

I quanta acustici e la teoria dell'udito^a

Dennis Gabor

Nelle esposizioni introduttive alla meccanica ondulatoria, si ricorre spesso ad esempi di natura acustica, come ha fatto con particolare successo Landé¹. In un articolo recente sulla «Teoria della comunicazione»² ho seguito la direzione opposta, descrivendo fenomeni acustici con esempi matematici strettamente collegati ai metodi della teoria quantistica. In fisica acustica non ci si può attendere che nuovi approcci formali a vecchi problemi possano rivelare molto più di quanto non sia già noto, ma nell'acustica soggettiva la situazione è piuttosto differente. In effetti, i nuovi metodi hanno già comprovato il loro valore euristico, e ci si può attendere che facciano maggior luce sulla teoria dell'udito. Nel mio articolo precedente, il punto di vista era principalmente quello dell'ingegneria della comunicazione: nel riepilogo che ne faccio qui di seguito, metto l'accento su caratteristiche che potranno essere d'interesse per fisici e fisiologi.

Che cosa udiamo? La risposta fornita dai normali libri di testo può essere accettata cum grano salis da ben pochi studiosi, o forse proprio da nessuno. Secondo la teoria legata principalmente ai nomi di Ohm e di Helmholtz, l'orecchio analizza il suono nelle sue componenti spettrali, e le nostre sensazioni uditive sono costituite dai componenti di Fourier, o meglio dal loro valore assoluto. Ma l'analisi di Fourier è una descrizione atemporale in termini di onde esattamente periodiche di durata infinita. D'altra parte l'esperienza comune ci dice che il suono ha un andamento temporale e un andamento di frequenza. Questa dualità delle nostre sensazioni non trova espressione né nella descrizione del suono come segnale $s(t)$ in funzione del tempo, né nella sua rappresentazione con componenti di Fourier $S(f)$. Si vuole una descrizione matematica che tenga conto *ab ovo* di questa dualità. Consideriamo, perciò, sia il tempo sia la frequenza come coordinate del suono, cercando di vedere che significato possiamo dare a una tale rappresentazione.

Se t e f sono messe in gioco come coordinate ortogonali, si ottiene un diagramma che può essere chiamato “diagramma informativo” (Fig. 1). Una semplice oscillazione armonica è rappresentata da una linea verticale con ascissa f , mentre un impulso istantaneo (funzione delta) è rappresentato da una linea orizzontale all'“istante” [*epoch*] t . Si tratta di casi estremi. In generale, i segnali non possono essere rappresentati da linee rette; però sarà possibile associare ad essi un rettangolo caratteristico o “cella” mediante il processo che segue, il quale a prima vista forse potrà apparire piuttosto complicato.

^aTraduzione a cura di Agostino Di Scipio di Dennis Gabor (1947), “Acoustical quanta and the theory of hearing”, apparso il 3 maggio 1947 su «Nature», n. 4044 (vol. 159), pp. 591-594. Laddove possibile, i riferimenti bibliografici vengono qui redatti per esteso. Le note del traduttore sono ordinate alfabeticamente. La fonte della traduzione è una copia su microfilm conservata presso la biblioteca della Simon Fraser University di Burnaby (Canada). Si ringrazia Davide Rochesso per consulenza e suggerimenti nel trattamento delle notazioni matematiche adottate da Gabor.

¹Landé 1930.

²Gabor 1946.

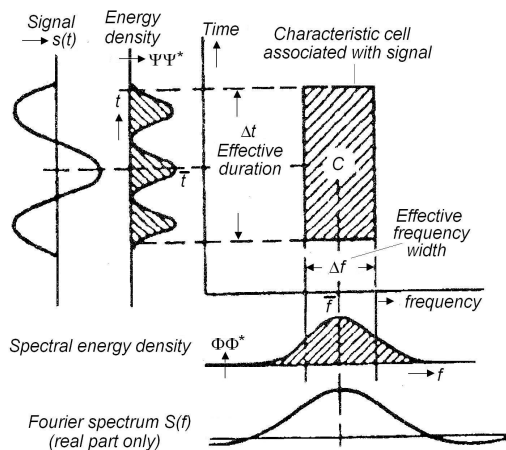


Figura 1: Rettangolo caratteristico di un segnale nel diagramma informativo.

