

# Musica / Tecnologia

# Music / Technology

10 • 2016

ISSN 1974-0042



# Musica/Tecnologia Music/Technology

RIVISTA DELLA FONDAZIONE EZIO FRANCESCHINI

10 • 2016

Firenze University Press

**Musica/Tecnologia**

**Music/Technology**

10 (2016)

Registrato al n. 5489 del 31/3/2006 presso il Tribunale di Firenze.

ISSN 1974-0050 (online)

Versione on-line: <http://www.fupress.com/mt>

*Direttore responsabile (Managing Director)*

Lino Leonardi

*Direttore scientifico (Research Director)*

Marco Ligabue (Conservatorio di Firenze)

*Condirettori scientifici (Research Co-directors)*

Francesco Carreras (ISTI-CNR di Pisa)

Paolo Zavagna (Conservatorio di Venezia)

*Comitato scientifico (Research Committee)*

Antonio Camurri (DIST – Università di Genova)

Pascal Decroupet (Université Nice Sophia Antipolis)

François Delalande (GRM Paris)

Giovanni De Poli (Università di Padova)

Agostino Di Scipio (Conservatorio de L'Aquila)

Ivan Fedele (Accademia Nazionale di S. Cecilia, Roma)

Marc Leman (University of Ghent)

Angelo Orcalli (Università di Udine)

Veniero Rizzardi (Conservatorio di Padova)

Curtis Roads (Media Arts and Technology, University of California, Santa Barbara)

Davide Rocchesso (IUAV, Istituto Universitario di Architettura di Venezia)

Dennis Smalley (London City University)

Marco Stroppa (Staatliche Hochschule für Musik und Darstellende Kunst, Stuttgart)

Alvise Vidolin (CSC – Università di Padova)

Con il contributo dell'Ente Cassa di Risparmio di Firenze



Con il patrocinio di:

Conservatorio di Musica "Luigi Cherubini" di Firenze



Istituto di Scienza e Tecnologie dell'Informazione CNR



Fondazione Rinascimento digitale



© 2016 Firenze University Press e Fondazione Ezio Franceschini ONLUS

Università degli Studi di Firenze – Firenze University Press – via Cittadella 7 – 50144 Firenze, Italy  
web: [www.fupress.com](http://www.fupress.com) – mail: [journals@fupress.com](mailto:journals@fupress.com)

# Indice

|   |     |
|---|-----|
| <i>Questo numero</i> di Paolo Zavagna   | 5   |
| <br>  |     |
| <i>Grani e texture sonore</i> a cura di Agostino Di Scipio e Paolo Zavagna  | 7   |
| – <i>I quanta acustici e la teoria dell’udito</i> di Dennis Gabor   | 9   |
| – <i>I quanta acustici di Gabor nelle tecnologie del suono e della musica</i><br>di Agostino Di Scipio  | 17  |
| – <i>Xenakis e il “circolo granulare”</i> di Makis Solomos  | 43  |
| – <i>Interagire con la complessità sonora interna ed esterna: dal microsuono<br/>alla composizione con suoni del paesaggio</i> di Barry Truax | 67  |
| – <i>Elaborazione di segmenti temporali</i> di Pierre Dutilleux, Giovanni De Poli,<br>Adrian von dem Knesebeck e Udo Zölzer                   | 75  |
| – <i>Stato dell’arte nella sintesi di texture sonore</i> di Diemo Schwartz  | 117 |
| <br>  |     |
| <i>Abstracts</i>  | 141 |



## *Questo numero*

Paolo Zavagna

Nel presentare l'argomento, in sé piuttosto articolato, del «paradigma granulare» nella musica elettroacustica, ci confrontiamo con diverse sfaccettature di questa odierna ‘visione’ del suono e della musica, provando a tenere insieme aspetti teorici, didattici e divulgativi. L’evoluzione storica del pensiero granulare in ambito musicale (esposta da Di Scipio nel suo articolo) ci pone in una prospettiva che non ha radici lontane nel tempo – possiamo farla risalire alla seconda metà degli anni quaranta del Novecento – ma che può avere un impatto radicale sul pensiero musicale, e non solo quello elettroacustico. «La differenza tra una concezione ‘ondulatoria’ e una concezione ‘granulare’ [...] non va certo intesa come una differenza di natura strettamente tecnica, musicalmente marginale, e riflette in realtà modi diversi di pensare il suono e dunque modi diversi di predisporsi a comporlo, a farne musica» (Di Scipio, p. 18).

Già a partire dall’articolo seminale di Gabor, del 1947, qui tradotto in italiano (pp. 9-16), è evidente come la trattazione di questo argomento induca a coniugare aspetti fisici (potremmo dire, semplificando, scientifici) e aspetti psicologici (sempre semplificando, umanistici) – circostanza che in definitiva, in questo numero di «Musica/Tecnologia», appare svilupparsi tematicamente. Inoltre, la produzione di tecnologie scaturita dalla necessità di verificare la teoria dei *quanta* acustici, già abbozzata negli studi di Gabor (si veda Di Scipio, pp. 26-32 per la parte storica e l’articolo di Diemo Schwarz per le ricadute attuali), ha dato origine a numerose ‘macchine’ capaci di alterare il suono secondo modalità alternative a quelle rese possibili dalle tecnologie derivate dal vocoder. Nel numero 7 di «Musica/Tecnologia», dedicato proprio a quest’ultimo, si nota che «cambiando il paradigma da vocoder a phase vocoder, da banco di filtri di durata ‘infinita’ a banco di oscillatori di durata ‘infinitesima’, assistiamo ad una radicale mutazione nel pensiero compositivo del musicista elettroacustico. Una mutazione molto simile – anche se parallela – a quella che avverrà in seguito alla ‘scoperta’ del pensiero di Dennis Gabor e dei quanta acustici: “the frames of the STFT are analogous to his acoustical quanta. The projection of the time-frequency plane onto the sonogram is analogous to a visual representation of the Gabor matrix”<sup>1</sup>»<sup>2</sup>. Ecco allora un altro aspetto che intendiamo qui tematizzare: la ‘contrapposizione’ tra *vocoder* e *granulatore*, fra continuo e discreto; contrapposizione da considerarsi tanto più fittizia quanto più il mondo digitale ci induce a trattare necessariamente con segnali quantizzati. Una certa ‘integrazione’ fra Fourier

