

## **Percepire la scena acustica con l'impianto cocleare**

Francesco Pavani

1. Center for Mind/Brain Sciences (CIMEC), Università degli Studi di Trento, Rovereto, Italia
2. Dipartimento di Psicologia e Scienze Cognitive (DiPSCo), Università degli Studi di Trento, Rovereto, Italia
3. Centre de Recherche en Neurosciences de Lyon, Lione, Francia

Autore per la corrispondenza:

Francesco Pavani, Center for Mind/Brain Sciences (CIMEC), Università degli Studi di Trento, Corso Bettini 31, 38068, Rovereto, Italia

E-mail: francesco.pavani@unitn.it. Telefono: 0464 808674

Parole chiave:

ASCOLTO IN RUMORE, SORDITÀ, IMPIANTO COCLEARE, IPOACUSIA, MULTISENSORIALITÀ

Per la versione editoriale si veda: Pavani, F. (2021). Percepire la scena acustica con l'impianto cocleare. *Logopedia e Comunicazione*, 17, 105-127.

## PERCEPIRE LA SCENA ACUSTICA CON L'IMPIANTO COCLEARE

### Riassunto

L'impianto cocleare (IC) permette alle persone sorde un accesso all'informazione acustica, benché non restituisca un'esperienza di udito comparabile a quella degli udenti. Una delle difficoltà maggiori nell'ascoltare con l'IC è legata alle situazioni di rumore, anche per coloro che mostrano un buon recupero uditivo. Questa rassegna descrive i meccanismi cognitivi coinvolti quando ascoltiamo nel rumore e, in generale, quando percepiamo una scena acustica complessa. Comprendere la natura di questi meccanismi è necessario, sia per capire le ragioni delle difficoltà di analisi della scena acustica con l'IC, sia per trovare soluzioni concrete a questo problema. La rassegna discute due diverse tipologie di soluzioni, fra loro non esclusive: da un lato quelle basate sui contributi offerti dalla tecnologia, dall'altro quelle basate sulle potenzialità insite nel nostro sistema cognitivo, con particolare riferimento ai vantaggi dell'esperienza multisensoriale. L'uso dell'informazione multisensoriale ha la potenzialità di ampliare le strategie riabilitative volte a migliorare le capacità di ascolto in rumore per le persone con IC, sia durante l'esperienza percettiva, sia come strategia di apprendimento. Le implicazioni delle ricerche descritte in questa rassegna, per quanto delineate in relazione agli IC, riguardano ugualmente l'ascolto della scena acustica con protesi acustiche o in condizioni di ipoacusia in generale.

Conteggio parole: 198 (massimo 200).

**Abstract**

Although cochlear implants (CI) allow unprecedented access to auditory information for deaf individuals, they do not reinstate normal-hearing experience. One of the major difficulties is associated with hearing-in-noise situations, even for those CI users who are considered good performers. This review describes the cognitive mechanisms subtending hearing-in-noise and, in general, listening in complex auditory scenes. Outlining the nature of these mechanisms, is essential for understanding the reasons behind the difficulty of CI users when listening in noise, and it is the basis for implementing possible solutions. This review will discuss two approaches to the problem, that are not mutually exclusive: on the one hand, the contributions currently offered by technology; on the other hand, the potentials linked to intrinsic abilities of our cognitive system, with particular reference to the advantages related to multisensory processes. The use of multisensory inputs, both during perception and for the purpose of learning, have the potentials to expand the rehabilitation strategies aiming to strengthen listening-in-noise abilities of CI users. The works described in this review, although primarily focused on CI, have implications for listening in complex auditory scenes with hearing aids and, more generally, with any condition of hearing loss.

Conteggio parole: 195 (massimo 200).

## PERCEPIRE LA SCENA ACUSTICA CON L'IMPIANTO COCLEARE

L'impianto cocleare (IC) è uno dei maggiori successi biomedici degli ultimi decenni. Diversamente dalle protesi acustiche, che amplificano i suoni per sfruttare al meglio le funzionalità residue dell'orecchio naturale, gli impianti cocleari si sostituiscono all'orecchio naturale. Grazie a un piccolo numero di elettrodi inserito all'interno della coclea, l'IC stimola elettricamente le cellule del nervo acustico per riprodurre – almeno in parte – la complessità dei suoni presenti nell'ambiente. Un microfono esterno capta i suoni, un processore li scompone in bande di frequenza e li converte in segnali, che possono essere trasmessi al nervo acustico attraverso il filamento di elettrodi inserito nella coclea.

In condizioni tipiche, la conversione di un suono in impulsi nervosi per il nervo acustico è realizzata da circa 3.500 cellule specializzate all'interno di ciascun orecchio (le cellule ciliate interne). Ciascuna di queste cellule realizza una conversione di energia: allo stesso modo in cui una pala eolica converte l'energia meccanica del vento in energia elettrica, le cellule ciliate interne convertono l'energia meccanica del suono in segnali elettrici per il cervello. Tanto per le pale eoliche quanto per le cellule ciliate interne questo processo di conversione è chiamato *trasduzione*. Benché le cellule ciliate interne siano le sole a realizzare la trasduzione del segnale acustico in impulsi nervosi, il loro lavoro è coadiuvato da circa 15.000 cellule ciliate *esterne*, che non contribuiscono direttamente alla trasduzione, ma svolgono un ruolo fondamentale nell'amplificare naturalmente l'energia meccanica che costituisce il suono.

Ad oggi, gli impianti cocleari stimolano il nervo acustico con 24 elettrodi al massimo, dei quali solo circa un terzo è attivo in ogni singolo istante (quindi circa 8 per volta). Già solo questo dato rende evidente come l'IC debba essere pensato come un'approssimazione dell'orecchio naturale e non come un suo sostituto perfetto. Questo, tuttavia, è solo uno degli aspetti che distinguono l'IC dall'orecchio naturale. Prima di poter essere trasmessa al nervo acustico tramite gli elettrodi, l'informazione acustica deve essere semplificata attraverso una

sequenza di filtri per le diverse bande di frequenza, di rilevatori delle variazioni di energia nel suono e di compressioni dell'intervallo dinamico di queste energie. Solo alla fine di questo processo di rielaborazione e semplificazione del suono è possibile produrre un segnale adatto per ciascun elettrodo (Wilson, 2019).

Come è possibile allora che gli impianti cocleari abbiano avuto un tale successo e una tale diffusione, nonostante le molte limitazioni a cui abbiamo appena accennato? Come possono restituire un'esperienza uditiva in grado di permettere al bambino e all'adulto una buona o molto buona percezione del parlato? La ragione è sintetizzata bene nelle parole di uno dei padri dell'IC moderno, Blake S. Wilson: «il lavoro dei progettisti di impianti cocleari è stato quello di fornire una quantità sufficiente di informazione in un formato chiaro alla periferia, per permettere al cervello di fare *tutto il resto del lavoro* necessario alla percezione del parlato e degli altri suoni con adeguata accuratezza e fedeltà» (traduzione e corsivo mio, tratto da Wilson, 2019, p. 53).

Se da un lato è evidente che senza l'orecchio – sia esso naturale o artificiale – non sarebbe possibile alcuna esperienza acustica, è altrettanto vero che è il cervello che ci permette di dare un senso alle informazioni uditive che riceviamo. Per usare una metafora, pensate al vostro smartphone. Se lo privaste del microfono non potreste certo impartirgli alcun comando. Ma senza il raffinato software e hardware che i progettisti hanno inserito nel dispositivo per interpretare i vostri comandi vocali il solo microfono sarebbe del tutto inutile. Anche quando fosse perfettamente funzionante non permetterebbe alcuna interazione fra voi e lo smartphone. Lo scopo di questo breve articolo è, anzitutto, mostrare il compito sorprendente che devono svolgere il nostro cervello e il nostro sistema cognitivo quando analizzano l'ambiente sonoro attorno a noi (la cosiddetta *scena acustica*); in secondo luogo, aiutare a comprendere le ragioni per cui è difficile ascoltare nel rumore con l'IC; infine, delineare possibili strategie di intervento.

### La scena acustica

Il problema dell'analisi della scena acustica è schematizzato nella Figura 1. In molte occasioni della vita quotidiana l'ambiente acustico attorno a noi è ricco di suoni. Pensate, ad esempio, al contesto di un locale affollato, a una stazione ferroviaria, o anche a una cena fra amici con un po' di musica o con una televisione accesa. Questi suoni hanno solitamente caratteristiche acustiche differenti: alcuni sono suoni brevi (un colpo di tosse), altri durano molto più a lungo nel tempo (una conversazione); alcuni hanno modulazioni ritmiche regolari (il ticchettio dell'orologio a muro), altri ancora hanno modulazioni meno prevedibili (il ritmo delle parole del nostro interlocutore); alcuni sono principalmente composti da suoni gravi (una voce maschile profonda), altri sono caratterizzati da suoni acuti (il miagolio del gatto) e così via.

