

Integrazione multisensoriale: la multimodalità e l'insegnamento a distanza

di **Gloria Galloni**

Introduzione

Per lungo tempo le diverse modalità sensoriali sono state indagate separatamente e considerate indipendenti, in quanto deputate all'analisi di forme di energia fisica specifiche (quali la luce per la vista o la pressione cutanea per il tatto o le onde sonore per l'udito e così via). Tradizionalmente si presumeva che l'integrazione tra le informazioni elaborate indipendentemente dalle diverse modalità avvenisse in aree neocorticali associative 'di alto livello'. La stessa caratteristica modulare che Fodor definisce 'incapsulamento informazionale' era stata teorizzata sulla base dell'assunto dell'isolamento delle modalità percettive, che egli chiamava trasduttori, il cui solo *output* veniva ad interagire con gli *output* delle altre modalità ai fini della generazione di una rappresentazione o dello svolgimento di un compito cognitivo di alto livello (Fodor, 1983).

Negli ultimi decenni l'attenzione si è invece sempre più spostata sulle interazioni tra modalità sensoriali, ora largamente indagate in numerosi settori di ricerca tra i quali quelli centrati sull'apprendimento, sull'attenzione, il riconoscimento di forme, oggetti o parole e via dicendo. In mol-

ti casi, infatti, l'elaborazione di informazioni provenienti da una modalità viene influenzata da informazioni disponibili in altre modalità, con effetti di facilitazione o di interferenza per i compiti cognitivi. Per lo più, avere a disposizione stimoli in più di una modalità sembra facilitarne l'elaborazione e il recupero (cf. Mastroberardino et al. 2008).

Inoltre, è importante considerare il fatto che nella quotidianità raramente ci troviamo di fronte ad oggetti le cui caratteristiche siano percepibili tramite una sola modalità sensoriale: per la maggior parte del tempo, siamo circondati da informazioni che ci giungono in più modalità percettive, le quali non si sommano banalmente assieme ma interagiscono fortemente tra di loro durante l'elaborazione del fenomeno percepito. Dunque, il pensare di doverle studiare nella loro interazione non è che la logica conseguenza del fatto che, se la percezione è un flusso continuo di informazioni in varie modalità sensoriali, i meccanismi cognitivi nel corso dell'evoluzione devono essersi adattati ad una percezione che è, per sua natura, multisensoriale. Senza alcun dubbio assumere una prospettiva di questo tipo permette di studiare i processi cognitivi in maniera più ecologica.

Studi 'storici' nel campo dell'integrazione multisensoriale

Da un punto di vista storico, è importante rilevare come già Müller (1838), nel teorizzare la cosiddetta 'legge dell'energia nervosa specifica', avesse citato l'effetto ventriloquo come un'illusione percettiva che presenterebbe un'ec-

cezione rispetto alla sua legge. Il ventriloquismo è un puro caso di interazione crossmodale, durante il quale la sincronizzazione tra il movimento della bocca di una bambola e il parlato fa sì che si percepisca quest'ultimo come proferto dalla bambola anziché dalla persona che la sostiene. Pensandoci un istante, è per lo stesso effetto che noi, vedendo un film doppiato, riusciamo ad attribuire le voci agli attori anche quando non perfettamente coerenti coi movimenti delle labbra e anche se il suono, invece che diffondersi direttamente tramite la bocca dei parlanti, fuoriesce da casse annesse al televisore.

Nonostante il fatto che, a partire da allora, vi siano stati sporadici lavori sulle 'illusioni multisensoriali', si dovrà attendere la seconda metà del Novecento per arrivare ad un movimento realmente sistematico di ricerche sulla percezione crossmodale. In tale contesto, un altro esempio di integrazione multisensoriale divenuto storico è il cosiddetto effetto McGurk (McGurk e MacDonald, 1976). Nell'effetto McGurk, la visione delle labbra del parlante influenza la percezione della sillaba udita: quando il parlante pronuncia un /ga/ ma lo stimolo uditivo che giunge è un /ba/, il soggetto tende a riconoscere la sillaba /da/ quale compromesso (McGurk e MacDonald, 1976). Secondo alcune interpretazioni (Ettlinger e Wilson, 1990), è l'attivazione sincrona la causa fondamentale del *binding*, ovvero della unificazione delle differenti sensazioni provenienti dallo stesso oggetto in un percetto unitario. Dunque, il riconoscimento del parlato avviene integrando le informazioni derivanti dalla modalità visiva e da quella acustica, così come sotto-

lineato anche nel titolo del loro lavoro ‘Sentire le labbra e vedere le voci’.