<sup>1</sup>In Curtis Roads (2001). *Microsound*. Cambridge, London: The MIT Press, p. 258.

<sup>2</sup>In Paolo Zavagna (2013). “La voce senz’ anima: origine e storia del Vocoder”. In: *Musica/Tecnologia* 7, pp. 27-63. URL: <http://www.fupress.net/index.php/mt/article/view/13206>, p. 57.

e Gabor è già chiara nell'articolo di Flanagan e Golden del 1966 sul "Phase Vocoder"<sup>3</sup> nel momento in cui il segnale da analizzare deve essere sottoposto ad una 'finestratura', di fatto discretizzandolo. Pensare i due 'metodi' di analisi come separati e contrapposti non è più sensato infine quando, per campionare la Short Time Fourier Transform, utilizziamo 'finestre' di Gabor<sup>4</sup>.

Poter comporre il suono (si vedano Di Scipio, p. 17 e Truax, p. 68) ha significato per i compositori la possibilità di operare su scale di osservazione e su dominii di rappresentazione diversi, integrabili – o addirittura sostituibili – a quelli già esistenti. Ma soprattutto ha cambiato paradigma di pensiero nell'affrontare i problemi compositivi, cosiccome è cambiato il paradigma nella fisica che, dovendo affrontare aspetti microscopici della materia ha creato modelli astratti – ed esperimenti ideali – diventando "fisica teorica". Il microtempo (si vedano Di Scipio, p. 17 e Truax, pp. 67-69) diventa protagonista, sia nell'analisi sia nella sintesi del suono, aprendo alla possibilità di pensare suono e musica in termini non più necessariamente deterministici (quale/i nota/e per quanto tempo?) ma statistici (quante 'note' in un dato tempo?).

Se il modello del vocoder è un modello facilmente interpretabile dai musicisti, con il suo rimandare ad un'idea 'armonica' del suono che a lungo è apparsa come 'giustificazione scientifica' – a torto o a ragione – dell'armonia tradizionale occidentale, il modello invece dei grani sonori ha avuto, e forse ha ancora, maggiore difficoltà ad essere assimilato dai musicisti, probabilmente perché, pur agevolmente applicabile alla sintesi del suono, risulta di maggiore difficoltà nel contesto della composizione strumentale (i casi riportati da Solomos nel suo articolo alle pp. 53-56 sono interessanti proposte di cambio di sensibilità da parte dei compositori ma, come lo stesso Solomos ammette parlando di Debussy e di Bartók a p. 53, appaiono problematici).

Per questi motivi ci è sembrato opportuno proporre al lettore una raccolta di articoli che cercano di dare una visione d'insieme sul problema dei grani sonori; visione che vorrebbe coprire sia gli aspetti più scientifico/tecnologici (Gabor, Schwarz, Dutilleux et al.) sia quelli più umanistico/musicali (Di Scipio, Solomos, Truax) dando al lettore – in particolare a quello italiano, che sull'argomento non trova facilmente materiale a disposizione – la possibilità di un'ampia prospettiva sull'argomento.

Infine, è auspicabile che la diversità di punti di vista offerta con questa raccolta di articoli – mirante ad offrire strumenti didattici (Dutilleux et al.), storici (Gabor, Di Scipio, Solomos), compositivi (Solomos, Truax), epistemologici (Di Scipio, Solomos) – possa costituire un precedente di più ampie riflessioni, in futuro, circa il destino di un tema, quello del paradigma granulare, che ha inciso e sta incidendo in modo non trascurabile sulle più diverse pratiche musicali elettroacustiche – da quelle più aperte e sperimentali a quelle più legate al consumo e all'intrattenimento.

---

<sup>3</sup>Si veda J. L. Flanagan and R. M. Golden (2013). "Phase Vocoder". In: *Musica/Tecnologia* 7, pp. 9–25. URL: <http://www.fupress.net/index.php/mt/article/view/13205>.

<sup>4</sup>Si veda Gianpaolo Evangelista, Monika Dörfler, and Ewa Matusiak (2013). "Arbitrary Phase Vocoders by means of Warping". In: *Musica/Tecnologia* 7, pp. 91–118. URL: <http://www.fupress.net/index.php/mt/article/view/13210>, pp. 96-106.

# **Grani e texture sonore**

a cura di Agostino Di Scipio e Paolo Zavagna



## *I quanta acustici e la teoria dell'udito<sup>a</sup>*

Dennis Gabor

Nelle esposizioni introduttive alla meccanica ondulatoria, si ricorre spesso ad esempi di natura acustica, come ha fatto con particolare successo Landé<sup>1</sup>. In un articolo recente sulla «Teoria della comunicazione»<sup>2</sup> 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.

