sound
studies
forum

La regola dichiarata.
Sonificazione deterministica come contro-modello all'oculocentrismo

di Alessio Greco
Ricercatore Indipendente

Pubblicato da Sound Studies Forum, Novembre 2026

Abstract
La sonificazione di informazioni visive può essere concepita come una traduzione computazionale governata da regole esplicite, anziché come una generazione sonora affidata a processi opachi. Questo articolo esamina il *RGB Sonification Framework* (RSF), un modello deterministico che traduce i valori cromatici RGB in strutture acustiche attraverso una mappatura parametrica e una banca sonora organizzata per canale. Il Manifesto del Sound Studies Forum individua nel ribaltamento dell'egemonia oculocentrica una posta politica ed epistemica. Il contributo propone di spostare quella domanda dal lato della ricezione, l'ascolto consapevole, a quello della produzione: che cosa cambia quando le regole che trasformano un'informazione visiva in esperienza acustica sono dichiarate e ispezionabili? La tesi è circoscritta. Una mappatura deterministica non è neutrale, universalmente accessibile né intrinsecamente più etica di una generativa. La dichiarazione delle regole rende però identificabili le decisioni progettuali e crea le condizioni per ispezionarle, contestarle e modificarle. Queste proprietà sono rilevanti nella progettazione di strumenti di accessibilità, ma non sostituiscono la verifica con gli utenti né la misurazione dell'impatto computazionale. L'analisi distingue le caratteristiche documentate nella versione 1.0 di RSF dagli sviluppi previsti, e tratta l'accessibilità come problema di progettazione, non come conseguenza automatica della disponibilità di un'interfaccia sonora.

1. Chi stabilisce come l'informazione diventa percepibile?

Gran parte delle infrastrutture digitali contemporanee organizza l'informazione attraverso interfacce visive. Testi, immagini, grafici e rappresentazioni spaziali costituiscono il livello ordinario attraverso cui dati e processi diventano disponibili. Questa centralità influenza anche il modo in cui vengono concepiti l'accesso all'informazione e i criteri con cui un sistema viene considerato utilizzabile.

Il Manifesto del Sound Studies Forum nomina questa condizione in termini che eccedono l'estetica:

> «Il progressivo ribaltamento di questa egemonia oculocentrica ha riportato al centro il valore politico, epistemico e relazionale del suono, contribuendo a ridefinire i sound studies come un'ecologia di ricerche e pratiche che sostengono forme di resilienza culturale e nuovi modi di stare nel mondo» (Sound Studies Forum, 2025).

È sul termine epistemico che vorrei insistere. Riconoscere al suono un valore epistemico significa ammettere che ascoltare è una forma di conoscenza e che, come ogni forma di conoscenza, ha procedure e condizioni di validità. Significa soprattutto ammettere che quelle procedure possono essere interrogate. La maggior parte del lavoro critico che discende da questa premessa si concentra sul lato della ricezione: chi ascolta, in quali condizioni, contro quali dispositivi di cattura dell'attenzione. È il terreno su cui si muove la lettura dell'ascolto come forma eco-politica di resistenza proposta da Daniela Gentile (Gentile, 2024). Resta in ombra una seconda domanda, che riguarda l'altro capo della catena: quando il suono non è registrato ma generato, quando cioè qualcuno ha scritto la regola che lo produce, quella regola fa parte del discorso o ne sta fuori?

L'oculocentrismo tecnologico può essere inteso come la tendenza a privilegiare la visione quale modalità predefinita di accesso e interpretazione. Non implica che la visione sia sempre dominante, né che un'interfaccia visiva sia necessariamente escludente: indica una condizione progettuale nella quale altre modalità percettive vengono trattate come alternative secondarie, introdotte successivamente per esigenze specifiche. La sonificazione consente di interrogare questa gerarchia. Tradurre informazioni visive in strutture acustiche significa costruire una relazione fra due modalità sensoriali attraverso un insieme di decisioni tecniche: quali dati selezionare, quali proprietà associare ai parametri sonori, quale intervallo acustico utilizzare, quali informazioni rendere distinguibili durante l'ascolto. Nessuna di queste operazioni è automatica. Il problema non consiste quindi soltanto nel rendere udibili dati inizialmente visivi, ma nel rendere verificabile il processo attraverso cui tale trasformazione avviene.