Si consideri un dato segnale descritto come $s(t)$ con il “linguaggio del tempo” e come $S(f)$, la sua trasformata di Fourier, con il “linguaggio della frequenza”. Se $s(t)$ è reale, $S(f)$ sarà in genere complesso e lo spettro si estenderà a frequenze sia positive sia negative. Il che crea un’ indesiderata asimmetria tra le due rappresentazioni, eliminabile operando, invece che col segnale reale $s(t)$, con un segnale complesso del tipo $\psi(t) = s(t) + i\sigma(t)$, dove $\sigma(t)$ è la trasformata di Hilbert di $s(t)$. Questa scelta rende la trasformata di Fourier $\varphi(f)$ di $\psi(t)$ pari a zero per tutte le frequenze negative. Definiamo poi la “densità di energia” del segnale come $\psi\psi^*$, dove l’asterisco denota il valore complesso coniugato, e definiamo quindi la [“densità spettrale di energia”] come $\varphi\varphi^*$. In Fig. 1 le due distribuzioni di energia sono mostrate come aree più scure. Esse hanno la medesima estensione: l’energia totale del segnale è la stessa per entrambe le definizioni. Possiamo ora definire un “tempo medio” \bar{t} del segnale, e similmente una “frequenza media” \bar{f} , quali coordinate dei centri di gravità delle due distribuzioni. Questo ci dà un punto C nel diagramma informativo come centro del segnale. Procediamo definendo la “durata efficace” Δt e la “banda di frequenza efficace” Δf del segnale per mezzo delle deviazioni medie quadratiche delle due distribuzioni di energia rispetto ai valori medi \bar{t} e \bar{f} . Nella figura queste corrispondono ai raggi inerziali delle due aree scure. Per ragioni che saranno evidenti più avanti, conviene definire Δt e Δf come $2\sqrt{\pi}$ volte i rispettivi raggi inerziali[†]. Con tali definizioni si può applicare una relazione matematica, la disuguaglianza di Schwarz, secondo la quale

$$\Delta t \Delta f \geq 1 : \tag{1}$$

vale a dire che *l’area del rettangolo caratteristico o cella di un segnale è almeno pari all’unità*^b. Questa è l’esatta formulazione della relazione di indeterminazione tra tempo e frequenza. È chiaro ora che quello che abbiamo ottenuto è un modello classico della meccanica di onde stazionarie unidimensionali nel quale l’unità sostituisce la costante di Planck h . Formalmente

[†]Nel mio articolo precedente, il fattore era $\sqrt{2\pi}$, invece che $2\sqrt{\pi}$, e pertanto la (1) era $\Delta t \Delta f \geq \frac{1}{2}$. La scelta presente è più vantaggiosa per una diretta comparazione con la teoria quantistica.

^bPossiamo dire, con terminologia più diretta, $\Delta t = \sqrt{2}$ ·durata rms, e $\Delta f = \sqrt{2}$ ·frequenza di banda rms.

questo “quantum” è rappresentato da una cella di area unitaria (adimensionale) nel diagramma informazionale. Il suo significato fisico diventa evidente dalle seguenti considerazioni.

Se in un intervallo di tempo T un segnale è espanso in una serie di Fourier, vi saranno T componenti complesse di Fourier per banda frequenziale di estensione unitaria [*unit frequency band-width*]. Perciò possiamo dire che il diagramma informazionale contiene un dato complesso per area unitaria. (Questo, naturalmente, va considerato come un teorema asintotico per aree molto grandi, dato che con la (1) non possiamo analizzare l’area di informazione in celle di ordine inferiore all’unità). Vediamo ora che i quanta in questo modello di meccanica d’onda sono *quanta di informazione*. Ogni quantum rappresenta un dato numerico complesso o due dati reali.