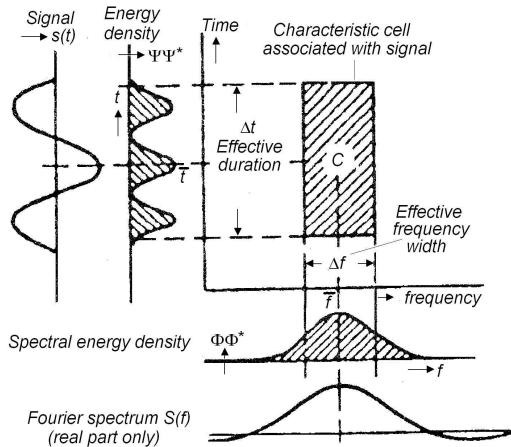
Se  $t$  e  $f$  sono messe in gioco come coordinate ortogonali, si ottiene un diagramma che può essere chiamato “diagramma informazionale” (Fig. 1). Una semplice oscillazione armonica è rappresentata da una linea verticale con ascissa  $f$ , mentre un impulso istantaneo (funzione delta) è rappresentato 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.

---

<sup>a</sup>Traduzione 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 Rocchesso per consulenza e suggerimenti nel trattamento delle notazioni matematiche adottate da Gabor.

<sup>1</sup>Landé 1930.

<sup>2</sup>Gabor 1946.



**Figura 1:** Rettangolo caratteristico di un segnale nel diagramma informazionale.

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 informazionale come centro del segnale. Procediamo definendo la “durata efficace”  $\Delta t$  e la “banda di frequenza efficace”  $\Delta f$  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  $\Delta t$  e  $\Delta f$  come  $2\sqrt{\pi}$  volte i rispettivi raggi inerziali<sup>†</sup>. Con tali definizioni si può applicare una relazione matematica, la disegualanza di Schwarz, secondo la quale

$$\Delta t \Delta f \geq 1 : \quad (1)$$

vale a dire che *l’area del rettangolo caratteristico o cella di un segnale è almeno pari all’unità*<sup>b</sup>. 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

<sup>a</sup>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.

<sup>b</sup>Possiamo dire, con terminologia più diretta,  $\Delta t = \sqrt{2} \cdot \text{durata rms}$ , e  $\Delta f = \sqrt{2} \cdot \text{frequenza di banda rms}$ .

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.

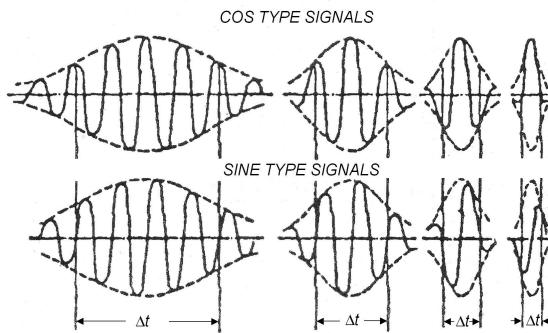
Questo risultato, in combinazione con la relazione di incertezza (1), suggerisce l’interrogativo se vi siano segnali per i quali la diseguaglianza diventi uguaglianza, e che possano essere utilizzati per la rappresentazione di quanta. Si può dimostrare che segnali semplicissimi o “elementari” di questo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

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

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  $\alpha$ , 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.



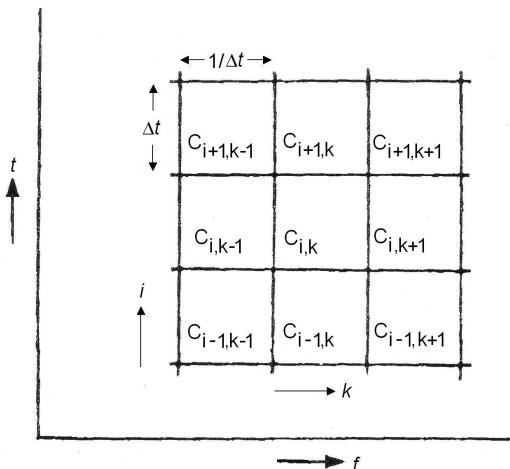
**Figura 2:** Segnali elementari.

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  $\alpha$ . Vale a dire, possiamo espandere un segnale dividendo

---

<sup>c</sup>Le 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:  $e^{-(\pi/\alpha)^2(f-f_0)^2} e^{i2\pi f_0 t}$ .

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 sconveniente, 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 pubblicazione<sup>3</sup>. 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 Lichte<sup>4</sup>. Il loro oggetto di studio era la soglia di discriminazione temporale in differenti registri di frequenza.

<sup>3</sup>Koenig, Dunn e Lacy 1946; Potter 1945.

<sup>4</sup>Bürck, Kotowsky e Lichte 1935a,b.

Mach aveva dimostrato che se un'oscillazione sinoidale 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<sup>d</sup>. 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 una 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.