Nella Figura 1, a sinistra, è mostrata un'ipotetica scena acustica composta da tre suoni: la sirena di un'auto della polizia, una conversazione fra un uomo e una donna, e l'abbaiare di un cane. Anche se nella realtà questi suoni provengono da fonti sonore distinte, quando raggiungono l'orecchio questa separazione naturale si perde. Indipendentemente dalla loro natura (suono meccanico, voce umana, verso animale) e dalla loro posizione nello spazio (molto vicino o più lontano) *tutti* i suoni presenti nella scena in ogni dato istante raggiungono il sistema acustico della persona udente attraverso due sole porte: il timpano destro e il timpano sinistro. Ciò significa che nella coclea, dove l'energia acustica diventa impulso nervoso per il nervo acustico, tutti i suoni presenti finiscono col sommarsi l'uno all'altro. Quella che è una scena formata da eventi acustici ben distinti, diventa una matassa intricata che deve essere dipanata. Se siamo in grado di percepire come distinti i molteplici eventi acustici che sono spesso presenti nella scena, è perché sull'informazione che ha raggiunto la coclea interviene il nostro *sistema cognitivo*, ovvero l'insieme delle capacità mentali che il nostro cervello è in grado di esprimere per permetterci un'interazione con l'ambiente fisico e sociale.

< Inserire la Figura 1 circa qui >

Si ritiene che il sistema cognitivo debba svolgere almeno tre operazioni per ricostruire la scena acustica originale, partendo dai segnali che raggiungono la coclea (Shinn-Cunningham, 2017). La prima è ricostruire i diversi *oggetti acustici* che erano presenti nell'ambiente, ovvero riuscire ad individuare nel segnale complessivo disponibile alla coclea (in rosso nella Figura 1) i diversi eventi acustici che l'hanno determinato (le tracce azzurre, verdi e gialle). La seconda è seguirli nel tempo, nel loro svolgersi istante dopo istante, come *flussi acustici*. A ben vedere, infatti, pochi oggetti acustici sono brevi e isolati. Il più delle volte ascoltiamo oggetti acustici che durano alcuni secondi (un'auto che passa in strada), alcuni minuti (una conversazione), alcune ore (il rumore del motore dell'aereo durante un viaggio); individuare gli oggetti acustici non è altro che il primo passo necessario per seguire poi il loro svilupparsi nel tempo. Infine, è necessario che i flussi acustici che abbiamo individuato siano agganciati dalla nostra *attenzione selettiva*. Questa fondamentale capacità cognitiva è lo strumento attraverso il quale il nostro cervello seleziona le informazioni che, di volta in volta, sono più rilevanti per il nostro comportamento, cercando di ignorare invece tutto ciò che non ci interessa. Durante un volo aereo, ad esempio, selezioniamo la conversazione del vicino sforzandoci di ignorare il rumore costante dei motori.

I meccanismi che permettono al cervello di identificare prima gli oggetti acustici e poi i flussi acustici sono studiati da tre decenni (Bregman, 1990). Per cercare di dipanare la matassa di suoni che si è venuta a creare a livello della coclea, il sistema cognitivo sfrutta anzitutto il fatto che oggetti acustici diversi hanno solitamente caratteristiche acustiche differenti: se nella scena acustica stanno parlando un uomo e un bambino, le due voci saranno caratterizzate da frequenze differenti, in quanto la voce dell'uomo avrà frequenze più gravi



rispetto alla voce del bambino. È possibile iniziare il processo di separazione dei due segnali sonori sfruttando queste differenze di frequenza. Evidentemente, tanto più sono simili fra loro i suoni presenti nell'ambiente, tanto più arduo è l'utilizzo di questo meccanismo di separazione. In secondo luogo, gli oggetti acustici possono essere separati gli uni dagli altri sulla base delle loro posizioni. Poiché nella vita reale due corpi distinti non occupano la stessa posizione nell'ambiente (il principio fisico della impenetrabilità dei corpi), è piuttosto improbabile che le fonti acustiche attorno a noi provengano dalla medesima posizione nello spazio. Dunque, nella misura in cui è possibile individuare la posizione dei suoni, sarà possibile separare fra loro gli oggetti acustici. Questo non è sempre facile, perché alcuni suoni (es., un ronzio di fondo, un sibilo) possono essere molto difficili da localizzare nell'ambiente. Infine, nel separare gli oggetti acustici il cervello utilizza una sua capacità fondamentale: la predizione.

Molti ricercatori sono ora concordi nel descrivere il cervello come macchina che genera continuamente predizioni (per una descrizione in italiano di questo principio generale del funzionamento del cervello e del sistema cognitivo, si veda Dehaene, 2019). In ogni singolo istante, il nostro cervello cerca di dare un senso alla realtà (perceptiva, sociale, concettuale) formulando predizioni sulla base delle esperienze precedenti. Ad esempio, se ascoltate la frase “Sono andato dal dentista perché avevo male ai ...”, il vostro cervello anticiperà la parola “denti” ancora prima che venga pronunciata dal vostro interlocutore. Se quella parola è all'interno della matassa dei suoni, sarà quindi più facile individuarla, proprio perché attesa. È questa la ragione per cui è molto più facile ascoltare nel rumore una conversazione nella nostra lingua madre, piuttosto che una conversazione in una lingua appresa più tardivamente, persino quando siamo perfettamente fluenti in questa seconda lingua (Sharenborg e van Os, 2019). Della lingua madre abbiamo modelli mentali estremamente precisi delle differenze fonologiche (es., un confine preciso fra la ‘b’ e la ‘v’), delle probabilità

lessicali (es., è molto più probabile ascoltare in una conversazione la parola 'casa' rispetto alla parola 'rasa') e delle strutture morfo-sintattiche (es., "Gianni e il nonno andavano al parco"). Quanto più siamo in grado di generare predizioni accurate circa i suoni che ascoltiamo, tanto più facile sarà dipanare i suoni nella matassa dei rumori.

### **Ascoltare nel rumore con l'impianto cocleare**

Come anticipato sopra, l'IC rende possibile l'ascolto anche per una persona con sordità severa o profonda. Per quanto si tratti di un'esperienza acustica molto diversa rispetto a quella di un udente, e diversa anche rispetto a quella di una persona che porta una protesi acustica, in molti casi l'IC consente un accesso al linguaggio orale assai migliore rispetto a quello veicolato dalla sola lettura labiale (vedi nota 1). In questo senso, l'IC è un dispositivo senza precedenti nel favorire l'accesso alla lingua parlata. Non affronterò qui il tema della variabilità nell'esito dell'impianto fra i diversi individui, che è un risultato molto noto e già affrontato in maniera chiara in altri contributi (per un contributo recente in italiano si veda Rinaldi, Caselli & Pavani, 2017). Piuttosto, mi concentrerò sulla grande differenza che esiste fra l'uso dell'impianto in condizione di silenzio e l'uso dell'impianto in condizione di rumore, quando, cioè, la scena acustica è complessa e occorre mettere in atto i meccanismi di analisi descritti nel paragrafo precedente.

In estrema sintesi, anche la persona che con l'IC ha un recupero uditivo buono o ottimo ha solitamente una difficoltà notevole quando deve ascoltare nel rumore. Questo concetto è riassunto nel grafico della Figura 2, che mostra la percentuale di frasi correttamente riportate da persone adulte con IC in tre diverse condizioni di ascolto: il silenzio, una situazione più facile di ascolto in rumore (+10 dB SNR) e una situazione più difficile di ascolto in rumore (+5 dB SNR; per il significato dei due valori di SNR si veda la nota 2). Si tratta di un compito di comprensione di frasi ascoltate; il rumore, quando presente, è un indistinto sovrapporsi di

molte voci simultanee (quello che viene definito solitamente il rumore da brusio multi-parlante o *multi-talker babble noise*). Per gli ascoltatori normo-udenti (NU), rappresentati dalle colonne bianche nel grafico, le tre condizioni sono ben poco differenti: si osserva un lieve decremento della prestazione nella situazione più difficile (+5 dB SNR), ma la prestazione rimane comunque oltre il 95% di risposte corrette. La situazione è però molto diversa se consideriamo le persone con IC (barre grigie): la loro prestazione decade dalla situazione di ascolto in silenzio (82% di risposte corrette) alla situazione di ascolto +10 dB SNR (54%) alla situazione più difficile di +5 dB SNR (36%). Si può dedurre che anche una differenza di rumore che non intacca la prestazione di persone udenti, può rendere molto difficile l'ascolto per le persone con impianto. Ho scelto di illustrare questo concetto usando i dati di Sphar e colleghi (2007; ridescritto in Dorman e Gifford, 2017), ma si tratta di un'osservazione molto comune nella letteratura sugli impianti (es., Caldwell & Nittrouer, 2013; Srinivasana et al., 2013) e ben nota anche nella pratica clinica.