Il *RGB Sonification Framework* (RSF) affronta questo problema mediante una mappatura deterministica dei tre canali cromatici in registri di frequenza distinti. La tesi qui sostenuta è circoscritta: la trasparenza non elimina le scelte autoriali né garantisce un risultato percettivo efficace, ma rende più semplice individuare dove tali scelte intervengono, discuterle e modificarle.

2. Sonificazione come traduzione progettata

La sonificazione viene generalmente intesa come trasformazione di dati in suono secondo una relazione definita. La definizione comprende sistemi molto diversi - rappresentazioni acustiche di misurazioni scientifiche, interfacce sonore, composizioni algoritmiche, dispositivi destinati all'accessibilità - che differiscono per obiettivi, utenti, vincoli temporali e criteri di valutazione (Hermann, Hunt, Neuhoff, 2011). Nel caso di RSF il dominio di partenza è costituito dai valori numerici dei canali RGB; il dominio di destinazione comprende parametri sonori organizzati in registri di frequenza differenti. La trasformazione non equivale alla riproduzione sonora di un colore: il colore non

possiede una frequenza acustica intrinseca che il sistema possa recuperare. La corrispondenza fra valore cromatico e parametro sonoro è una relazione progettata, la cui utilità dipende dalla struttura adottata e dalle condizioni in cui viene appresa.

Occorre pertanto distinguere almeno tre livelli:

1. **Dato:** il valore cromatico acquisito dal sistema;
2. **Regola di trasformazione:** la funzione che associa quel valore a un parametro acustico;
3. **Esperienza percettiva:** il risultato dell'ascolto da parte di una persona, in un determinato contesto.

La corrispondenza fra i primi due livelli può essere definita matematicamente; il terzo non discende automaticamente dai primi due. Una specifica tecnica può stabilire che un determinato valore RGB produca un parametro acustico preciso, ma non che quel parametro sia riconoscibile, memorizzabile o utile a chi usa il sistema. La distinzione è fondamentale per evitare che la precisione computazionale venga confusa con la validità percettiva.

3. Il modello RGB Sonification Framework

La specifica *RGB Sonification Framework for Accessible Visual Experience*, versione 1.0, descrive un sistema deterministico per la traduzione di informazioni cromatiche in strutture acustiche (Greco, 2026). L'architettura prevede l'acquisizione dei valori RGB, la separazione dei tre canali in flussi indipendenti, l'applicazione di una mappatura parametrica, la selezione di campioni da una banca sonora pre-generata e un missaggio a tre canali con uscita stereo o multicanale.

Il principio operativo è che, date le stesse condizioni iniziali e le stesse regole, la trasformazione produca lo stesso risultato. Ciascun canale è assegnato a un registro spettrale distinto e non sovrapposto, così che i tre canali restino percettivamente distinguibili nel risultato miscelato:

Canale	Registro	Intervallo di frequenza
R — Rosso	Alte frequenze	2 kHz – 16 kHz
G — Verde	Medie frequenze	500 Hz – 2 kHz
B — Blu	Basse frequenze	20 Hz – 500 Hz

Per ciascun canale, un valore *v* compreso fra 0 e 255 è mappato su una frequenza *f* all'interno dell'intervallo assegnato [*f*min, *f*max] mediante una funzione logaritmica che preserva l'uniformità percettiva:

f(*v*) = *f*min · (*f*max / *f*min) elevato a (*v* / 255)

L'ampiezza è mappata linearmente dal valore di canale al guadagno, con una correzione logaritmica per la sonorità percepita:

A(*v*) = (*v* / 255) elevato a 0,6

L'esponente 0,6 approssima la legge di potenza di Stevens per la sonorità. La distribuzione logaritmica delle frequenze è coerente con la natura logaritmica della percezione dell'altezza, ma resta una scelta progettuale per organizzare l'intervallo acustico: non dimostra, da sola, che la mappatura corrisponda in modo ottimale alla percezione di tutte le persone.