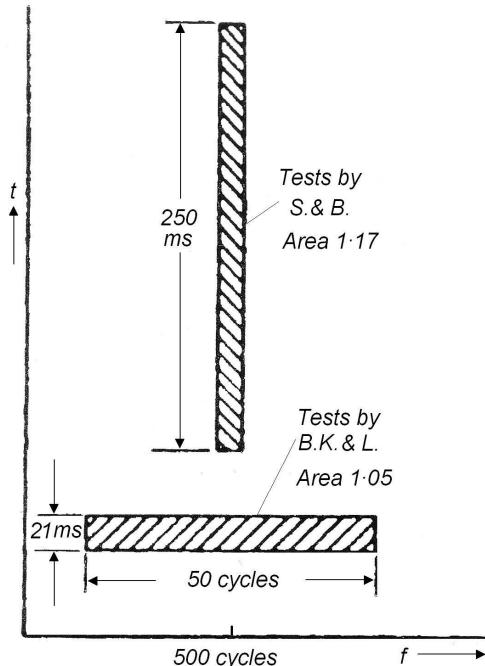
Un'altra serie di esperimenti è quella di Shower e Biddulph<sup>5</sup>, e riguarda la soglia di discriminazione frequenziale dell'orecchio. Il loro metodo sperimentale era del tutto diverso. La frequenza di una nota veniva modulata entro limiti ridotti, e veniva annotata la minima deviazione che desse 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 sinoidale 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.

<sup>d</sup>Si 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.

<sup>5</sup>Shower 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,

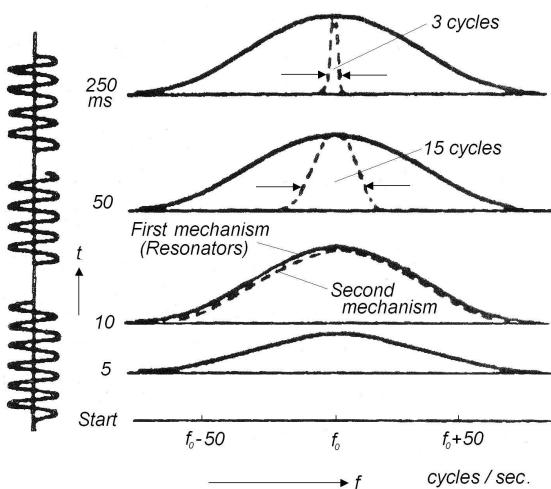


**Figura 4:** Aree di soglia dell'orecchio a 500 Hz.

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 Lane<sup>6</sup> 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.

<sup>6</sup>Wegel 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 accettare 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 Davis<sup>7</sup>.

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 insegni 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 segue 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

<sup>7</sup>Galambos 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: Academische Verlagsgesellschaft.
- Potter, Ralph K. (1945). «Visible Patterns of Sound». In: *Science* 102.2654, pp. 463–470.
- Shower, 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.

# *I quanta acustici di Gabor nelle tecnologie del suono e della musica*

Agostino Di Scipio\*

(Ricevuto il 10 novembre 2015. Versione rivista ricevuta il 2 dicembre 2015)

## Sommario

In una serie di contributi scientifici della fine degli anni quaranta del Novecento, Dennis Gabor sviluppò un quadro concettuale e operativo di analisi del segnale sonoro che, innestandosi in una tradizione scientifica allora piuttosto recente, apriva orizzonti di ricerca e applicazione musicale al tempo del tutto inediti. L'articolo segue appunto il formarsi di tale quadro di riferimenti scientifici e delle relative ripercussioni pratiche – particolarmente a partire da due contributi di primaria importanza (“Theory of communication” del 1946, e “Acoustical quanta and the theory of hearing”, del 1947). Dopo alcune osservazioni preliminari, l'articolo approfondisce alcuni elementi di senso “genealogico” (teoria dei quanti, indeterminazione di Heisenberg, analisi delle sensazioni di Mach) e contestualizza poi la ricerca di Gabor rispetto ad interessi scientifici più o meno coevi ad esso e ad esso coerenti. Seguendo un percorso ibrido tra storia e critica della scienza e delle tecnologie, l'articolo descrive poi alcune delle sperimentazioni e dei dispositivi tecnologici, tra quelli più direttamente ricollegabili ai contributi di Gabor, che hanno avuto maggiori ripercussioni in ambito scientifico e, soprattutto, nell'ambito di nuovi percorsi di teoria della musica e degli sviluppi storici – certo pionieristici, ma in un certo senso fondativi – delle pratiche compositive legate alle tecnologie analogiche e digitali del suono.

## 1 Introduzione

In queste pagine riprendo un testo apparso ormai parecchi anni fa (Di Scipio 1998), apportandovi necessarie revisioni e aggiornamenti. Ho conservato l'impianto complessivo della versione iniziale, ma ho riscritto molti passaggi, inserito titoli di sezione e sottosezione, e aggiornato la bibliografia laddove opportuno. Ho infine aggiunto alcune interpolazioni utili a rendere più chiara l'esposizione e meno incompleta la trattazione dei temi toccati.

### 1.1 (S)composizione del suono e dominii di rappresentazione del segnale

Uno degli elementi più caratteristici e trasversalmente condivisi nelle molteplici esperienze di musica elettroacustica e sperimentale, e nelle moderne tecnologie del suono in generale, è certamente l'idea di poter ‘comporre il suono’ facendone dimensione autonoma rispetto alla più usuale dimensione del ‘comporre con i suoni’, alla quale naturalmente si affianca e si interseca. Nel quadro della cultura musicale contemporanea, questa possibilità ha dischiuso un enorme potenziale linguistico-musicale e si è tradotta in una grande varietà di prassi estetiche.

---

\*Conservatorio di Musica “A. Casella” – L’Aquila.