Cosa avviene nel nostro cervello

L'integrazione multisensoriale a livello neuronale avviene a livello di alcuni neuroni singoli (Meredith et al., 1986, 1987, 1996; Harting et al., 1997; Jiang et al., 2001), tra i quali molto studiati sono quelli del collicolo superiore nei mammiferi, sui quali convergono le attivazioni di neuroni unimodali appartenenti ad aree modali differenti. L'attivazione di questi neuroni può essere di tipo eccitatorio o inibitorio (Meredith, 2002; Calvert et al., 2000), ovvero più o meno della somma delle attivazioni dovute ai singoli stimoli a livello unisensoriale. Inoltre, l'attivazione può retroagire dalle zone di convergenza multisensoriale alle aree unisensoriali (Calvert et al., 1997; Vroomen e de Gelder, 2000; Macaluso et al., 2000; Driver e Spence, 2000), le quali sarebbero dunque influenzate ed influenzabili a livello crossmodale: ad esempio, la percezione del movimento delle labbra può attivare la corteccia uditiva primaria anche in assenza di stimoli uditivi (Calvert et al., 1997; Sams et al., 1991).

L'estensione dei meccanismi di integrazione multisensoriale dalle aree di convergenza multimodale alle aree primarie e dunque a tutti i livelli di elaborazione corticale ha fatto sì che venga proposto un ripensamento di tutti quei meccanismi cognitivi fino a non molto tempo fa modellizzati in una prospettiva essenzialmente unisensoriale.

Modalità dominanti

Un argomento rilevante qualora ci si interroghi sull'interazione tra le modalità sensoriali è quello della dominanza tra modalità, soprattutto in casi di discordanza tra informazioni provenienti da canali sensoriali differenti. Al riguardo vi sono teorie diverse. Secondo l'ipotesi dell'attenzione diretta, risulterà dominante la modalità sensoriale 'attesa' (Welch e Warren, 1980); tale ipotesi è poi stata modificata nell'ipotesi dell'appropriatezza, per cui il peso di una modalità nell'interazione con le altre dipende non dalle caratteristiche delle fonti d'informazione ma dal grado di adeguatezza o appropriatezza della modalità in quel dato compito (Welch, 1999). Per l'ipotesi dell'attendibilità, ad essere invece dominante in un compito cognitivo sarebbe la modalità le cui stimolazioni apportano, per quel compito, le informazioni maggiormente pregnanti (Schwartz et al., 1998). Infine, secondo l'ipotesi della discontinuità, risulta dominante la modalità in cui la stimolazione è maggiormente discontinua (Shams et al., 2002). Andersen e colleghi tuttavia offrono considerazioni teoriche a favore dell'idea che "all of these [hypotheses] should be considered as factors which contribute to the relative dominance of each modality and not as all-or-nothing conditions" (Andersen et al., 2004).

Multimodalità ed *e-learning*

Come precedentemente accennato, sembra dunque che avere una codifica in più modalità sensoriali possa facilitare la memorizzazione di informazioni. Ciò risulta cruciale per l'insegnamento, sicuramente in presenza ma ancora di

più per quanto concerne l'*e-learning*. Noi siamo organismi essenzialmente ed intrinsecamente multimodali, la nostra relazione col mondo e le nostre azioni in esso si basano sulla possibilità di cogliere i fenomeni integrativi della realtà.

Già nella metà del secolo scorso Edgar Dale (1946), nel suo famoso cono dell'esperienza, indagava le relazioni tra differenti tipi di materiali audiovisivi e apprendimento, sostenendo che siamo in grado di ricordare il 10% di ciò che leggiamo, il 20% di ciò che ascoltiamo, il 30% di ciò che vediamo, il 50% di ciò che sentiamo e vediamo, il 70% di ciò che diciamo e il 90% di ciò che diciamo e facciamo. Ciò sarebbe dovuto al fatto che il maggior coinvolgimento dell'individuo, grazie all'integrazione tra informazioni e al coinvolgimento motorio, facilita l'apprendimento.

Negli ultimi due decenni abbiamo assistito ad un notevole incremento del numero di piattaforme per *e-learning*, e si sta analizzando la differenza tra piattaforme classiche e piattaforme che utilizzino interfacce multimodali, alla ricerca di un vantaggio nell'apprendimento in caso di interfacce multimodali.

Secondo una ricerca appena pubblicata, per esempio, tra una piattaforma di interfaccia unicamente testuale e un'interfaccia multimodale (con video e avatar espressivi) vi è una differenza significativa. Infatti, l'uso di una piattaforma multimodale innalza il livello di usabilità, nel senso che gli utenti impiegano meno tempo a completare i compiti richiesti e li completano con meno errori. Inoltre, gli utenti risultano essere più soddisfatti dell'apprendimento e più motivati (Sallam e Rigas, 2010).

Conclusioni

L'e-learning è oggi una realtà in continua espansione, e si arricchisce sempre più di strumenti di apprendimento, di comunicazione, di verifica, strumenti che lo stanno rendendo un elemento davvero prezioso per lo sviluppo culturale della società. Da diversi anni sono tutor e docente in programmi di apprendimento a distanza, ed ho utilizzato piattaforme differenti con differenti tipi di studenti. Nella mia esperienza ho maturato l'idea che, nel promuovere la motivazione degli utenti, sia cruciale la capacità dell'interfaccia di coinvolgere i fruitori, di esprimere e far esprimere emozioni e vissuti personali. È evidente, anche nei risultati ottenuti dagli studenti e nel loro entusiasmo, la differenza tra un rapporto costruito su mail formali oppure su forum, e sulla presenza o assenza di *emoticons*, nonché sui differenti tipi di materiali utilizzati. La mia speranza è dunque che le interfacce si adattino sempre più alla peculiarità del nostro sistema cognitivo, un sistema dinamico e fortemente interattivo, che integra costantemente le informazioni per coglierne pienamente il significato e per sviluppare al meglio nuove conoscenze.

