produrre la parola associata ad un significato. Questo è in effetti ciò che succede: ad esempio, con una probabilità di connessione inter-rete dello 0.2, la rete è in grado sia di comprendere che di produrre tutte le parole apprese.

Per verificare la presenza di eventuali asimmetrie fra produzione e comprensione abbiamo quindi studiato il comportamento della rete con gradi di connettività molto bassi tra le due sotto-reti, in particolare, con probabilità di connessione tra 0.020 e 0.001. Tale basso grado di interconnettività vuole riflettere, nel modello, la progressiva diminuzione di collegamenti sinaptici che si manifesta con l'avanzare dell'età.

Per entrambe le condizioni di richiamo semantico e di richiamo lessicale, e per ciascun grado di connettività, prendiamo due misure: (a) la percentuale di pattern correttamente recuperati e (b) il tempo medio di recupero dei pattern accuratamente recuperati (misurato come il numero di cicli necessari al rilassamento).

La Figura 3 mostra la percentuale di pattern recuperati nelle due condizioni in funzione del tasso di connettività tra le due sotto-reti, mentre la Figura 4 mostra il tempo medio di recupero. I dati riportati in entrambe le figure rappresentano le medie di 10 repliche di ciascun test con valori iniziali casuali diversi.

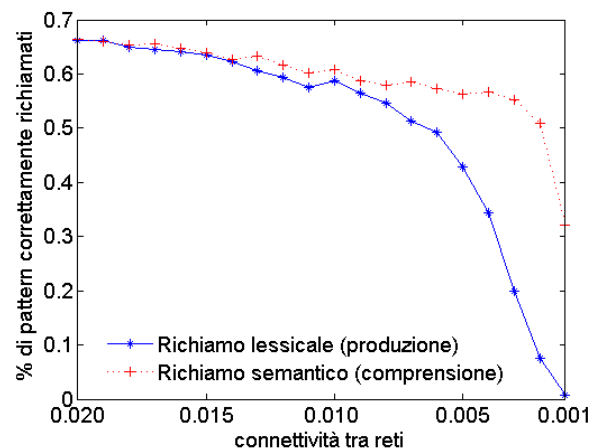


Figura 3: Percentuale di pattern recuperati correttamente.

In generale, i risultati mostrano una chiara asimmetria tra produzione e comprensione, che va nello stesso senso di quella riscontrata nella realtà: la produzione è infatti più difficile della comprensione, sia in termini di numero di pattern corretti richiamati (Figura 3) che in termini di tempo di richiamo (Figura 4).

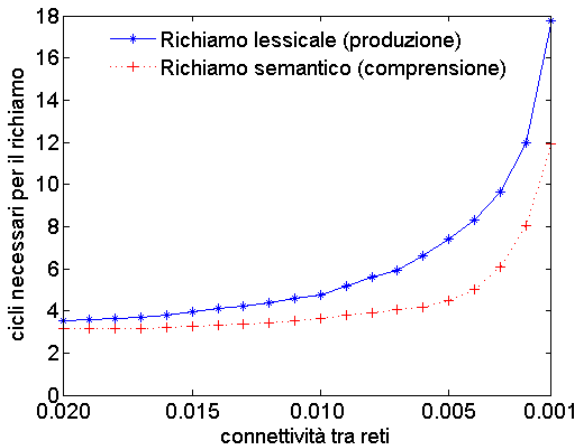


Figura 4: Tempo di recupero dei pattern correttamente recuperati.

Confrontiamo, ad esempio, i risultati delle due condizioni con un grado di connettività tra le due reti pari a 0.002. In tal caso, mentre dato il suono di una parola viene richiamato il significato giusto in circa il 50% dei casi, dato un significato la probabilità che venga richiamata la parola corrispondente è circa l'8%. Inoltre, con la stesso grado di connettività, mentre il tempo di recupero medio nel test di richiamo semantico è di circa 8 cicli, il tempo medio nel test di richiamo lessicale è di circa 12 cicli.

Un altro risultato interessante delle nostre simulazioni riguarda il differente profilo mostrato dalle curve di produzione e comprensione via via che la connettività tra le due sotto-reti diminuisce: infatti, le differenze tra richiamo lessicale e richiamo semantico sembrano aumentare parallelamente alla diminuzione del grado di connettività tra le due reti. Ciò è in linea con il fatto che con l'invecchiamento la capacità di produzione linguistica peggiora costantemente, mentre quella di comprensione rimane, almeno apparentemente, stabile: in altre parole, più connessioni vengono eliminate tra la rete semantica e quella lessicale (per effetto di meccanismi di degenerazione cellulare) e maggiore è l'asimmetria tra produzione e comprensione.