Ciò è avvenuto mediante la previa riduzione del *suono a segnale* implicitamente assunta nella grande maggioranza delle tecnologie disponibili (un ‘riduzionismo’ per nulla ovvio o aproblematico, benché accettato maggioritariamente nella cultura contemporanea del suono). Si pone quindi la questione dei modelli di *rappresentazione del segnale* sulla cui base è possibile sviluppare operazioni di analisi, sintesi ed elaborazione del suono. Da un certo punto di vista, le soluzioni tecnologiche praticate nel corso dei decenni possono essere ricondotte a due prospettive essenziali: da una parte, la possibilità di studiare e progettare lo spettro del suono, anche nella sua morfologia dinamica (si tratta quindi di operare sulla base di rappresentazioni del suono *nel dominio della frequenza*, lasciando implicita o derivata la determinazione del tempo); dall’altra, la possibilità di studiare e gestire direttamente le curve di oscillazione, ovvero l’andamento microtemporale del segnale (si tratta allora di operare *nel dominio del tempo*, lasciando implicita e derivata la determinazione della frequenza). Nei concreti modelli operativi costruiti a partire da queste due strade, le diverse possibilità di relazionarsi alla dimensione temporale implicano diverse conseguenze pratiche e concettuali, anche a livello di pensiero e di prassi musicale.

Si parla talvolta di ‘microcomposizione’ alludendo alla possibilità di scomporre e ricomporre il suono e più estese strutture musicali a partire da particelle di durata finita, a partire cioè da ‘atomi’ o ‘grani’ sonori. In campo musicale, il primo a muoversi in tale prospettiva è stato Iannis Xenakis<sup>1</sup>, in brani davvero pionieristici realizzati alla fine degli anni cinquanta del Novecento presso il GRM di Parigi. Il suo tentativo di lavorare con *quanta sonori* (termine adoperato dal compositore stesso) inaugurava un’operatività piuttosto differente rispetto alle prospettive più comuni, nelle quali il suono è concepito come somma di oscillazioni sinusoidali periodiche o aperiodiche, alla stregua di quanto avveniva per l’acustica classica, ma anche nei primi esempi di *elektronische Musik* realizzati alla WDR di Colonia e nella progettazione di strumenti musicali elettronici inventati nel corso della prima metà del secolo, a partire dal *telharmonium* di Thaddeus Cahill (risalente agli ultimi anni del XIX secolo). La differenza tra una concezione ‘ondulatoria’ e una concezione ‘granulare’ – comunque non esclusiva della poetica di Xenakis, dato che esempi diversi se ne trovano anche in altre esperienze pionieristiche di musica elettronica (Di Scipio 1994) – non va certo intesa come una differenza di natura strettamente tecnica, musicalmente marginale, e riflette in realtà modi diversi di pensare il suono e dunque modi diversi di predisporsi a comporlo, a farne musica.

Queste circostanze possono essere ricondotte a corrispondenti momenti della storia della scienza: da una parte abbiamo allora l’acustica classica, che vede nel fenomeno sonoro una meccanica ondulatoria e che trova nella ‘teoria analitica del calore’ di Jean-Baptiste Fourier (1822) adeguata rappresentazione matematica, perfettamente funzionale ai successivi contributi di George Ohm (1843) e di Hermann von Helmholtz (1885), a loro volta decisivi per lo studio dei fenomeni della percezione uditiva e per la teoria del suono musicale; dall’altra una visione di tipo ‘atomistico’ o ‘corpuscolare’, la quale – pur potendosi inscrivere in un’antica tradizione che annovera, a vario titolo, nomi di illustri scienziati del XVII secolo (come Gassendi e Beekman)<sup>2</sup> – trova adeguati mezzi di rappresentazione matematica solo nel corso del XX secolo, particolarmente nelle ricerche portate avanti negli anni quaranta da Dennis Gabor (1900-1979). In tempi moderni, questa dualità di prospettive non è posta come una netta divisione tra scuole di pensiero contrastanti: piuttosto, come osserveremo proprio a partire

<sup>1</sup>Cfr. i primi capitoli del suo *Musiques Formelles* (1963), anche nella traduzione inglese (1992). Si veda anche Xenakis (1985, pp. 178 e sgg.).

<sup>2</sup>Si veda il breve riepilogo storico in Roads (2001a, pp. 50-55).

da Gabor, le due idealizzazioni alternative di segnale (variazione di ampiezza in funzione del tempo) e di spettro (analisi di Fourier) possono essere ricondotte all'interno di un formalismo unitario, diventando casi particolari ma diametralmente opposti dentro uno spazio di fenomeni omogeneo.

In sintonia con altri ricercatori, Gabor muove dalla constatazione dello statuto teorico, e astratto dal tempo, dell'analisi di Fourier, arrivando infine a proporre, sullo sfondo della fisica quantistica e della teoria dell'informazione della sua epoca, un quadro concettuale nuovo, ricco di implicazioni conoscitive e tecnologiche. Cercheremo di darne un profilo complessivo, mettendo in evidenza vari aspetti tecnici e teorici, e le principali implicazioni musicali. Sarà importante valutare bene come l'impianto formale del lavoro di Gabor si ponga sin dall'inizio in connessione diretta con lo studio della percezione uditiva. Già il titolo del suo articolo scientifico più noto, «Acoustical quanta and the theory of hearing» (Gabor 1947a), indica la volontà di affrontare simultaneamente due ambiti d'interesse solitamente ardui da trattare all'interno di una visione unitaria, quello dei fenomeni fisico-acustici e quello della percezione uditiva, entrambi visti da una certa angolazione scientifica con la sua specifica dimensione epistemologica. Più in generale, nel contesto della proposta di Gabor, «le due parole *teoria* e *pratica* si incontrano» permettendo, come vedremo, non solo che «matematici, fisici e ingegneri [possano] comunicare tra loro in vista di sviluppi fruttuosi» (Feichtinger e Strohmer 1998, p. 1), ma anche che compositori e musicisti possano trarne importanti conseguenze concettuali ed operative nel loro lavoro sul suono.