Questo risultato, in combinazione con la relazione di incertezza (1), suggerisce l’interrogativo se vi siano segnali per i quali la diseuguaglianza diventi uguaglianza, e che possano essere utilizzati per la rappresentazione di quanta. Si può dimostrare che segnali semplicissimi o “elementari” di questo genere esistono, e che la loro forma è

$$s(t) = \exp -\alpha^2(t - t_0)^2 \cdot \exp i2\pi f_0 t \quad (2)$$

dove α è una costante reale. Perciò i segnali elementari sono oscillazioni armoniche di qualsiasi frequenza f_0 , modulate da un impulso gaussiano. La loro trasformata di Fourier è della stessa forma matematica

$$S(f) = \exp -(\pi/\alpha)^2(f - f_0)^2 \cdot \exp i2\pi t_0 f [^c]. \quad (3)$$

Le componenti reali dei segnali elementari sono convenientemente specificate come segnali tipo coseno e seno, e sono illustrate in Fig. 2 per diversi valori della costante α , che è connessa alla durata effettiva e alla larghezza di banda effettiva per mezzo delle equazioni $\Delta t = \sqrt{\pi}/\alpha$ e $\Delta f = \alpha/\sqrt{\pi}$. Se $\alpha = 0$ i segnali elementari diventeranno semplici oscillazioni armoniche; se invece $\alpha = \infty$, diventano la funzione delta e la sua derivata.

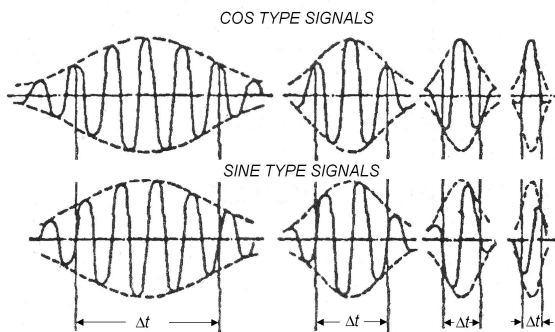


Figura 2: Segnali elementari.

Questi segnali elementari possono essere utilizzati per la rappresentazione di quanta di informazione, poiché qualsiasi segnale arbitrario può essere espanso nei loro termini, per qualsiasi valore della costante α . Vale a dire, possiamo espandere un segnale dividendo

^cLe formule (2) e (3) possono essere rese oggi più leggibili. La prima: $e^{-\alpha^2(t-t_0)^2} e^{i2\pi f_0 t}$. La seconda: $e^{-(\pi/\alpha)^2(f-f_0)^2} e^{i2\pi t_0 f}$.

l'area di informazione in celle unitarie rettangolari di qualsiasi aspetto $\Delta t / \Delta f = \pi / \alpha^2$ ed associando a ciascuna cella un segnale elementare con fattore di ampiezza complesso c_{ik} . Otteniamo quindi una sorta di matrice, come si vede in Fig. 3. Questo metodo di analisi contiene il "linguaggio del tempo" e il "linguaggio della frequenza" come casi speciali estremi. Se le celle sono di lunghezza infinita nella direzione del tempo otteniamo l'analisi di Fourier, se sono di lunghezza infinita nella direzione della frequenza otteniamo un'espansione in funzioni delta, vale a dire lo stesso segnale $s(t)$.

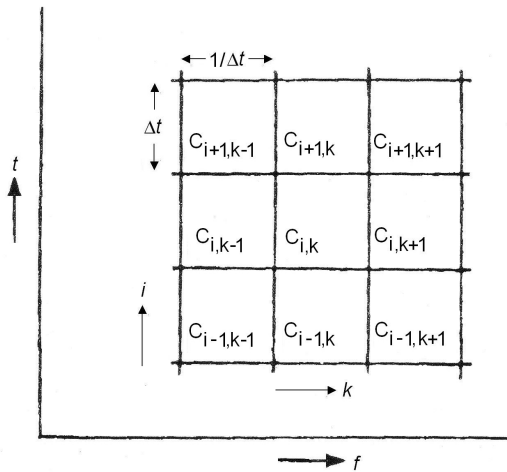


Figura 3: Matrice di segnali elementari rappresentante un segnale arbitrario.