C'è un aspetto di questo risultato che è importante sottolineare: tutti i partecipanti con IC dello studio citato ( $n = 65$ ) avevano almeno un anno di esperienza di ascolto con l'impianto ed erano stati scelti per la loro buona o molto buona capacità di comprensione di parole singole nel silenzio. In altri termini, rispetto al noto problema della variabilità *inter*-individuale nell'esito dell'impianto, questi partecipanti appartengono tutti al gruppo di coloro che hanno una buona prestazione con l'IC (i cosiddetti *good-performers*). Ciononostante, l'aggiunta della condizione di rumore li mette in difficoltà, creando una variabilità nella loro prestazione che potremmo definire *intra*-individuale. Ovvero, lo stesso individuo può beneficiare più o meno dell'impianto in misura diversa in funzione delle condizioni di ascolto nelle quali si viene a trovare, anche quando nel silenzio è un *good-performer*.

Ma quanto è probabile che una persona con impianto si venga a trovare in condizioni di ascolto difficoltose, perché rumorose? La risposta è nota perché, diversamente da quanto accade per gli udenti, è molto facile tenere traccia delle situazioni di ascolto che una persona con impianto ha incontrato durante la sua giornata: basta registrare il segnale che arriva al microfono dell'impianto nell'arco delle 24 ore. Una recente ricerca italiana (Cristofari et al., 2017) ha registrato le scene acustiche in cui si erano venuti a trovare 1.366 portatori di IC di diverse età, dall'infanzia (0-2 anni) all'età matura (>75 anni). Ciò che è emerso è che il gruppo pre-scolare e scolare (0-18 anni) era quello maggiormente esposto all'ascolto in rumore. Soprattutto nella fascia d'età fra i 6 e i 18 anni (periodo scolare), i bambini e i ragazzi passavano in media 4-5 ore al giorno in questo contesto acustico difficoltoso. Al contrario, il gruppo di soggetti più adulti (>19 anni) passava più tempo in contesti silenziosi. Questo dato certamente non sorprende, ma rivela che proprio coloro che hanno maggiormente bisogno di estrarre informazioni dalla scena acustica con l'IC – i bambini che stanno apprendendo il linguaggio, e bambini e ragazzi che apprendono contenuti comunicati oralmente a scuola – sono i più esposti alla situazione che è più difficile gestire con l'IC.

< Inserire la Figura 2 circa qui >

### **Perché è difficile ascoltare in rumore con l'impianto cocleare?**

Per comprendere come mai sia così problematico l'ascolto in rumore con l'IC, dobbiamo tornare ai meccanismi cognitivi che rendono possibile l'ascolto nella scena acustica, ovvero: la capacità di individuare gli oggetti acustici, la capacità di seguire i flussi acustici nel tempo e la capacità di dirigere le proprie risorse attenzionali verso il flusso acustico che interessa.

Come anticipato sopra, tanto più gli oggetti e i flussi acustici sono diversi nelle frequenze dei suoni che contengono, tanto più facile sarà individuarli e seguirli nel tempo. Dunque, ascoltare una voce femminile in una scena acustica con altri parlanti sarà più facile se le altre voci sono maschili, tipicamente caratterizzate da frequenze più gravi di quelle femminili. Tra le caratteristiche ben note dell'IC c'è, tuttavia, un sostanziale impoverimento delle frequenze percepibili, che rende difficile discriminare le differenze qualitative fra i diversi suoni (Moore e Shannon, 2009). In particolare, sappiamo che persone con IC possono avere difficoltà nel riconoscere se una voce è maschile o femminile (Clearly e Pisoni, 2002; Fu et al., 2004), o anche nel discriminare fra voci umane e altri suoni. Ad esempio, uno studio condotto in Francia (Massida et al., 2011) ha chiesto a 30 persone con sordità post-verbale e IC di ascoltare dei suoni di brevissima durata (0.5 secondi) per decidere se fossero campioni di voce umana (es., parlato, risate, colpi di tosse) o piuttosto suoni ambientali di altra natura (es., rumori di auto, campane, rumori di acqua). Quanto più avevano esperienza con l'IC, tanto più i pazienti miglioravano nella loro comprensione di parole, raggiungendo una prestazione oltre il 70% dopo 18 mesi dall'accensione dell'impianto. Tuttavia, la loro capacità di discriminare fra stimoli voce e stimoli non-voce era molto più deficitaria: erano in grado di fare questa distinzione solo dopo 18 mesi di utilizzo dell'impianto e la loro prestazione era comunque molto inferiore a quella misurata negli udenti.

La prima ragione per cui è difficile dipanare la matassa di una scena acustica con l'IC risiede quindi nel fatto che l'analisi delle frequenze del suono è molto meno efficiente nelle persone con IC rispetto alle persone udenti. Ciò che è importante notare è che anche una ridotta efficienza costituisce un problema in questo caso: anche laddove una discriminazione è possibile (nell'esempio sopra, quella fra voci e non voci, che abbiamo visto emergere attorno ai 18 mesi dall'impianto), il fatto che sia più difficoltosa rende l'analisi della scena acustica uno sforzo continuo. A parità di prestazione, le singole persone possono fare uno sforzo molto

diverso nell'ascoltare la scena acustica. Una minore efficienza nella separazione fra gli oggetti acustici può quindi spiegare anche per quale ragione le persone con IC si stanchino prima, e quindi smettano prima di ascoltare rispetto alle persone udenti, quando la situazione di ascolto è difficile.

Una seconda importante ragione che rende difficile l'ascolto in rumore con l'IC ha a che fare con l'analisi dello spazio acustico (Litovski, 2017). Se chiudiamo gli occhi e ascoltiamo lo spazio acustico attorno a noi, l'esperienza che abbiamo come udenti è quella di uno spazio tridimensionale dei suoni. In questo momento alcuni suoni sono lontani, come il rumore delle automobili in strada; altri provengono da un po' più vicino, come l'occasionale colpo di tosse della persona nella stanza a fianco; altri ancora possono essere molto prossimi, come il ronzio del mio computer. Inoltre, alcuni sono a destra, altri a sinistra, altri in alto, e così via. Questa spazialità insita nella scena acustica è molto utile per distinguere gli oggetti e i flussi sonori, è inoltre necessaria per orientare l'attenzione verso la posizione spaziale dove sta accadendo ciò che ci interessa seguire (Pavani et al., 2017).

Purtroppo anche la capacità di analisi dello spazio acustico è alquanto compromessa nelle persone con IC (Litovski, 2017; Pavani et al., 2017). Nel localizzare i suoni nello spazio attorno a noi il cervello sfrutta due informazioni principali. Le prime sono le differenze di tempo e di intensità fra i suoni che raggiungono le due orecchie – i cosiddetti *indizi acustici binaurali*. I suoni a destra arriveranno prima all'orecchio destro rispetto al sinistro, perché quest'ultimo è più lontano dalla sorgente del suono. Al contempo, i suoni a destra saranno più intensi all'orecchio destro rispetto al sinistro, perché fra le due orecchie c'è la testa a fare da schermo acustico. Queste differenze del suono alle due orecchie svolgono un ruolo fondamentale nella nostra capacità di individuare la direzione del suono sul piano orizzontale, ovvero quanto un suono proviene da destra o da sinistra. Le seconde sono le variazioni nel suono che dipendono dal modo in cui l'onda acustica interagisce con il padiglione auricolare,

l'orecchio esterno. Questi indizi acustici, detti monoaurali, sono fondamentali per stabilire, ad esempio, la posizione verticale del suono o la sua provenienza dallo spazio anteriore o posteriore.

Se una persona utilizza un solo IC nessuno dei due indizi binaurali che ho appena descritto sarà pienamente disponibile. Gli indizi binaurali saranno assenti (nel caso di sordità completa all'orecchio non impiantato) o comunque alterati (in caso di protesi acustica indossata sull'orecchio non impiantato, qualora sia presente un residuo uditivo sufficiente). Gli indizi monoaurali saranno anch'essi ampiamente compromessi per effetto del fatto che l'IC non permette un'analisi fine delle frequenze contenute nel suono. Di conseguenza, la sua percezione della provenienza del suono sarà molto incerta e frequentemente il paziente attribuirà il suono al lato dal quale porta l'impianto, perché inevitabilmente ogni suono arriverà sempre prima e sarà sempre più intenso a quell'orecchio rispetto all'orecchio sordo. La situazione è in parte migliore per coloro che utilizzano due impianti cocleari, perché in questo caso la binauralità dell'ascolto è parzialmente ristabilita. In effetti, le persone con impianto bilaterale sono maggiormente in grado di localizzare i suoni e, in genere, commettono meno errori sul lato del suono (Litovski, 2017). Tuttavia, i dati della ricerca mostrano che l'incertezza orizzontale può ancora essere piuttosto ampia e le persone impiantate non sono in grado di stabilire facilmente se un suono si trovi davanti o dietro, in alto o in basso.