L'assegnazione R→acuto, G→medio, B→grave non è arbitraria nel senso di casuale: è coerente con le corrispondenze crossmodali documentate fra colori caldi e altezze acute, colori freddi e altezze gravi, riscontrate anche in popolazioni non sinestetiche (Marks, 1974; Cytowic, Eagleman, 2009; Spence, 2011). Resta però una decisione d'autore, presa fra alternative possibili. Ciò che la distingue da una decisione implicita è che è scritta.

3.1 Banca sonora e limiti di rappresentazione

La banca sonora costituisce l'asset acustico primario del framework ed è organizzata in tre collezioni parallele, una per canale. Ciascuna copre l'intervallo completo [0–255] alla risoluzione predefinita di 256 valori discreti, un campione per valore intero, in formato 44,1 kHz a 24 bit: 768 campioni complessivi (256 × 3). Ogni file è nominato con l'identificatore di canale e il valore a tre cifre, così che l'indirizzamento non richieda una tabella di corrispondenza. Questa descrizione non deve essere confusa con una rappresentazione sonora di tutte le combinazioni RGB possibili: lo spazio completo a 8 bit per canale comprende 256^3, ossia 16.777.216 combinazioni cromatiche. La banca rappresenta i valori dei tre canali attraverso collezioni distinte, non una registrazione indipendente di ogni combinazione. La sintesi di ciascun campione è interamente procedurale e priva di elementi stocastici: dato un valore di canale e un identificatore, l'algoritmo produce un risultato deterministico, e la banca può essere rigenerata in modo identico su piattaforme diverse. La riproducibilità dell'intero risultato dipende però anche dalle modalità con cui vengono implementate selezione, interpolazione, temporizzazione e uscita audio: il determinismo della mappatura non copre da solo l'intera catena.

RSF v1.0 va pertanto considerato, sulla base della sua specifica, come un nucleo di traduzione cromatico-acustica. Il documento indica il riconoscimento cromatico in tempo reale, la sonificazione spaziale, l'adattamento assistito da machine learning e un livello di integrazione fra i moduli successivi previsti: funzionalità che non possono essere attribuite alla versione 1.0. La delimitazione del perimetro tecnico non riduce il valore del modello, permette di valutare ciò che è effettivamente documentato senza confondere un nucleo funzionale con un sistema completo di accessibilità visiva.

4. La regola dichiarata come condizione di contestabilità

La trasparenza algoritmica viene spesso presentata come un valore generale. Per comprenderne la rilevanza nel caso della sonificazione occorre specificare che cosa diventa visibile quando le regole sono dichiarate. Una mappatura esplicita consente di individuare le decisioni progettuali: l'assegnazione di intervalli di frequenza ai canali, la scelta della funzione, la distribuzione dei valori e le condizioni di uscita. Queste decisioni possono essere discusse senza inferire il comportamento del sistema dall'ascolto del risultato: è possibile interrogare perché un canale sia associato a un determinato intervallo, se gli intervalli siano sufficientemente distinguibili, quali informazioni vengano preservate e quali trascurate.

La questione incontra qui il terreno che Steve Goodman ha mappato in *Sonic Warfare*. Il suono come arma, come controllo ambientale, come modulazione dell'affetto agisce proprio perché il soggetto non sa come agisce (Goodman, 2010). L'opacità non è un difetto accidentale di quei dispositivi: ne è la condizione di funzionamento. Un sistema di sonificazione e un dispositivo sonoro di controllo condividono la struttura formale, poiché in entrambi i casi una regola interna al meccanismo determina che cosa un corpo ascolterà. Ciò che li separa non è la bontà delle intenzioni, che non è verificabile, ma l'accessibilità della regola, che lo è.

Questa proprietà è rilevante soprattutto nei sistemi che mediano l'accesso all'informazione: quando un dispositivo traduce dati in un'altra modalità sensoriale, la mappatura contribuisce a stabilire quali differenze siano rese disponibili, quali attenuate e quali non rappresentate. La trasparenza non elimina questi problemi. Una mappatura può essere formalmente coerente e produrre risultati poco distinguibili in condizioni d'ascolto reali; può essere deterministica e imporre una curva di apprendimento eccessiva. Ciò che la dichiarazione della regola consente è formulare questi problemi con precisione e collegarli a decisioni verificabili. È in questo senso circoscritto che la regola dichiarata può costituire un contro-modello all'opacità: non un sistema privo di scelte, ma un sistema nel quale le scelte possono essere riconosciute e discusse.