## 1.2 Dennis Gabor

Nato a Budapest nel 1900, Denés Gabor studiò nella città natale, iniziò poi a lavorare a Berlino dove rimase finché nel 1934 (con la definitiva conquista del potere da parte di Hitler) si trasferì in Gran Bretagna. Morì a Londra nel 1979. Attivo in numerosi campi di ricerca, egli si occupò di teoria dell'informazione, di microscopia elettronica, di oscilloscopia, e di molto altro ancora. Il suo nome è di solito associato allo sviluppo dell'olografia, al quale iniziò a dedicarsi verso la fine degli anni quaranta (per i suoi studi in questo campo, fu Premio Nobel per la fisica nel 1971)<sup>3</sup>.

Tra il 1944 e il 1945, Gabor è attivo presso i laboratori della British Thomson-Houston, a Rugby, e prepara una memoria di ricerca molto articolata, apparsa nel 1946 col titolo «Theory of communication» (Gabor 1946)<sup>4</sup>. All'interno di un ampio quadro di riferimenti a studi di teoria dell'informazione e della comunicazione, Gabor propone un modello di analisi che procede scomponendo il segnale in *quanta* di informazione. La proposta viene ripresa in modo più snello e diretto in «Acoustical quanta and the theory of hearing» (Gabor 1947a), dove se ne rimarcano le implicazioni per la rappresentazione di segnali sonori (vocali e musicali), anche in rapporto a certi aspetti particolari del funzionamento dell'apparato uditivo, ma tralasciando quasi del tutto gli aspetti più tecnologici presenti nel primo articolo<sup>5</sup>. Nel seguito,

<sup>3</sup>Si veda Gabor (1979). Il primo contributo nella direzione che porterà Gabor ai principi di olografia è in Gabor (1948a).

<sup>4</sup>Questo lungo scritto è diviso in tre parti (“The analysis of information”, “The analysis of hearing” e “Frequency compression and expansion”), ciascuna avente un proprio apparato di Appendici e di riferimenti bibliografici. Un riepilogo estremamente ridotto apparve l'anno successivo sulla stessa rivista, con lo stesso titolo (Gabor 1947c).

<sup>5</sup>Pur presentandosi come un testo unitario, privo di ripartizioni interne, «Acoustical quanta...» è chiaramente diviso in due parti, grosso modo corrispondenti alle prime due parti di «Theory of communication», di cui può essere considerato una sorta di versione agile, adeguata alle finalità di comunicazione e divulgazione scientifica della rivista che ne accoglieva la pubblicazione, *Nature*.

ci soffermeremo sui contenuti di questi due contributi, pur tenendo presente alcuni altri scritti di Gabor su questi temi apparsi negli anni subito successivi, come «*New possibilities in speech transmission*» (Gabor 1947b) e *Lectures on communication theory* (Gabor 1952)<sup>6</sup>. Vanno anche segnalati due articoli pubblicati in Francia, uno dei quali apparso in un volume collettivo curato dal celebre fisico Louis de Broglie nel contesto del dibattito fisico-matematico sul dualismo ‘onda-particella’ (Gabor 1951), l’altro apparso su *L’Onde électrique*, una rivista destinata a professionisti della radiofonia (Gabor 1948b)<sup>7</sup>.

## 2 Quanta acustici e teoria dell’informazione

### 2.1 Tempo e frequenza

Gabor (1947a, p. 591) pone una domanda molto diretta: «Che cosa udiamo?». L’interrogativo in realtà presenta una sua intrinseca duplicità semantica. In senso ‘fisicalista’ vuol dire: cosa desta in noi le sensazioni uditive? In senso ‘fenomenista’ può significare: cosa sono le sensazioni uditive stesse? Evidentemente i due punti di vista vanno collegati: la rappresentazione matematica del suono deve tener conto sin dall’inizio dell’esperienza uditive, dove tempo e frequenza sono inevitabilmente compresenti e interrelati. All’udito, un suono si presenta come una configurazione energetica corrispondente a determinate componenti spettrali di data frequenza e ampiezza, ma anche dotata di un’evoluzione temporale che consiste proprio nel dinamico mutare dei rapporti di frequenza (e quindi di fase) tra le componenti. L’analisi di Fourier, «inattaccabile» [*beyond reproach*] sul piano matematico (Gabor 1946, p. 431), tuttavia conduce ad una descrizione atemporale in termini di onde esattamente periodiche, cioè di durata infinita. L’esperienza rivela invece un decorso temporale nel suono, una condizione di finitezza della durata e di aperiodicità delle oscillazioni. Gabor scrive:

il metodo basato sull’integrale di Fourier considera i fenomeni in un intervallo infinito, *sub specie aeternitatis*, quanto di più lontano dal punto di vista dell’esperienza di tutti i giorni. Se il termine ‘frequenza’ viene usato nel senso strettamente matematico che si applica solo a treni d’onda infiniti, una ‘frequenza mutevole’ [*changing frequency*] diventa una contraddizione in termini, perché sarebbe una definizione che riguarda sia il tempo che la frequenza (*ibid.*, p. 431)<sup>8</sup>.