Infine, c'è un'ulteriore ragione per cui può essere difficile per una persona con IC ascoltare in rumore. Una ragione che questa volta non ha nulla a che fare con le limitazioni tecniche dell'impianto, ma dipende piuttosto da aspetti più centrali del funzionamento cognitivo: le capacità predittive del cervello circa i suoni che la persona potrà ascoltare nel rumore.

Nelle persone con IC le capacità trasversali necessarie a sostenere i meccanismi di predizione possono essere diverse rispetto a quelle degli udenti. L'esempio più evidente è

probabilmente legato alle predizioni che nascono da una competenza linguistica pienamente matura. Un dato noto, confermato anche da una recente meta-analisi (ovvero una ri-analisi di risultati ottenuti attraverso più ricerche; Lund, 2015), è che i bambini con IC hanno capacità di vocabolario minori rispetto ai bambini udenti. Una ricerca recente (Amenta et al., in revisione) ha inoltre mostrato che, anche all'interno dell'insieme delle parole note, i ragazzi con sordità preverbale e IC sono più suscettibili agli effetti di frequenza della parola scritta. Pur mostrando di conoscere le parole che venivano presentate, i ragazzi con impianto erano rapidi tanto quanto i coetanei udenti nell'elaborare parole frequenti del lessico (es., "casa"), ma impiegavano più tempo degli udenti per le parole a più bassa frequenza (es., "altero"). Avere un vocabolario più o meno ampio, o più o meno rapidamente accessibile, può plausibilmente rendere i meccanismi predittivi utili all'ascolto in rumore più o meno efficienti.

Un secondo esempio di natura linguistica deriva da uno studio elettroencefalografico condotto in Germania su persone con sordità tardiva e IC. Hahne e colleghi (2012) hanno dimostrato che le persone con impianto possono avere delle difficoltà a elaborare anomalie sintattiche in frasi udite. I partecipanti di questo studio ascoltavano frasi corrette, come "Der Tisch wurde gedeckt" (La tavola fu preparata), o frasi che contenevano una violazione della struttura argomentale del verbo, come "Der Hund wurde gebellt" (Il cane fu abbaiato). Registrando attraverso l'elettroencefalogramma (EEG) le risposte cerebrali evocate da queste due tipologie di frasi si osserva un dato caratteristico: l'onda cerebrale innescata dall'ascolto della frase, chiamata P600, è maggiore per le frasi sbagliate rispetto alle frasi corrette. Si tratta di una risposta automatica del cervello all'anomalia contenuta nella frase sbagliata, che prende la forma di una maggiore positività elettrica (P), che si verifica 600 millisecondi circa dopo l'inizio della frase udita (600). Questa risposta alle frasi sbagliate è considerata un indicatore del fatto che il cervello si è accorto dell'anomalia sintattica e sta cercando di risolverla. Usando questo approccio, gli autori dello studio hanno potuto mostrare che le persone con IC sono



meno sensibili degli udenti a queste violazioni sintattiche ascoltate nelle frasi. Come se l'ascolto alterato tramite l'IC rendesse, giorno dopo giorno, la percezione di alcuni elementi sintattico-morfologici (es., morfemi liberi, morfemi legati) più incerta, e di conseguenza potesse modificare le aspettative che le persone formulano nell'ascoltare le frasi. Nuovamente, un meccanismo predittivo alterato che potrebbe influenzare i meccanismi di ascolto in rumore.

### **Migliorare l'ascolto in rumore attraverso la tecnologia**

Descrivere e comprendere i principali motivi alla base delle difficoltà di ascolto in rumore per le persone con IC non è solo un esercizio teorico: è il passo necessario se vogliamo provare ad agire su questo problema cruciale, con l'obiettivo di migliorare l'ascolto in rumore per chi ha problemi di udito. Negli ultimi anni, molta ricerca empirica sta perseguendo questo scopo, attraverso uno sforzo multidisciplinare che da un lato, mira a sviluppare nuove tecnologie per migliorare l'ascolto in rumore e, dall'altro, si interroga su come potenziare le capacità del sistema cognitivo essenziali per dipanare la matassa di una scena acustica. In questa sezione descriverò anzitutto alcune delle soluzioni emerse dall'approccio tecnologico al problema. Nella sezione successiva descriverò invece le soluzioni basate sul potenziamento del sistema cognitivo. È importante ricordare fin d'ora che si tratta di due soluzioni complementari, non alternative.

La prima soluzione tecnologica al problema dell'ascolto in rumore con l'IC deriva da quanto è già stato sperimentato con le protesi acustiche: l'utilizzo di microfoni direzionali, in grado di captare il segnale proveniente da una posizione dello spazio, può efficacemente ridurre il rumore che proviene dal resto dell'ambiente. Ad esempio, un microfono che amplifichi ciò che proviene da davanti e attenui ciò che proviene da dietro o dai lati. Confrontata con una strategia *omnidirezionale* (nella quale il microfono capta in maniera uguale i suoni che provengono da tutte le direzioni), la strategia *direzionale* porta ad un

miglioramento della prestazione durante l'ascolto in contesti di rumore. Nella Figura 3A, questo tipo di strategia direzionale è detta 'adattiva', nel senso che si adatta al contesto acustico, dando priorità ad alcune posizioni quando viene rilevato un incremento del rumore. Confrontate fra loro le barre del grafico: le barre grigie o a righe, che rappresentano la prestazione in presenza di rumore, sono più basse rispetto a quelle nere dell'ascolto nel silenzio, indipendentemente dal fatto che il rumore aggiunto al parlato sia la simulazione di un ristorante o la simulazione di un *cocktail party* (nota 3). Tuttavia, le barre a righe (microfono in modalità *adattiva*) rivelano una prestazione migliore rispetto alle barre grigie (microfono in modalità *omnidirezionale*).

Si tratta di una soluzione efficace, come potranno confermare le persone che l'hanno sperimentata sulle loro protesi o i loro IC. Ma è importante essere consapevoli di un limite insito in questa soluzione: è il microfono che sceglie a cosa dare priorità nel rumore, non il vostro sistema cognitivo. In altre parole, questa soluzione funziona bene fin tanto che voi volete prestare attenzione a ciò che il microfono sta amplificando (es., la persona che parla davanti a voi). Ma se la vostra attenzione si vuole spostare altrove – magari senza darlo a vedere, come quando ad una cena o durante una festa volete prestare attenzione al vostro vicino senza risultare scortesi con chi vi sta davanti – questa soluzione tecnologica mostra immediatamente i suoi limiti. Considerate ad esempio gli effetti del microfono direzionale in un'aula scolastica. Durante la spiegazione da parte del professore, la direzionalità del microfono permetterà di ottimizzare l'ascolto della lezione. Tuttavia, nel momento in cui la classe diventa più interattiva, con domande e commenti da parte dei compagni, la stessa direzionalità diventa un ostacolo alla partecipazione dell'alunno sordo a questi scambi. Infine, immaginate che il microfono stia amplificando in maniera 'adattiva' ciò che accade davanti a voi, senza dare priorità acustica al motorino che proviene da dietro e del quale sarebbe meglio

accorgersi. In questo caso la risposta del microfono potrebbe essere ben poco 'adattiva' per voi stessi!

< Inserire la Figura 3 circa qui >

Una seconda soluzione tecnologica molto importante è quella che separa il flusso acustico del parlato all'interno del rumore, indipendentemente dalla sua posizione nello spazio. In questa soluzione troviamo il tentativo di imitare ciò che, in condizioni tipiche, il nostro cervello riesce a fare spontaneamente durante l'ascolto della scena acustica: identificare gli oggetti acustici presenti nella scena e seguirli nel tempo. È una soluzione efficace e in rapidissima evoluzione, perché implementata anche in Google Assistant, Amazon Alexa o Apple Siri, per permettere a questi assistenti virtuali di riconoscere comandi verbali pronunciati nel rumore. Chi ha già provato a interagire con questi assistenti virtuali sa che, ormai, possono eseguire i nostri comandi vocali (es., "Alexa accendi le luci", "Siri chiama Agata") anche in presenza di altri suoni interferenti. Questo non è forse sorprendente quando pronunciamo le nostre istruzioni direttamente al microfono del cellulare, ma gli assistenti virtuali possono funzionare bene anche quando diamo i nostri ordini a diversi metri di distanza.