5. Oculocentrismo e accessibilità multisensoriale

Il problema dell'accessibilità non consiste nell'aggiungere un canale sonoro a un'interfaccia visiva, ma nel progettare una relazione alternativa con le informazioni, tenendo conto delle condizioni percettive e delle pratiche delle persone che utilizzeranno il sistema. La sonificazione può offrire un accesso complementare ad alcune proprietà dei dati visivi. Una trasformazione acustica non equivale però alla sostituzione integrale della visione, né garantisce che ogni proprietà dell'immagine diventi riconoscibile attraverso l'ascolto. La distinzione è importante quando il sistema viene presentato in relazione alla cecità e all'ipovisione: le persone cieche e ipovedenti non costituiscono una categoria percettivamente omogenea, e familiarità con il suono, esperienza d'uso e obiettivo dell'attività modificano in modo sostanziale l'utilità di una rappresentazione.

La specifica v1.0 indica la valutazione percettiva formale con ascoltatori ciechi e ipovedenti come attività di ricerca prioritaria, da condurre in collaborazione con istituzioni dell'accessibilità e partner museali. In assenza di risultati documentati, non è possibile affermare che il framework sia già stato dimostrato efficace per una popolazione specifica o che garantisca accessibilità universale. Una verifica adeguata dovrebbe considerare la capacità di distinguere i registri, riconoscere differenze cromatiche pertinenti, apprendere le associazioni e utilizzare il sistema senza un carico attentivo sproporzionato, e dovrebbe includere le persone interessate nella definizione dei requisiti, nella prova e nella revisione. L'accessibilità deve quindi essere trattata come una proprietà da progettare, verificare e migliorare. La disponibilità di una regola trasparente offre un punto di partenza per questo processo, non il suo esito.

6. Determinismo, generatività, apertura

Il rapporto fra sistemi deterministici e sistemi generativi richiede una distinzione concettuale. La generatività riguarda la produzione di configurazioni nel tempo; il determinismo riguarda la relazione fra condizioni iniziali, regole e risultati. Le due

proprietà non sono incompatibili: un sistema generativo può essere deterministico, e un sistema che produce risultati differenti non è per questo stocastico, poiché la variabilità può dipendere dagli ingressi o da uno stato interno che evolve. La distinzione è rilevante per la pratica artistica contemporanea, nella quale la generatività costruisce installazioni e sistemi audiovisivi che si sviluppano nel tempo. Non è necessario attribuire a ogni processo generativo un'opacità intrinseca, né considerare ogni sistema deterministico eticamente preferibile. La questione decisiva non è opporre in assoluto le due proprietà, ma rendere leggibili le responsabilità progettuali: quali elementi sono fissati, quali variano, quali dipendono dagli ingressi e quali restano condizionati dall'interazione.

Una cautela analoga vale per l'apertura e per la sostenibilità computazionale. La documentazione delle regole, la disponibilità del codice e la chiarezza delle licenze sono proprietà distinte, e la verificabilità ha soglie di competenza: una regola pubblicata in un linguaggio che pochi padroneggiano è pubblica solo in senso formale, limite che riguarda ogni pratica open source e che non intendo minimizzare. Allo stesso modo, un sistema basato su funzioni deterministiche e una banca sonora finita può evitare alcune operazioni associate a processi generativi più complessi, ma la sola architettura non dimostra un minore consumo energetico in ogni contesto: servono misure riferite a configurazioni concrete e la definizione di quale sistema costituisca il termine di confronto. La computazione responsabile va intesa come criterio progettuale da verificare, non come superiorità ecologica già dimostrata.

7. Conclusione: rendere la traduzione discutibile

Il *RGB Sonification Framework* propone una relazione esplicita fra informazioni cromatiche e strutture acustiche. Il suo interesse non risiede nell'aver dimostrato che una particolare mappatura sia universalmente efficace, ma nell'aver reso dichiarabile una parte significativa delle decisioni che organizzano la trasformazione. Restano aperti almeno quattro fronti di verifica: la riproducibilità dell'intera pipeline, comprese selezione e interpolazione dei campioni; la valutazione percettiva delle associazioni con utenti reali, incluse persone cieche e ipovedenti; lo studio del tempo di apprendimento e del carico cognitivo nei contesti d'uso proposti; la misurazione ripetibile del comportamento computazionale. Questi limiti non costituiscono una confutazione del modello: definiscono il perimetro delle conclusioni che è possibile trarre dalla documentazione disponibile.