Ma «tempo e frequenza sono complementari, piuttosto che mutualmente esclusivi» (*ibid.*, p. 431), e dovrebbero quindi essere espressi in maniera esplicita in una rappresentazione del segnale che dia conto di entrambi, pur nei rispettivi ruoli «simmetrici» (*ibid.*, p. 429). L’obiettivo di ricerca perseguito è appunto quello di dimostrare l’esistenza di «metodi matematici adeguati a questo scopo» (*ibid.*, p. 431). Si tratterà di metodi che avranno necessariamente la «strana caratteristica» (*ibid.*, p. 432) di non poter avere un grado molto elevato di precisione

<sup>6</sup>La datazione di questo testo, che riprende i contenuti di interventi fatti al MIT di Boston, talvolta è parsa incerta. Stando a Léon Brillouin, pioniere della teoria dell’informazione il cui lavoro viene citato da Gabor proprio in questo testo, la data sarebbe quella del 1951 (Brillouin 1956, pp. 98, 168). Stessa data si trova in un contributo del grande linguista Roman Jakobson (1966, p. 32). Sembra plausibile che Brillouin e Jakobson si riferissero in realtà all’anno in cui le *lectures* furono effettivamente tenute, non alla pubblicazione dei testi corrispondenti.

<sup>7</sup>Qualche anno dopo, questa rivista, pubblicata a Parigi dalla Société des Radioélectriciens, ospitò scritti di Jacques Poullin, primo collaboratore di Pierre Schaeffer al GRM di Parigi (Poullin 1954) e di Abraham Moles, ricercatore e collaboratore di Schaeffer sin dai primi anni cinquanta (Moles 1953a,b). La circostanza induce ad ipotizzare che Gabor (1948b) possa essere stato letto nell’ambiente del GRM.

<sup>8</sup>Qui e altrove si intenda ‘traduzione mia’.

*simultaneamente* per la determinazione del tempo e per la determinazione della frequenza – problematica non a caso affrontata con gli strumenti concettuali della meccanica quantistica e richiamandosi al ‘principio di indeterminazione’ che Werner Heisenberg aveva identificato due decenni prima (Heisenberg 1927).

Inoltre, mentre i metodi basati sul modello di Fourier implicano il *continuum* del tempo e della frequenza, Gabor considera che la capacità dell’apparato uditivo ha invece limiti di risoluzione o di ‘discriminazione’ tra valori contigui in entrambe le dimensioni, limiti che sono accettabili e che non possono essere disattesi. Come vedremo, per identificare le unità minime di informazione rilevanti per l’orecchio, il suo discorso rinvia a ricerche di psicofisiologia dell’udito condotte nei decenni precedenti a proposito delle soglie di discriminazione uditiva (particolarmente quelle relative al tempo).

## 2.2 Diagramma informazionale e segnali elementari

Gabor costruisce un diagramma informazionale [*diagram of information*], cioè una griglia o matrice bidimensionale, discretizzando sia il continuo temporale (asse orizzontale) sia quello frequenziale (asse verticale). Ciascuna delle celle di tale matrice è assimilabile alla risposta caratteristica di un risonatore acustico, secondo l’analogia meccanica classica che vede nell’analisi spettrale la risposta di molteplici risonatori intonati (corrispondenti ad un banco di filtri). Nel caso della matrice di Gabor, idealmente ciascun risonatore ha larghezza di banda finita ( $\Delta f$ ) e risposta nel tempo finita assai breve ( $\Delta t$ ). Ogni cella vale allora come *quantum* di informazione geometricamente definito da un passo differenziale di frequenza e di tempo. L’area dovrà essere pari almeno all’unità:  $\Delta f \Delta t \geq 1$ . «Questa è l’esatta formulazione della relazione di indeterminazione» (Gabor 1947a, p. 591; qui p. 10). Inizialmente, in Gabor (1946), l’area è messa in rapporto all’identità statistica  $\Delta t \Delta f \approx 1$ . Il senso è che ciascuna cella, prodotto di un’incertezza quantistica, è centrata su un valor medio di frequenza e su un valor medio di tempo: al crescere di  $\Delta t$  dovrà decrescere  $\Delta f$  e viceversa, rimanendo mediamente pari all’unità il loro prodotto<sup>9</sup>. Attraverso un’analisi matematica della questione, Gabor infine stabilisce nella diseguaglianza  $\Delta t \Delta f \geq 1/2$  la relazione di incertezza minima attribuibile ad ogni cella nel diagramma informazionale, che d’altra parte sta «alla radice del principio fondamentale delle comunicazioni» (ibid., p. 434)<sup>10</sup>.

Cosa contiene, in concreto, l’area  $\Delta t \Delta f$  di ciascuna cella del diagramma informazionale? Essa va pensata come un ‘segnale elementare’ [*elementary signal*] che possa far sì che «la diseguaglianza diventi uguaglianza» (Gabor 1947a, p. 591; qui p. 11) e che  $\Delta t \Delta f$  assuma «il valore più piccolo possibile» (Gabor 1946, p. 435). Gabor dimostra che ciò si può ottenere con segnali che sono «il prodotto di modulazione di un’oscillazione armonica di qualsiasi frequenza con un impulso avente la forma di una funzione di probabilità» (ibid., p. 435). In altre parole, a ciascuna cella del diagramma va associato un segnale costituito da una parte reale di tipo seno o coseno, e da una parte immaginaria corrispondente ad una curva di tipo gaussiano (la quale garantisce minima dispersione spettrale, cioè minima deviazione rispetto alla relazione

<sup>9</sup>In un passaggio ulteriore (Gabor 1946, p. 442), leggiamo che «G. W. Stewart nel 1931 fu il primo a chiedersi se i limiti della sensazione uditiva [*aural sensation*] non seguano una relazione di incertezza, che egli scrisse come  $\Delta t \Delta f = 1 \dots$ ». Possiamo forse considerare l’articolo di Stewart il primo esempio di un avvicinamento a problematiche di acustica in termini quantistici.