Un gruppo di ricerca statunitense, ad esempio, ha cominciato a testare questo genere di algoritmi nelle persone ipoacusiche per misurare il guadagno che possono portare nell'ascolto in rumore. Come si vede dall'esempio in Figura 3B, per tutti i partecipanti ipoacusici (HI: Hearing Impaired) testati in una di queste ricerche (Healy et al., 2013), la prestazione durante l'ascolto di frasi in un brusio indistinto di altre voci migliorava grazie all'uso di questa tecnologia (ogni freccia indica un partecipante: il punto alla base indica la prestazione senza il sistema di separazione fra parlato e rumore, la punta della freccia indica

invece la prestazione quando si sfrutta la soluzione tecnologica; si noti come tutte le frecce puntano verso prestazioni più alte).

Anche questa seconda soluzione, per quanto efficace, ha tuttavia alcuni punti deboli, quantomeno nello scenario tecnologico attuale. Uno di questi consiste nel fatto che per estrarre il segnale dal rumore bisogna rinunciare a molti aspetti qualitativi del parlato. Il suono che risulta, dopo operazioni di filtraggio che separano il segnale dal rumore, è impoverito di molte delle caratteristiche naturali della voce: la prosodia, le emozioni nel parlato, o anche l'identità della voce, per effetto del filtraggio vengono fortemente attenuate, quando non cancellate del tutto. Se questo può essere indifferente per un assistente virtuale, per il quale l'unica informazione che conta è il contenuto dell'ordine che riceve, di certo non è indifferente per un essere umano.

### **Migliorare l'ascolto in rumore sfruttando il sistema cognitivo**

Un diverso approccio al problema dell'ascolto in rumore per le persone con IC consiste nello sfruttare le risorse del nostro sistema cognitivo, che, vale sempre la pena di ricordarlo, è l'interfaccia di cui l'evoluzione ci ha dotato per interagire con il mondo fisico e sociale. Come già anticipato, nel separare la scena acustica in oggetti e flussi acustici, il sistema cognitivo sfrutta il fatto che le diverse fonti sonore hanno una loro coerenza interna (es., la voce di un determinato individuo manterrà le medesime frequenze nell'arco del tempo), hanno posizioni spaziali distinte e, non ultimo, che è possibile esercitare un certo grado di predizione su ciò che ascoltiamo. Possiamo aiutare le persone con IC a sfruttare al meglio queste abilità insite nel loro sistema cognitivo?

Una prima risposta affermativa a questa domanda la troviamo nel vantaggio che le persone con IC ricavano dal combinare informazioni acustiche e visive nell'ascolto del parlato. Esattamente come accade per gli udenti, anche per le persone con IC la percentuale di

frasi riportate correttamente è maggiore quando possono ascoltare e vedere il labiale, rispetto a quando devono svolgere lo stesso compito attraverso la sola modalità acustica (Figura 3C; Dorman et al., 2016). Il vantaggio dato dall'esperienza visuo-acustica del parlato nelle persone udenti è stato descritto già negli anni '50 del secolo scorso. Nel 1954, Sumbly e Pollack mostrarono che, rispetto a una condizione di ascolto basata sulle sole informazioni acustiche, l'aggiunta del labiale migliorava quasi del 40% il riconoscimento delle parole udite in una condizione di ascolto rumorosa. Questo tipo di dato è stato confermato e esteso da molte ricerche successive (per una rassegna si veda Navarra et al., 2012).

Oggi sappiamo che le persone con IC sfruttano questi meccanismi di integrazione audio-visiva nella percezione del parlato addirittura in misura maggiore rispetto agli udenti. La prima evidenza di questa capacità risale a oltre un decennio fa (Rouger et al., 2007), e è emersa da una ricerca che aveva esaminato le capacità di comprensione di parole in 97 adulti sordi post-linguali, prima e dopo l'IC. Ai partecipanti veniva chiesto di ripetere parole solo udite, parole solo viste dal labiale di un attore, o parole che erano al contempo udite e viste. La prestazione solo uditiva aumentava nei primi mesi dopo l'impianto, raggiungendo un plateau attorno all'80% di parole correttamente identificate già nella seconda metà del primo anno post-impianto. Questo era un risultato atteso: nelle persone con sordità post-linguale, l'IC apporta un significativo beneficio rispetto alle protesi acustiche, pur non raggiungendo il livello dell'esperienza uditiva naturale.

Più interessante è ciò che accadeva per le capacità di sola lettura labiale. Prima dell'impianto, le persone sorde identificavano correttamente il 30-40% circa delle parole (una prestazione superiore a quella degli udenti, che raggiungevano appena il 10%). Dopo più di 8 anni dall'IC questa abilità visiva rimaneva sostanzialmente invariata, a indicare che l'esercizio alla lettura labiale continuava a verificarsi. Significa che le persone con IC si affidano completamente alla lettura labiale per comprendere quanto ascoltano? Certamente no.

L'interpretazione corretta di questo dato va cercata in ciò che accadeva nella prestazione audio-visiva. La prestazione nella condizione audio-visiva arrivava a un plateau del 90% già due mesi dopo l'accensione dell'impianto, dunque era molto più elevata, sia rispetto alla sola abilità di lettura labiale, sia rispetto all'abilità di comprensione di parole solo udite. Questo risultato indica che la preservata abilità di lettura labiale serviva durante l'esperienza audio-visiva per *sostenere* una migliore comprensione del messaggio uditivo percepito tramite l'IC (nota 4).

In sintesi, un primo modo efficace per aiutare le persone con impianto ad ascoltare meglio in rumore è permettere loro di sfruttare al massimo la multisensorialità dell'esperienza linguistica. Questa multisensorialità si realizza in maniera spontanea quando guardiamo un parlante in volto: la combinazione fra le informazioni dalle labbra e dal volto aiuta a disambiguare meglio ciò che viene percepito attraverso l'udito. Alcuni ricercatori stanno però cominciando a estendere lo stesso principio anche a altre combinazioni multisensoriali. Ad esempio, una ricerca recente (Cieśła et al., 2019) ha testato il contributo di indizi vibro-tattili alle dita nel migliorare l'ascolto di frasi pronunciate nel rumore. Si tratta di una ricerca pilota, condotta ancora su persone udenti che ascoltavano i suoni *come se* fossero trasmessi dall'IC (nota 5). Ciononostante, la ricerca ha dimostrato che il semplice fatto di stimolare le dita con la frequenza fondamentale contenuta in ciascuna frase migliorava la prestazione nell'ascolto di frasi.

Una seconda strategia cognitiva per migliorare l'ascolto in rumore in caso di materiale linguistico è quella che sfrutta segnali visivi non legati al volto per risolvere l'ambiguità (o la carenza) presente nell'informazione acustica. Questo è il caso, ad esempio, di ricerche che hanno valutato il contributo di singoli gesti o veri e propri segni nel supportare l'ascolto in rumore per le persone con ipoacusia. Un primo esempio deriva da una ricerca che ha valutato il contributo del cosiddetto *cued speech* in aggiunta all'informazione acustica e del labiale

(Bayard et al., 2019). Proposto a partire dalla fine degli anni '60 (Cornett, 1967), il *cued speech* prevede gesti della mano realizzati in prossimità del volto per rendere visibili aspetti della fonologia del linguaggio orale, altrimenti non visibili nel labiale. In questo sistema di supporto alla comprensione del labiale ogni sillaba pronunciata è accompagnata da un gesto chiamato 'segnale manuale' (in inglese, *manual cue* da cui il termine *cued speech*). Bayard e colleghi (2019) hanno chiesto a un gruppo di 20 adolescenti sordi, tutti portatori di IC, a eccezione di tre con protesi acustiche, di ripetere delle frasi ascoltate nel rumore in tre diverse condizioni sperimentali: una prima condizione solo acustica, una seconda condizione nella quale erano disponibili le informazioni dal labiale, e una terza condizione nella quale oltre al labiale erano disponibili le informazioni di *cued speech*. Tutti i partecipanti sordi avevano una pregressa esperienza con il *cued speech* (lo avevano appreso attorno ai 4 anni e mezzo). I risultati di questo studio hanno rivelato che l'ascolto in rumore incrementava dal 15% di frasi correttamente identificate nella condizione solo acustica, al 55% della condizione in cui erano visibili le informazioni del labiale (si noti nuovamente il vantaggio descritto sopra dell'esperienza audio-visiva del parlato), all'83% di frasi corrette in presenza dei segnali manuali del *cued speech*.