La regola dichiarata non elimina la mediazione: la rende identificabile. Non rende il sistema neutrale: permette di discutere le scelte attraverso cui viene costruito. Non garantisce accessibilità: offre un'architettura che può essere sottoposta a verifica e revisione. Un sistema opaco non può essere criticato nel merito, solo accettato o rifiutato in blocco; un sistema la cui regola è scritta può essere smontato, riscritto, usato contro le intenzioni di chi l'ha costruito. Il contro-modello all'oculocentrismo non consiste quindi nel sostituire una gerarchia sensoriale con un'altra, né nel sostenere che il suono sia intrinsecamente più inclusivo della visione. Consiste nel mettere in discussione l'idea che l'accesso all'informazione debba essere progettato a partire da un'unica modalità percettiva, e nel trattare le traduzioni multisensoriali come sistemi di decisioni esplicite. Il decentramento del visivo, se deve essere qualcosa di più di uno spostamento di gerarchia, richiede che il dispositivo che lo opera sia esso stesso esposto all'ascolto di chi vorrà verificarlo.

La domanda decisiva non è soltanto se un'immagine possa essere trasformata in suono. È quali relazioni vengano stabilite da questa trasformazione, per chi siano utili e in quali condizioni possano essere comprese, verificate e modificate.

Riferimenti bibliografici

Cytowic Richard E.; Eagleman David M., *Wednesday Is Indigo Blue. Discovering the Brain of Synesthesia*, The MIT Press, Cambridge (MA), 2009;

Goodman S., *Sonic Warfare. Sound, Affect, and the Ecology of Fear*, The MIT Press, Cambridge (MA), 2010;

Hermann Thomas; Hunt Andy; Neuhoff John G. (a cura di), *The Sonification Handbook*, Logos Verlag, Berlin, 2011;

Marks L. E., *On Associations of Light and Sound. The Mediation of Brightness, Pitch, and Loudness*, in «The American Journal of Psychology», 87 (1/2), 1974, pp. 173–188;

Spence C., *Crossmodal Correspondences. A Tutorial Review*, in «Attention, Perception & Psychophysics», 73 (4), 2011, pp. 971–995.

Sitografia

Gentile D., *L'Ascolto Come Forma Eco-Politica Di Resistenza*, Sound Studies Forum, 2024, https://soundstudiesforum.org/publication/lascolto-come-forma-eco-politica-di-resistenza (consultato il 30/10/2026);

Greco A., *RGB Sonification Framework for Accessible Visual Experience*, versione 1.0, Zenodo, 2026, https://doi.org/10.5281/zenodo.20474031 (consultato il 30/10/2026);

Sound Studies Forum, *Manifesto*, https://soundstudiesforum.org/manifesto (consultato il 30/10/2026).

Alessio Greco è sound artist e computational audio researcher, con base tra Latina e Roma. La sua pratica, interamente autonoma e priva di affiliazione istituzionale, si colloca nell'ambito della cultura aperta: sistemi deterministici e documentati, accessibilità multisensoriale e una concezione della computazione come scelta progettuale di cui dichiarare i costi. È autore del *RGB Sonification Framework* (versione 1.0, Zenodo, CC

BY 4.0), modello deterministico di traduzione cromatico-acustica sviluppato come livello fondativo per sistemi di accessibilità multisensoriale, con applicazioni previste nella mediazione culturale e museale. Tra i suoi lavori: l'installazione *Panta Rei* (Tower Festival, 2025), il progetto audiovisivo processuale *AvIS* (Venezia e Roma, 2024), la partitura per *Palumedda* di Mandragora Dance Co. e le sonorizzazioni elettroacustiche dal vivo di film muti, tra cui *Ballet Mécanique* e *Anémic Cinéma*. Porta avanti il progetto *Kintsugi Sonoro* con l’artista multidisciplinare e compagna Valentina Di Nunno