In generale, quello dell'espansione in funzioni elementari è un processo piuttosto sconvolgente, poiché tali funzioni non formano un insieme ortogonale. Tuttavia esiste un metodo sperimentale che, sebbene non proprio equivalente, offre risultati piuttosto simili e molto validi. Si tratta del metodo di "spettrografia sonora" sviluppato presso i laboratori della Bell Telephone durante la Guerra, e pubblicato quando il precedente articolo di chi scrive era già stato redatto e accettato per la pubblicazione³. Non c'è spazio qui per descrivere questo metodo e la differenza che esiste tra i suoi "ritratti sonori" [*sound portraits*] e le mie matrici di segnali elementari, ma si può notare che tali "ritratti sonori" non sono rappresentazioni complete del suono, pur contenendo gran parte delle caratteristiche soggettivamente importanti.

Ci si attende che l'analisi delle sensazioni uditive possa rivelarsi terreno fertile per nuovi metodi di rappresentazione; i risultati ottenuti fino ad oggi sembrano indicare molto chiaramente che il "quantum di suono" [*quantum of sound*] è un concetto di considerevole significato fisiologico. Chi lavora in acustica ha raccolto una gran mole di materiale circa la soglia di discriminazione dell'orecchio per differenze di tempo e frequenza. Sembra che almeno una buona parte di questi risultati possa riassumersi con l'affermazione che *l'orecchio possiede un'area di soglia di discriminazione di ordine unitario*.

Una serie di esperimenti presi in esame è quella di Bürck, Kotowsky e Lichte⁴. Il loro oggetto di studio era la soglia di discriminazione temporale in differenti registri di frequenza.

³Koenig, Dunn e Lacy 1946; Potter 1945.

⁴Bürck, Kotowsky e Lichte 1935a,b.

Mach aveva dimostrato che se un'oscillazione sinusoidale risuona per pochi cicli, l'orecchio la percepirà come un rumore, mentre al di sopra di una certa durata minima, di circa 10 millisecondi, sarà ascoltata come una nota musicale breve ma di frequenza accertabile^d. Gli autori hanno condotto una serie di esperimenti su questo fenomeno, e in una seconda occasione sono andati oltre, realizzando esperimenti in cui l'intensità del suono, dopo una certa durata dall'inizio, veniva raddoppiata. Al di sotto di un certo valore minimo (21 msec. a 500 Hz), il suono non poteva essere distinto rispetto ad una nota che suonasse con intensità doppia sin dall'inizio, mentre solo al di sopra di quel valore poteva essere udito per quel che era. Si potrebbe dire che nell'intervallo tra 10 e 21 msec. l'orecchio sia appena pronto a rilevare una seconda sensazione distinta.

Un'altra serie di esperimenti è quella di Shower e Biddulph⁵, e riguarda la soglia di discriminazione frequenziale dell'orecchio. Il loro metodo sperimentale era del tutto diverso. La frequenza di una nota veniva modulata entro limiti ridotti, e veniva annotata la minima deviazione che desse l'impressione di un "trillo". A 500 Hz questa deviazione minima era di 2.3 Hz. Al di sotto, il trillo non poteva essere distinto da una nota non modulata. Il ciclo di modulazione fu scelto in modo tale che desse la discriminazione di frequenza più fine (0.5 sec. per l'intero ciclo di modulazione). In questi esperimenti, come in quelli richiamati prima, l'intensità effettiva non aveva un ruolo essenziale, essendo le prove svolte ad un livello di buona udibilità.