Un ulteriore esempio di informazioni visive in grado di supportare l'ascolto in rumore per le persone con impianto deriva da uno studio che ha sfruttato come indizio per migliorare l'ascolto i segni dell'American Sign Language. Blom e colleghi (2016) hanno chiesto a 40 studenti sordi con IC di guardare tre brevi presentazioni a schermo: la prima conteneva un messaggio audio prodotto in una situazione silenziosa, la seconda conteneva un messaggio audio in situazione di brusio indistinto di altre voci, e la terza era ancora una volta presentata in rumore ma con il supporto di segni. È importante precisare che quest'ultima condizione si avvaleva della cosiddetta *Comunicazione Simultanea*, e non di vera e propria lingua dei segni (nota 6). Quando i materiali verbali erano accompagnati da segni, la comprensione del

messaggio era migliore rispetto alla condizione di ascolto in rumore senza segni. Come nell'esempio precedente, un supporto visivo (questa volta di natura lessicale-semantic) poteva aiutare la comprensione del parlato nel rumore.

### **Addestrare il sistema cognitivo**

Gli esempi descritti nella sezione precedente mettono in evidenza l'utilità di aggiungere informazioni visive (o anche presenti in altri sistemi sensoriali, come nell'esempio della ricerca di Cieřla et al., 2019 che usava segnali vibrotattili) per aiutare il sistema cognitivo a risolvere il problema dell'ascolto in rumore. È utile notare che molti dei casi descritti sopra la sola informazione visiva non è in grado di veicolare pienamente il messaggio linguistico. Il labiale da solo non lo permette, il cued-speech può riuscirci nella misura in cui coloro che lo utilizzano sono molto esperti, e i singoli segni della comunicazione simultanea non contengono la struttura sintattica che serve per veicolare il significato di una breve storia. Siamo quindi di fronte a esempi nei quali l'*aggiunta* di informazione visiva permette al cervello di formulare migliori predizioni linguistiche, favorendo così una migliore comprensione del parlato in rumore.

Tutto questo è senz'altro molto utile durante l'esperienza d'ascolto, ragion per cui è decisamente opportuno dare alle persone con IC (o portatrici di protesi acustiche, o ancora genericamente ipoacusiche) la possibilità di vedere il parlante da vicino, in classe, in una riunione, in un *cocktail party*. Analogamente, al di là dei dibattiti ideologici o culturali che contrappongono la comunicazione simultanea alle vere e proprie lingue dei segni, può rivelarsi utile per sostenere l'apprendimento in classe dell'alunno sordo (con o senza impianto) che conosce i segni attraverso l'uso di elementi del lessico visuo-gestuale. (Per dati a supporto del ruolo che la comunicazione simultanea può giocare nei processi di apprendimento e memorizzazione degli studenti sordi, si veda: Marschark et al., 2008).



Una prospettiva complementare e aggiuntiva è quella di cercare di sfruttare gli stessi meccanismi di facilitazione efficaci durante l'ascolto, per potenziare *a lungo termine* le abilità necessarie all'ascolto in rumore. In breve, sfruttarli come strategie per apprendere ad ascoltare meglio in rumore, anche quando le informazioni multisensoriali aggiuntive non saranno più presenti. Per esemplificare, pensiamo al caso di una persona con IC alla quale sia data la possibilità di migliorare il suo ascolto della lingua attraverso un addestramento che sfrutti la multisensorialità del parlato (ascoltare e vedere il labiale) e che possa poi usare questa abilità appresa anche in un contesto in cui il labiale non sia più disponibile (es., l'ascolto al telefono). È possibile questo scenario? Le ricerche sugli apprendimenti multisensoriali suggeriscono che possa effettivamente essere così (Shams e Seitz, 2008; Moradi et al., 2013). Si è infatti osservato che esercitare un'abilità in maniera multisensoriale (es., un'abilità acustica addestrata attraverso un approccio che combini segnali acustici e visivi) può avere poi ricadute anche unisensoriali (ovvero sulle sole abilità acustiche, in assenza di indizi dagli altri sistemi sensoriali, come nel caso dell'ascolto al telefono).

Benché molte di queste ricerche siano ancora da realizzare, la strada appare promettente e alcuni gruppi di ricerca stanno cominciando a sfruttare esattamente questo principio, per potenziare le abilità utili all'ascolto in rumore. Un esempio recente è dato da una ricerca condotta in Svezia su persone anziane ipoacusiche che usavano protesi acustiche (Moradi et al., 2017). Gli autori hanno testato le capacità di ascolto in rumore dei partecipanti, prima e dopo una fase di addestramento. Metà dei partecipanti faceva l'addestramento con frasi presentate solo acusticamente, l'altra metà con frasi presentate in maniera audio-visiva (ovvero vedendo anche il volto del parlante). I risultati di questa ricerca hanno mostrato che solo il gruppo con addestramento audio-visivo migliorava la prestazione dell'ascolto in rumore, tanto nel primo test dopo la fine dell'addestramento, quanto in un test realizzato a

distanza di un mese dall'addestramento. Si tratta di un risultato preliminare, che tuttavia mette in evidenza l'importanza di estendere simili ricerche anche alle persone con IC.

Un secondo esempio deriva dalle ricerche che mirano a migliorare la percezione spaziale dei suoni nelle persone con impianto, attraverso protocolli di addestramento basati su feedback multisensoriali. Come ho anticipato, le persone con IC hanno grandi difficoltà a localizzare i suoni nell'ambiente, soprattutto quando l'impianto è unilaterale. La capacità di localizzare i suoni è però una delle poche abilità percettive che rimane modificabile per tutto l'arco della vita, potendo quindi essere addestrata (Carlile, 2017). Ad oggi sono stati fatti pochissimi tentativi di migliorare la capacità di localizzazione dei suoni nelle persone con IC, o anche in coloro che portano una protesi acustica, e persino nel contesto clinico non sono presenti protocolli di addestramento volti a migliorare la spazializzazione del suono. Ricerche in questa direzione, però, iniziano ad esserci. Ad esempio, in una ricerca stiamo cercando di addestrare la percezione spaziale dei suoni nelle persone con IC, usando paradigmi di apprendimento multisensoriale validati su persone udenti, nelle quali era stata simulata una difficoltà di ascolto tappando temporaneamente un orecchio (Rabini, Altobelli e Pavani, 2019). L'obiettivo è produrre un miglioramento delle capacità di localizzazione dei suoni che possa fornire al sistema cognitivo un elemento utile a separare gli oggetti e i flussi nella scena acustica.

### **Conclusioni**

In questa rassegna ho cercato di illustrare le principali ragioni per cui le persone con IC (e in verità praticamente tutte le persone con problemi di ipoacusia) hanno difficoltà ad ascoltare in contesti rumorosi. Quella che nella realtà fisica è una scena fatta di eventi sonori distinti, a livello della coclea (o dell'impianto) diventa una matassa da dipanare. Riuscire o meno a districare i suoni gli uni dagli altri per ricostruire la complessità degli ambienti sonori nei quali

## PERCEPIRE LA SCENA ACUSTICA CON L'IMPIANTO COCLEARE

solitamente viviamo è un compito complesso, che spesso richiede uno sforzo attivo. Per le persone con IC tutto ciò può essere un'ulteriore causa di variabilità nelle loro prestazioni e un limite alle interazioni con l'ambiente fisico e sociale. Promuovere le capacità di ascolto in rumore per le persone con IC è quindi una priorità fondamentale della ricerca: sia di quella ingegneristica, che sta cercando di proporre soluzioni sempre nuove per separare i segnali dal rumore, sia di quella cognitiva, che sta cercando di trovare strade per sfruttare al meglio le risorse che il nostro cervello ha a disposizione per realizzare questo compito complesso.

### **Ringraziamenti**

Questo articolo è il risultato della rielaborazione di un mio intervento in occasione della Terza conferenza nazionale sulla sordità organizzata dall'Ente Nazionale Sordi (Napoli, 20-22 Febbraio 2020). Ringrazio Pasquale Rinaldi e Camillo Galluccio per avermi dato l'opportunità di presentare questa relazione. Sono infine riconoscente anche a coloro che durante e dopo questa presentazione mi hanno stimolato a riflettere ulteriormente su queste tematiche, in particolare Mariapaola Scuderi e Valerio Leonetti. A loro, come anche a Fioralba Burnelli e Daniele Tasso, devo inoltre l'attenta rilettura di una precedente versione del manoscritto e diversi suggerimenti su come migliorarlo. Durante la stesura di questo articolo l'autore era finanziato dalla Fondazione MEDISITE (Francia), della Fondazione Neurodis (Francia), dalla Agence Nationale de la Recherche (Francia) e da un fondo del Ministero dell'Università e della Ricerca (PRIN, prot. 20177894ZH).