Discussione e conclusione

Il modello proposto rappresenta un tentativo di spiegare l'asimmetria tra produzione e comprensione linguistica che si manifesta in varie forme: varie linee di evidenza empirica indicano infatti come sia più difficile produrre parole che comprenderle, sia per i bambini, gli adulti e gli anziani normali che, sembra, per pazienti che hanno subito lesioni cerebrali. Tutte queste linee di evidenza riguardano certamente fenomeni ben diversi, ed è più che probabile che le varie asimmetrie fra produzione e comprensione dipendano anche da fattori diversi. Per fare solo un esempio, un certo grado di asimmetria tra produzione e comprensione nello sviluppo è sicuramente dovuto al fatto che le parole sono apprese socialmente. La capacità di comprendere una parola, ossia di capire in quali contesti essa è applicata correttamente, è certamente un prerequisito per la capacità di produrla nei contesti

corretti. Detto altrimenti, non si possono produrre adeguatamente parole che non si comprendono. Dunque, sembra inevitabile che il numero di parole comprese rappresenti il limite superiore rispetto al numero di parole prodotte, e che la produzione segua cronologicamente la produzione linguistica.

Nonostante ci possano essere motivi differenti dietro alle diverse forme in cui l'asimmetria tra produzione e comprensione linguistica si manifesta, è certamente possibile che tali forme di asimmetria abbiano anche una causa comune.

In questo contributo abbiamo proposto un modello neurale dei rapporti tra semantica e lessico che spiega la maggiore difficoltà di produrre parole rispetto al comprenderle come effetto della grandezza degli spazi computazionali deputati al processamento di semantica e lessico. Assumendo che le parti del cervello deputate a rappresentare i significati delle parole siano considerevolmente più grandi rispetto a quelle deputate al rappresentare le parole stesse (in termini di numero di neuroni coinvolti), abbiamo mostrato che, con il diminuire del numero di connessioni che collegano le due aree il richiamo di parole diventa sempre più difficile, mentre il richiamo dei significati si mantiene più stabile.

Ovviamente, il nostro modello ha diversi limiti. Un primo limite consiste nel fatto che l'asimmetria si manifesta quando ormai la rete è in grado di richiamare (in entrambe le direzioni) solo il 70% delle parole o dei significati appresi. Nella realtà l'asimmetria sembra manifestarsi prima: ad esempio, un anziano inizia ad avere difficoltà con la produzione linguistica quando ancora è in grado di comprendere tutte le parole che conosce. Stiamo sperimentando vari modi in cui tale limite può essere superato: ad esempio, una possibilità consiste nell'aumentare le differenze di grandezza tra le due sotto-reti; un'altra è quella di diminuire la connettività interna alle reti (nel nostro modello, come nelle classiche reti di Hopfield, la connettività interna è completa, ma nel cervello tale connettività è certamente più bassa); ancora un'altra possibilità consiste nel vedere cosa succede nel caso in cui le connessioni tra due neuroni non siano necessariamente simmetriche (rendere cioè possibile che un neurone i sia collegato con un altro neurone j , ma non viceversa: anche questo renderebbe il sistema sicuramente più realistico).

Ci sono poi limiti del modello più generali. Ad esempio, nel presente modello il sistema semantico non ha alcuna strutturazione interna. I significati sono rappresentati come pattern di attivazione distribuiti su tutto il 'sistema semantico'. Nella realtà, sappiamo che diversi tipi di 'significati', ad esempio nomi e verbi, sono rappresentati nel cervello in gruppi di neuroni diversi (vedi, ad esempio, Caramazza & Hillis, 1991). E anche parole dello stesso tipo sembrano essere processate in parti diverse a seconda del loro significato (Martin, Wiggs, Ungerleider, & Haxby, 1996, Pulvermüller, 1999; Tettamanti, Buccino, Saccuman, Gallese, Danna, Scifo, Fazio, Rizzolatti, & Cappa, 2005). Inoltre, la connettività tra diversi sottogruppi di neuroni deputati a processare tipi diversi di parole non è certamente uniforme: non tutti i sottogruppi sono connessi con tutti gli altri, o, perlomeno, la percentuale di connettività tra diversi gruppi è

certamente diversa. Versioni successive del nostro modello possono sicuramente tener conto di questi vincoli empirici, e, di conseguenza, potrebbero contribuire a spiegare altri importanti fenomeni, come le varie forme di dissociazione che si riscontrano in neuropsicologia (Caramazza & Mahon, 2005).