Bibliografia

- Andersen T.S., Tiippana K., Sams M. (2004) Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*. 21: 301-308.
- Calvert G.A., Campbell R., Brammer M.J. (2000) Evidence from functional magnetic resonance imaging of crossmodal binding

- in the human heteromodal cortex. *Current Biology*. 10:649–657.
- Calvert G.A., Bullmore E.T., Brammer M.J., Campbell R., Williams S.C., McGuire P.K., Woodruff P.W., Iversen S.D., David A.S. (1997) Activation of auditory cortex during silent lipreading. *Science*. 276:593-596.
- Dale E.E. (1946) *Audio-Visual Methods in Teaching*. New York: Dryden Press.
- Driver J., Spence C. (2000) Multisensory perception: Beyond modularity and convergence. *Current Biology*. 10:R731–R735.
- Ettlinger G., Wilson W.A. (1990) Cross-modal performance: behavioural processes, phylogenetic considerations and neural mechanisms. *Behavioral Brain Research*. 40:169-192.
- Fodor J.A. (1983) *The Modularity of Mind. An Essay on Faculty Psychology*. Cambridge. The MIT Press.
- Giard M.H., Peronnet F. (1999) Auditory-visual integration during multimodal object recognition in humans: a behavioural and electrophysiological study. *Journal of Cognitive Neurosciences*. 11:473-490.
- Harting J.K., Feig S., Van Lieshout D.P. (1997) Cortical somatosensory and trigeminal inputs to the cat superior colliculus: light and electron microscopic analyses. *The Journal of Comparative Neurology*. 388:313-326.
- Jiang H., Wallace M.T., Jiang H., Vaughan J.W., Stein B.E. (2001) Two cortical areas mediate multisensory integration in superior colliculus neurons. *Journal of Neurophysiology*. 85:506-522.
- Macaluso E., Frith C., Driver J. (2000) Modulation of human visual cortex by crossmodal spatial attention. *Science*. 289:1206-1208.
- Mastroberardino S., Santangelo V., Botta F., Marucci F.S., Olivetti Belardinelli M. (2008) How the bimodal format of presentation affects working memory: an overview. *Cognitive Processing*. 9:69–76.

- McGurk H., MacDonald J. (1976) Hearing lips and seeing voices. *Nature*. 264:746-748.
- Meredith M.A. (2002) On the neural basis for multisensory convergence: a brief overview. *Cognitive Brain Research*. 14:31-40.
- Meredith M.A., Stein B.E. (1996) Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology*. 75:1843-1857.
- Meredith M.A., Nemitz J.W., Stein B.E. (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience*. 7:3215-3229.
- Meredith M.A., Stein B.E. (1986) Visual, auditory, and somatosensory convergence on cells in the superior colliculus results in multisensory integration. *Journal of Neurophysiology*. 56:640-662.
- Müller J. (1838) *Elements of Physiology*.
- Olivetti Belardinelli M., Sestrieri C., Di Matteo R., Delogu F., Del Gratta C., Ferretti A., Caulo M., Tartaro A., Romani G. (2004) Audio-visual crossmodal interactions in environmental perception: an fMRI investigation. *Cognitive Processing*. 5:167-174.
- Pourtois G., de Gelder B. (2002) Semantic factors influence multisensory pairing: a transcranial magnetic stimulation study. *NeuroReport*. 13 (12):1567-1573.
- Sallam M., Rigas D. (2010) Comparing effectiveness and efficiency between multimodal and textual note-taking interfaces. *International Journal of Computers*. 2(4):70-77.
- Sams M., Aulanko T., Hamalainen H., Hari R., Lounesmaa O.V., Lu D.T., Simola J. (1991) Seeing speech: Visual information from lip movements modifies activity in the human auditory cortex. *Neuroscience Letters*. 127:141-145.
- Schwartz J.L., Robert-Ribes J., Escudier P. (1998) Ten years after Summerfield: a taxonomy of models for audio-visual fusion in speech perception. In R. Campbell, B. Dodd and D. Burnham (a cura di), *Hearing by Eye II: Advances in the Psychology of*

Speechreading and Auditory-visual Speech, Psychology Press, Hove, U.K., pp. 85-108.

Shams et al 2002

Vroomen J., de Gelder B. (2000) Sound enhances visual perception: cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*. 26(5):1583-1590.

Welch R.B. (1999) Meaning, attention and the unity assumption in the intersensory bias of spatial and temporal perceptions. In Ascherrsleben G. et al. (a cura di), *Cognitive contributions to the perception of spatial and temporal events*. Elsevier. Amsterdam. pp. 371-387.

Welch R.B., Warren D.H. (1980) immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*. 88:638-667.