**Note**

1. Vale la pena di ricordare qui che l'accesso alla lingua orale attraverso la lettura labiale è molto limitato, anche per le persone sorde che in media hanno prestazioni migliori in compiti di lettura labiale. Ad esempio, una ricerca condotta in Gran Bretagna su un campione di oltre 100 persone sorde e udenti (età 18-68 anni) ha mostrato che le persone sorde hanno prestazioni migliori degli udenti in compiti di lettura labiale. Tuttavia, la stessa ricerca mostra anche che l'accuratezza media si colloca tra l'80 e il 90% di risposte corrette per le parole singole in un compito di scelta multipla (il partecipante deve scegliere fra un insieme di disegni quello che raffigura la parola pronunciata, dopo aver osservato il labiale), ma decade al 65-80% per frasi semplici o per le storie (anche in questo caso in compiti di scelta multipla). Se consideriamo che questo è ciò che si ottiene in condizioni di laboratorio controllate (quando il parlante è visto su uno schermo, in condizioni di illuminazione ideale, si rivolge all'ascoltatore in maniera frontale, è consapevole di dover articolare in maniera chiara il messaggio verbale e i materiali verbali sono pochi e ben distinti fra loro) è ragionevole pensare che questi numeri siano una sovrastima delle reali capacità di lettura labiale nella vita quotidiana.
2. SNR è l'acronimo per l'espressione inglese *Signal-to-Noise Ratio* (rapporto tra segnale e rumore): una grandezza numerica che deriva dalla divisione fra la potenza di un segnale e la potenza del rumore. SNR è espresso in decibel (dB) e dunque questa grandezza numerica è calcolata come 10 volte il logaritmo in base 10 della divisione fra la potenza del segnale ( $P_{segnale}$ ) e la potenza del rumore ( $P_{rumore}$ ). In formula:  $SNR = 10 \log_{10} (P_{segnale} / P_{rumore})$ . Quello che è utile ricordare è che quanto più grande è il valore di SNR, tanto più il segnale prevale sul rumore. Quindi il segnale è più intenso del rumore nella condizione +10dB SNR rispetto alla condizione +5dB SNR.

3. Lo studio dell'ascolto in rumore comincia negli anni '50 in Gran Bretagna, grazie al lavoro di un ingegnere delle telecomunicazioni: Edward Colin Cherry (Cherry, 1953). Questo scienziato cognitivo *ante litteram* fu uno dei primi studiosi ad analizzare problematiche della psicologia della percezione e del linguaggio attraverso le nozioni di ingegneria e di analisi del segnale. Il termine 'cocktail party', che evoca uno scenario piuttosto anglosassone e borghese di ritrovo nel quale le conversazioni per noi rilevanti si sovrappongono a un rumore di fondo indistinto, è diventato sinonimo del problema dell'ascolto in rumore. Per questo, solitamente, non viene tradotto in italiano.
4. Risultati analoghi li troviamo anche nel riconoscimento delle emozioni presenti nella voce. Sebbene le persone con IC abbiano maggiori difficoltà rispetto ai controlli udenti in questo compito uditivo, esse beneficiano ampiamente di informazioni audio-visive congruenti (es., voce triste e volto triste; Agrawal et al., 2013; Chatterjee et al., 2014; Fengler et al., 2017).
5. Ad oggi non ci è dato di sapere come percepiscano i suoni le persone con sordità bilaterale e IC, perché si tratta di un'esperienza soggettiva che non è facile sondare. Gli adulti con sordità post-verbale che portano un impianto riportano esplicitamente l'esperienza di ascoltare la loro lingua con suoni diversi. I bambini con sordità pre-verbale invece non possono darci facilmente una risposta, perché la loro esperienza acustica è cominciata direttamente con l'IC. Ciononostante è pratica comune nella ricerca simulare l'ascolto con l'IC in persone udenti usando i cosiddetti suoni 'vocoded'. In estrema sintesi, si tratta di filtrare i suoni che si vogliono proporre all'ascoltatore attraverso le medesime procedure di trasformazione del segnale adottate dal processore dell'IC. Il suono che ne

risulta è spesso metallico e difficilmente intellegibile (per una simulazione di parlato inglese vocoded si veda: <https://auditoryneuroscience.com/prosthetics/noise-vocoded-speech>). Bisogna però essere consapevoli che si tratta di un'approssimazione: questo è forse il segnale trasformato dal processore e inviato al nervo acustico, ma non ci è dato di sapere come questo segnale sarà percepito dopo le successive trasformazioni che portano lo stimolo dal nervo alle altre regioni del cervello, fino a farlo diventare un oggetto acustico per la nostra consapevolezza percettiva.

6. Il termine *Comunicazione Simultanea* (o *SimCom*) è usato in Nord-America per descrivere quello che in Europa è noto come *Sign-Supported Speech* (parlato con supporto di segni) e corrisponde a ciò che in Italiano chiamiamo Italiano Segnato. Si tratta di una forma di comunicazione audio-visiva che utilizza segni presi dalla lingua dei segni del paese di riferimento e li abbina al parlato, seguendo l'ordine delle parole della lingua orale. Nel caso dell'Italiano Segnato quindi utilizza i segni della Lingua dei Segni Italiana (LIS) e li abbina all'Italiano. Affinché questa forma di comunicazione audio-visiva sia efficace, è necessario che la persona che la riceve o la utilizza abbia una conoscenza dei diversi segni utilizzati, ma non necessariamente che conosca la grammatica della lingua dei segni da cui provengono, perché questa non viene utilizzata (come detto, si usa l'ordine delle parole dell'Italiano). In questo senso è dunque evidente che l'Italiano Segnato non è la Lingua dei Segni Italiana e, più in generale, che la Comunicazione Simultanea non è una lingua naturale come la lingua dei segni.

**Bibliografia**

- Agrawal, D., Thorne, J. D., Viola, F. C., Timm, L., Debener, S., Büchner, A., et al. (2013). Electrophysiological responses to emotional prosody perception in cochlear implant users. *NeuroImage: Clinical*, 2, 229-238.
- Amenta, S., Artesini, L., Musola, D., Frau, G.N., Vespignani, F., & Pavani, F. (under revision). Probing Language Processing in Cochlear Implant Users with Visual Word Recognition: Effects of Lexical and Orthographic Word Properties. *Language, Cognition and Neuroscience*.
- Bayard, C., Machart, L., Strauß, A., Gerber, S., Aubanel, V., & Schwartz, J. L. (2019). Cued Speech Enhances Speech-in-Noise Perception. *The Journal of Deaf Studies and Deaf Education*, 24(3), 223-233.
- Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press.
- Caldwell, A., & Nitttrouer, S. (2013). Speech perception in noise by children with cochlear implants. *Journal of Speech, Language and Hearing Research*, 56, 13–30.
- Carlile, S. (2014). The plastic ear and perceptual relearning in auditory spatial perception. *Frontiers in neuroscience*, 8, 237.
- Chatterjee, M., Zion, D. J., Deroche, M. L., Buriánek, B. A., Limb, C. J., Goren, A. P., et al. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing research*, 322, 151-162.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 25(5), 975-979.
- Cleary, M., Pisoni, D.B., 2002. Talker discrimination by prelingually deaf children with cochlear implants: preliminary results. *The Annals of Otology, Rhinology & Laryngology*, 119, 113-118.



- Cornett, R. O. (1967). Cued speech. *American Annals of the Deaf*, 112, 3–13.
- Cristofari, E., Cuda, D., Martini, A., Forli, F., Zanetti, D., Di Lisi, D., et al. (2017). A Multicenter clinical evaluation of data logging in cochlear implant recipients using automated scene classification technologies. *Audiology and Neurotology*, 22(4-5), 226-235.
- Dehaene, S. (2019). *Imparare. Il talento del cervello, la sfida delle macchine*. Raffaello Cortina Editore.
- Dorman, M. F., & Gifford, R. H. (2017). Speech understanding in complex listening environments by listeners fit with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 60(10), 3019-3026.
- Dorman, M. F., Liss, J., Wang, S., Berisha, V., Ludwig, C., & Natale, S. C. (2016). Experiments on auditory-visual perception of sentences by users of unilateral, bimodal, and bilateral cochlear implants. *Journal of Speech, Language, and Hearing Research*, 59(6), 1505–1519.
- Fengler, I., Nava, E., Villwock, A. K., Büchner, A., Lenarz, T., & Röder, B. (2017). Multisensory emotion perception in congenitally, early, and late deaf CI users. *PloS one*, 12(10), e0185821.
- Fu, Q.J., Chinchilla, S., & Galvin, J.J. (2004). The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users. *Journal of the Association for Research in Otolaryngology*, 5, 253e260.
- Hahne, A., Wolf, A., Müller, J., Mürbe, D., & Friederici, A. D. (2012). Sentence comprehension in proficient adult cochlear implant users: On the vulnerability of syntax. *Language and cognitive processes*, 27(7-8), 1192-1204.