Sia negli esperimenti di Shower e Biddulph, volti alla discriminazione di frequenza, sia in quelli di Bürck, Kotowsky e Lichte, che riguardano la risposta temporale dell'orecchio, in realtà i fenomeni osservati riguardano naturalmente sia il tempo sia la frequenza. Negli esperimenti di Bürck, Kotowsky e Lichte, veniva considerata l'effettiva ampiezza di un'oscillazione sinusoidale di durata finita. Nelle prove di Shower e Biddulph è necessario un tempo di 0.25 sec. per notare un ciclo di modulazione completo, verso l'alto o verso il basso, della frequenza. Perciò in entrambi i casi, con i rispettivi intervalli di tempo e frequenza considerati, è possibile formare un prodotto, ovvero una "area di soglia" all'interno della quale l'orecchio era in grado di rilevare una sola sensazione. Occorreva oltrepassare quell'area di soglia per rilevare una seconda sensazione. Il valore numerico di tale area può essere definito, in una certa misura. Con le convenzioni adottate nel mio articolo precedente, l'area di soglia relativa ai dati di Bürck, Kotowsky e Lichte era 1.05 a 500 Hz e 1.5 a 1000 Hz; quella relativa a Shower e Biddulph è di 1.17 sia a 500 Hz sia a 1500 Hz. Naturalmente non si può vantare particolare accuratezza, dato che il tutto dipende dalla precisione dei dati forniti; tuttavia è ragionevole pensare che l'area di soglia tenda al valore unitario, vale a dire approssimi il "limite di quantum acustico". Si deve comprendere, naturalmente, che vi è una differenza importante tra il quantum acustico rilevato da uno strumento di misurazione e quello rilevato dall'orecchio. Negli esperimenti considerati sopra, l'orecchio veniva chiamato a rispondere soltanto "sì" o "no" a delle semplici domande. Ad uno strumento di misura, invece, un quantum di informazione condurrebbe un dato numerico complesso (due dati reali), e ciascun dato porterebbe con sé un numero infinito di "sì" e di "no". La differenza va espressa parlando di "discriminazione" nel caso dell'orecchio, e di "rilevamento" [*registering*] nel caso di uno strumento di misura. In questo senso possiamo dire che, almeno in esperimenti semplici, le migliori orecchie possono appena discriminare un solo quantum acustico nel registro frequenziale ottimale.

^dSi riferisce all'articolo di Ernst Mach del 1865, "Untersuchungen über den Zeitsinn des Ohres", in: *Sitzungsberichte der Wiener Akademie der Wissenschaften: Mathematisch-Naturwissenschaftliche* 51, pp. 133-150.

⁵Shower e Biddulph 1931.

L'importanza di tale risultato sta nell'ambito notevolmente vasto di rapporti cui sembra applicarsi. Come mostra la Fig. 4, nel caso di segnali di circa 500 Hz, l'area di soglia non varia apprezzabilmente mentre invece l'intervallo di tempo varia da circa 20 a 250 msec., vale a dire di un fattore superiore a dieci. Entro questi ampi confini l'orecchio ha la capacità di cercare,

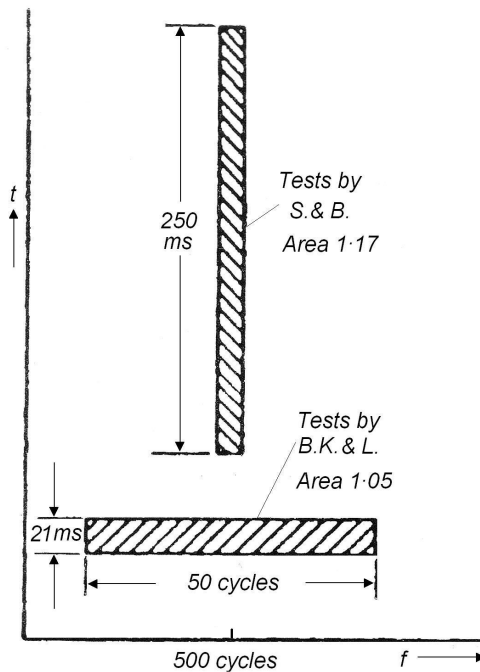


Figura 4: Aree di soglia dell'orecchio a 500 Hz.