Un altro limite del nostro modello, ancora più generale, consiste nel fatto che le rappresentazioni semantiche (ed anche quelle linguistiche) sono fornite da noi e non vengono apprese (sviluppate) dalla rete stessa (come avviene nella realtà ed anche in molte simulazioni connessioniste). Per risolvere questo tipo di limite il modello andrebbe modificato in modo considerevole, in quanto le reti di Hopfield, essendo costituite da un solo gruppo di neuroni, non sono in grado di costruirsi da sole rappresentazioni interne che medino tra l'input e l'output, e dunque richiedono che i pattern appresi vengano generati direttamente dallo sperimentatore.

Nonostante i limiti qui discussi, riteniamo che il nostro modello rappresenti un interessante tentativo di trovare una spiegazione all'asimmetria ampiamente documentata tra comprensione e produzione linguistica. Un'altra linea di ricerca futura consiste nel fornire un modello più dettagliato delle varie forme in cui tale asimmetria si manifesta: ad esempio, riteniamo sia possibile non solo simulare l'invecchiamento ed il conseguente deterioramento della capacità di trovare le parole come l'effetto della diminuzione di connessioni, ma anche il fenomeno di avere una parola sulla punta della lingua, presente ad ogni età, mediante l'introduzione di rumore (che sappiamo essere ampiamente presente nel cervello), o nella trasmissione dell'informazione tra i neuroni o nella funzione di attivazione dei neuroni stessi.

References

- Bates, E., & Goodman, J.C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes*, 12(5-6), 507-584.
- Bates, E., Thal, D., Finlay, B.L., & Clancy, B. (2002). Early language development and its neural correlates. In I. Rapin, & S. Segalowitz (eds.), *Handbook of neuropsychology*, Vol. 7: Child neurology. Amsterdam: Elsevier.
- Brown, A.S. (1991). A review of the tip-of-the-tongue experience, *Psychological Bulletin*, 109(2), 204-223.
- Burke, D.M., & MacKay, D.G. (1997). Memory, language and ageing. *Philosophical Transactions of the Royal Society: Biological Sciences*, 352, 1845-1856.
- Burke, D.M., MacKay, D.G., & James, L.E. (2000). Theoretical approaches to language and aging. In T. Perfect, & E. Maylor (Eds.), *Models of cognitive aging*. Oxford: Oxford University Press.
- Caramazza, A.H., & Hillis, A.E. (1991). Lexical organization of nouns and verbs in the brain, *Nature*, 349, 788-790.
- Caramazza, A., & Mahon, B. (2005). The organisation of conceptual knowledge in the brain: the future's past and some future directions, *Cognitive Neuropsychology*, 22, 1-25.
- Dick, F., Bates, E., Wulfeck, B., Utman, J., & Dronkers, N. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasics and normals. *Psychological Review*, 108 (4), 759-788.
- Hebb, D. O. (1949). *The Organization of Behavior*. New York: John Wiley.
- Heine, M., Ober, B., & Shenaut, G. (1999). Naturally occurring and experimentally induced tip-of-the-tongue experiences in three adult age groups. *Psychology and Aging*, 14(3), 445-457.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 9, 2554-2558.
- Kelso, S., Ganong, A., & Brown, T. (1986). Hebbian synapses in hippocampus. *Proceedings of the National Academy of sciences of the USA*, 83, 5326-5330.
- Rolls, E. T., & Treves, A. (1998). *Neural Networks and Brain Function*. Oxford: Oxford University Press.
- Sharkey, A.J.C., & Sharkey, N.E. (1992) Weak contextual constraints in text and word priming. *Journal of Memory and Language*, 31, 543-572.
- Stanton, P., & Sejnowski, T. (1989). Associative long-term depression in the hippocampus induced by hebbian covariance, *Nature*, 339(6221), 215-218.
- Martin, A., Wiggs, C., Ungerleider, L., & Haxby, J. (1996). Neural correlates of category-specific knowledge. *Nature*, 379(6566), 649-52.
- Pulvermüller, F. (1999). Words in the Brain's Language. *Behavioral and Brain Sciences*, 22, 253-279.
- Tettamanti, M., Buccino, G., Saccuman, M., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., & Cappa, S.P. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17, 273-281.