- Healy, E. W., Yoho, S. E., Wang, Y., & Wang, D. (2013). An algorithm to improve speech recognition in noise for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *134*(4), 3029-3038.
- Lidestam, B., Moradi, S., Pettersson, R., & Ricklefs, T. (2014). Audiovisual training is better than auditory-only training for auditory-only speech-in-noise identification. *The Journal of the Acoustical Society of America*, *136*(2), EL142-EL147.
- Litovski, R.Y., Goupell, M.J., Misurelli, S.M., & Kan, A. (2017). Hearing with cochlear implants and hearing aids in complex auditory scenes. In *The auditory system at the cocktail party* (pp. 261-291). New York: Springer.
- Lund, E. (2016). Vocabulary knowledge of children with cochlear implants: A meta-analysis. *Journal of Deaf Studies and Deaf Education*, *21*(2), 107-121.
- Marschark, M., Sapere, P., Convertino, C., Pelz, J. (2008). Learning via direct and mediated instruction by deaf students. *Journal of Deaf Studies and Deaf Education*. *13*, 446–461.
- Massida, Z., Belin, P., James, C., Rouger, J., Fraysse, B., Barone, P., & Deguine, O. (2011). Voice discrimination in cochlear-implanted deaf subjects. *Hearing research*, *275*(1-2), 120-129.
- Moore, D. R., & Shannon, R. V. (2009). Beyond cochlear implants: awakening the deafened brain. *Nature Neuroscience*, *12*(6), 686.
- Moradi, S., Lidestam, B., & Rönnerberg, J. (2013). Gated audiovisual speech identification in silence vs. noise: Effects on time and accuracy. *Frontiers in Psychology*, *4*, 359.
- Navarra, J., Yeung, H. H., Werker, J. F., & Soto-Faraco, S. (2012). Multisensory Interactions in Speech Perception. In B. E. Stein (Ed.), *The new handbook of multisensory processing*. MIT Press.

- Pavani, F., Venturini, M., Baruffaldi, F., Artesini, L., Bonfioli, F., Frau, G. N., & van Zoest, W. (2017). Spatial and non-spatial multisensory cueing in unilateral cochlear implant users. *Hearing research, 344*, 24-37.
- Rabini, G., Altobelli, E., & Pavani, F. (2019). Interactions between egocentric and allocentric spatial coding of sounds revealed by a multisensory learning paradigm. *Scientific reports, 9*(1), 1-12.
- Rinaldi, P., Caselli, M.C., & Pavani, F. (2017). Sviluppo del linguaggio e processi cognitivi nei bambini sordi. In S. Vicari e M.C. Caselli (Eds), *Neuropsicologia dello Sviluppo*. Bologna: Il Mulino.
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences, 104*(17), 7295-7300.
- Scharenborg, O., & van Os, M. (2019). Why listening in background noise is harder in a non-native language than in a native language: A review. *Speech Communication, 198*, 53-64.
- Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in cognitive sciences, 12*(11), 411-417.
- Shinn-Cunningham, B., Best, V., & Lee, A. K. (2017). Auditory object formation and selection. In J.C. Middlebrooks, J.Z., Simon, A.N. Popper, R.R. Fay (Eds), *The auditory system at the cocktail party* (pp. 7-40),. New York: Springer.
- Spahr, A. J., Dorman, M. F., & Loisel, L. H. (2007). Performance of patients using different cochlear implant systems: Effects of input dynamic range. *Ear and Hearing, 28*(2), 260–275.

Srinivasana, A. G., Padilla, M., Shannon, R. V., & Landsberger, D. M. (2013). Improving speech perception in noise with current focusing in cochlear implant users. *Hearing Research*, 299, 29–36.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26(2), 212-215.

Wilson, B. S. (2019). The remarkable cochlear implant and possibilities for the next large step forward. *Acoustics Today*, 15, 53-61.

Figura 1

Figura 1. Il problema della scena acustica. (Modificato da:  
<https://spectrum.ieee.org/consumer-electronics/audiovideo/deep-learning-reinvents-the-hearing-aid>)

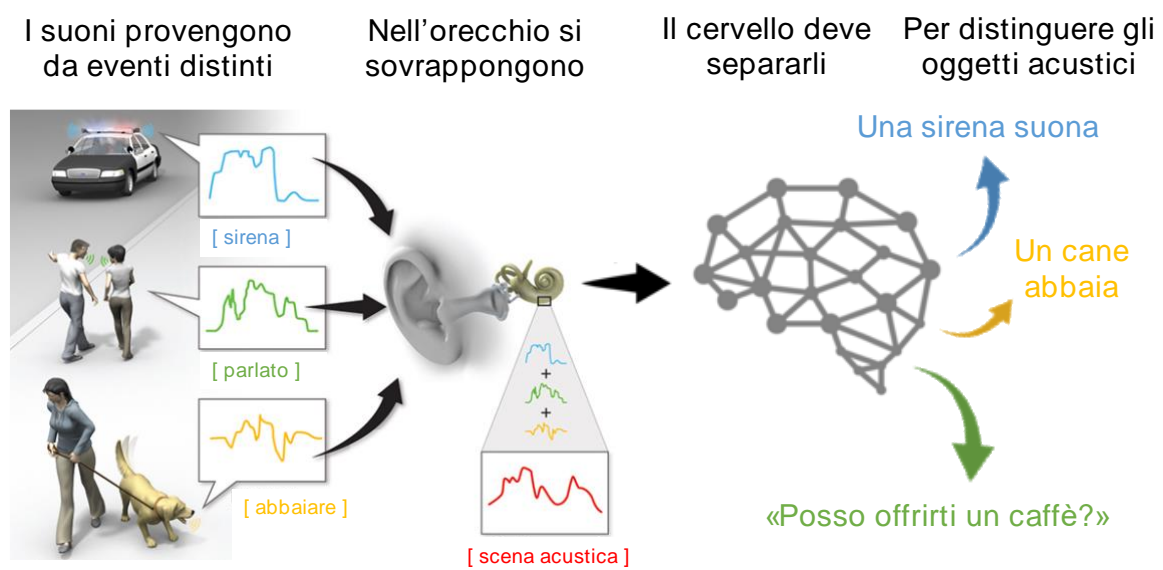
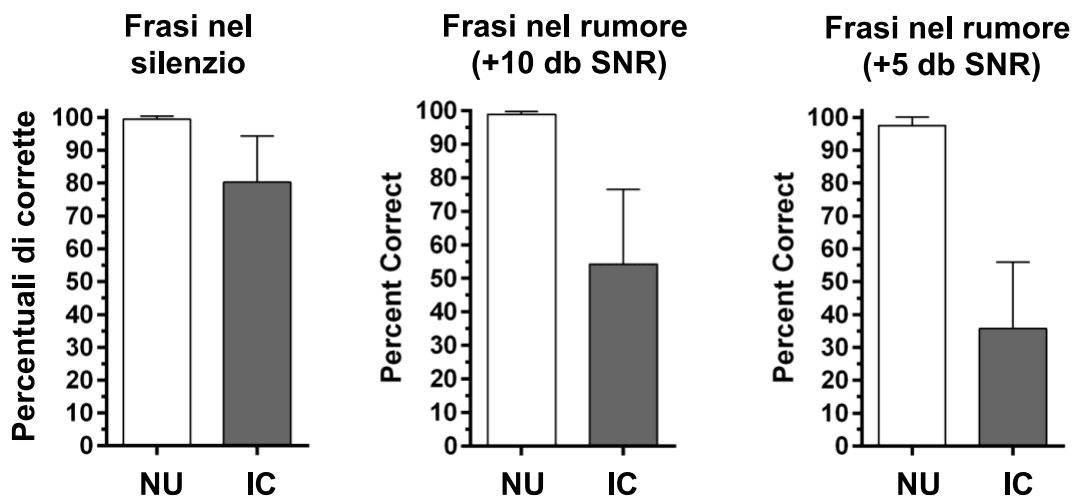


Figura 2

Figura 2. Le persone con IC hanno difficoltà ad ascoltare nel rumore. NU: Normo-udenti; IC: Impianto Cocleare. (Modificato da: Dorman e Gifford, 2017)



**Figura 3**

Figura 3. Migliorare l'ascolto in rumore. Soluzioni tecnologiche: (A) Ascolto in silenzio o rumore per persone con IC: le barre nere indicano la prestazione in contesto di silenzio, quelle grigie e quelle a righe la prestazione in contesto di rumore (ristorante o cocktail party); tratto da: Dorman e Gifford, 2017; (B) Miglioramento osservato in contesto di brusio indistinto per diversi partecipanti ipoacusici (HI: Hearing Impaired); tratto da: Healy et al., 2013. Soluzioni cognitive: (C) Effetto dell'ascolto in rumore in persone con IC, in funzione della presenza o assenza di concomitanti informazioni visive dal labiale; tratto da: Dorman et al., 2016. (D) Effetto dell'ascolto in rumore in persone con IC bilaterale, in funzione del fatto che l'ascolto avvenisse con uno solo dei due impianti (quello con le migliori prestazioni audiometriche) o con entrambi gli impianti; tratto da: Loïselle et al., 2016.