con un aggiustamento automatico, i dettagli più fini della configurazione sonora che gli si presenta. Se un suono breve colpisce l'orecchio, con frequenza scarsamente determinabile, la discriminazione temporale è dell'ordine di 20 msec. Se il suono è prolungato, in modo tale che la banda di frequenza si riduce soltanto a pochi cicli al secondo, la discriminazione frequenziale dell'udito si mantiene al passo con questo processo per almeno un quarto di secondo. Ricompare qui con tutta evidenza un vecchio problema dell'acustica fisiologica. Non esistono sistemi fisici di risonatori che possano vantare questo grado di perfezione. I risonatori possono essere tarati su bande larghe o su bande strette, ma non su entrambe allo stesso tempo. Da dati fisiologici, e in particolare dagli esperimenti di Wegel e Lane⁶ sul "mascheramento", è praticamente accertato che i risonatori dell'orecchio sono ampiamente smorzati, con tempi di decadimento dell'ordine di 10 msec. o anche meno. Essi sono tarati su bande larghe all'incirca quanto necessario per spiegare i risultati di Bürck, Kotowsky e Lichte, ma non quelli di Shower e Biddulph, per i quali sarebbero necessari risonatori con decadimento almeno dieci volte più lungo. Pertanto sembra inevitabile dover postulare un secondo meccanismo nell'orecchio che spieghi la capacità di adeguarsi in modo progressivamente più preciso nel caso di suoni prolungati.

⁶Wegel e Lane 1924.

Il ruolo di questi due “meccanismi” è illustrato in modo approssimativo in Fig. 5. Il primo è probabilmente il vero e proprio meccanismo dei risonatori dell’orecchio; il secondo è quasi certamente un fenomeno “non-meccanico”, e si sarebbe tentati di localizzarlo nel cervello, anche se tale ipotesi non è inevitabile. Potrebbe ben trattarsi di un nuovo fenomeno

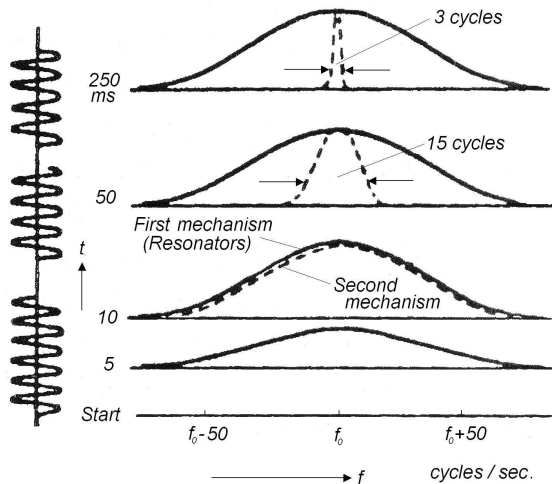


Figura 5: I due meccanismi dell’udito. Se si suona una nota per più di 10 msec., il pattern di risonanza dell’orecchio interno rimane immutato; ma un secondo meccanismo parte alla ricerca del massimo di eccitazione e lo localizza con accuratezza crescente per il tempo successivo fino a 250 msec.

di conduzione nervosa, e potrebbe spiegarsi con l’ipotesi che la conduzione in fibre nervose adiacenti è in qualche misura instabile, in modo che la fibra più fortemente eccitata gradualmente sopprime l’eccitazione delle fibre vicine eccitate meno fortemente. Quindi se viene suonata una nota pura per un tempo sufficiente, saranno eccitate sempre meno fibre, fino a quando rimane in azione solo la fibra corrispondente alla posizione di massima ampiezza nella membrana basilare. Ciò rende possibile accertare la frequenza con molta maggiore accuratezza di quanto è possibile attendersi dai risonatori meccanici. Sarebbe molto interessante, anche se probabilmente molto difficile, verificare o negare questo effetto, magari attraverso una versione più raffinata delle tecniche sperimentali di Galambos e Davis⁷.

È interessante notare che cominciamo a percepire il suono come “musicale” proprio nel punto in cui subentra il secondo meccanismo. Affinché il parlato sia perfettamente intelligibile basterebbe il primo meccanismo, ma per permetterci di apprezzare la musica, è necessario il secondo.

Infine, ci si può chiedere se questo modello acustico insegna qualcosa a chi si occupa di teoria quantistica. In senso formale la risposta deve essere ovviamente negativa. Tuttavia dal fatto che due differenti campi ammettano lo stesso trattamento formale consegue che, almeno per quanto riguarda la matematica, non vi può essere nulla nell’uno che non sia implicito nell’altro. Modelli familiari di fenomeni poco familiari hanno almeno il vantaggio

⁷Galambos e Davis 1943, 1944.

che semplificano la distinzione tra caratteristiche intrinseche del fenomeno e caratteristiche introdotte dal metodo di analisi. Si può essere inclini a pensare che condizioni nettamente definite, caratterizzate con numeri interi, siano peculiari dei fenomeni quantistici, o almeno che richiedano meccanismi speciali per farne un'imitazione con metodi classici, come corde e membrane. Ma abbiamo visto che nel modello acustico gli interi emergono come parte dello sfondo matematico prima che a questo stadio appaia un qualche fenomeno fisico. La ragione ultima per l'emergenza di "quanta acustici" è che abbiamo osservato lo stesso fenomeno simultaneamente sotto due aspetti, e lo abbiamo descritto con due "quantità di interesse", cioè tempo e frequenza. Si possono trovare innumerevoli altri modelli classici se soltanto ci poniamo simultaneamente due tipi di domanda sulla medesima cosa, o se seguiamo due linee di esperimenti su di essa, posto che le domande non siano né identiche né indipendenti.

Bibliografia

- Bürck, W., P. Kotowsky e H. Lichte (1935a). «Der Aufbau des Tonhöhen bewußtseins». In: *Elektrische Nachrichten Technik* 12, pp. 326–333.
- (1935b). «Die Hörbarkeit von Laufzeitdifferenzen». In: *Elektrische Nachrichten Technik* 12, pp. 355–362.
- Gabor, Dennis (1946). «Theory of communication». In: *Journal of the Institution of Electrical Engineers* 93.26. Part III: Radio and Communication Engineering, pp. 429–457.
- Galambos, Robert e Hallowell Davis (1943). «The response of single auditory-nerve fibers to acoustic stimulation». In: *Journal of Neurophysiology* 6.1, pp. 39–57. URL: <http://jn.physiology.org/content/jn/6/1/39.full.pdf> (visitato il 28/12/2015).
- (1944). «Inhibition of activity in single auditoru nerve fibers by acoustic stimulation». In: *Journal of Neurophysiology* 7.5, pp. 287–303. URL: <http://jn.physiology.org/content/jn/7/5/287> (visitato il 28/12/2015).
- Koenig, W., H. K. Dunn e L. Y. Lacy (1946). «The Sound Spectrograph». In: *The Journal of the Acoustical Society of America* 18.1, pp. 19–49.
- Landé, A. (1930). *Vorlesungen über Wellenmechanik*. Leipzig: Akademische Verlagsgesellschaft.
- Potter, Ralph K. (1945). «Visible Patterns of Sound». In: *Science* 102.2654, pp. 463–470.
- Showers, E. G. e R. Biddulph (1931). «Differential pitch sensitivity of the ear». In: *Journal of the Acoustical Society of America* 3, pp. 275–287.
- Wegel, R. L. e C. E. Lane (1924). «The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear». In: *Physical Review* 23, pp. 266–285.