<sup>10</sup>Gabor non chiarisce l’osservazione. Va detto che nel suo modo di argomentare riecheggiano ripetutamente i principi del trasferimento di informazione che erano stati illustrati in contributi fondamentali di Harry Nyquist e di Ralph Hartley negli anni venti (riportati da Gabor in bibliografia), che poi trovarono proprio tra il 1948 e il 1949 nel teorema di Shannon e Weaver una formulazione efficace per l’ingegneria dell’informazione (Shannon e Weaver 1949).

tempo/frequenza)<sup>11</sup>. A quel punto, determinando l'ampiezza del segnale elementare, ogni cella del diagramma informazionale costituisce un quantum del segnale acustico, cioè individua *la quantità di energia dispersa all'interno dell'unità minima di frequenza durante l'unità minima di durata*. Gabor procede a dimostrare matematicamente che qualsiasi segnale può essere espanso in termini di queste unità elementari, per cui il diagramma informazionale sarà la rappresentazione del segnale scomposto in *quanta acustici* trattabili come *quanta di informazione* (Gabor 1946, p. 435).

Abbiamo allora un modello di analisi del suono che «parla sia il ‘linguaggio del tempo’ sia il ‘linguaggio della frequenza’...» (Gabor 1947a, p. 592; qui p. 12). Come anticipato sopra, tale modello presenta due casi particolari ed opposti: celle infinitamente estese nella direzione del tempo equivarranno ad oscillazioni sinusoidali periodiche (analisi di Fourier); celle infinitamente estese nella direzione della frequenza equivarranno a ‘funzioni delta’, cioè impulsi di Dirac (durata infinitesima). Insomma *spettro* (variazioni di ampiezza nel dominio della frequenza) e *segnaile* (variazione di ampiezza nel dominio del tempo, discretizzato su campioni così prossimi tra loro da pervenire al continuo) «sono entrambi casi speciali del modello di Gabor» (Brillouin 1956, p. 99). In questo risultato riverbera un tema importante nel dibattito scientifico dell’epoca, e fondamentale per tutta la conoscenza contemporanea, quello della dualità tra particella e onda. Nel modello avanzato da Louis de Broglie una ventina di anni prima, le particelle di materia che si muovono a velocità identica rivelano, ad una certa scala di osservazione, le proprietà di un’onda, e possono allora essere ricondotte alle leggi della meccanica ondulatoria; le stesse particelle, guardate ad un’altra scala di osservazione, possono essere descritte nei rispettivi moti individuali, il che richiede un passaggio alla meccanica quantistica.

Il modello teorico di Gabor si presenta come la prima rappresentazione formale (in senso matematico) nella storia delle teorie corpuscolari del suono. Nella storia dell’acustica (Lindsay 1966), a partire dall’atomismo greco tramandatoci da Lucrezio, e passando per il corpuscolarismo seicentesco, si è spesso ritenuto che la concezione corpuscolare abbia prodotto poco più che semplici intuizioni meccanicistiche. Con Gabor si ha un vero modello teorico e analitico (che poi risulterà preliminare ad ulteriori modelli dello stesso tipo, ma più sofisticati). In tal senso, sul piano storico, il suo ruolo appare simile al ruolo svolto dal modello di Fourier un secolo prima rispetto a precedenti intuizioni circa la natura composita del moto oscillatorio, cioè rispetto all’idea per cui ogni movimento, quindi anche ogni moto oscillatorio complesso, è scomponibile in una serie di oscillazioni semplici con frequenze in rapporto armonico tra loro (intuizione che si trova già nel lavoro di Joseph Saveur, all’inizio del XVIII secolo, e comprovata sperimentalmente da Daniel Bernoulli e da Eulero intorno alla metà di quello stesso secolo). Entrambi i contributi danno spessore teorico ad un bagaglio di intuizioni e di osservazioni empiriche precedenti.

### 2.3 L’analogia meccanica e l’orecchio

Riprendiamo l’analogia meccanica posta inizialmente in relazione all’area unitaria del diagramma informazionale. Ciascuna cella, si è detto, può essere pensata come risposta nel

<sup>11</sup>In termini più consueti per la prassi musicale elettroacustica, ogni cella corrisponde ad un ‘grano sonoro’: la parte reale del segnale è la sua ‘forma d’onda’, la parte immaginaria è il suo ‘inviluppo’. In un passaggio successivo di Gabor (1946), viene valutata la possibilità di utilizzare come parte immaginaria (indicata ora come ‘finestra’ del segnale reale) forme triangolari o trapezoidali, che viene però scartata perché comporterebbe inutili complicazioni teorico-matematiche senza portare a «risultati differenti rispetto a finestre probabilistiche» (*ibid.*, p. 449).

tempo (o «tempo di decadimento», come dice Gabor) di un risonatore tarato ad una certa frequenza. L'analogia permette di pensare tutta la problematica anche in relazione all'idea che descrive l'orecchio in termini di un sistema di risonatori.

La questione ha implicazioni importanti naturalmente in rapporto all'individuazione dell'area della cella,  $\Delta t \Delta f$ , essendo «praticamente accertato che i risonatori dell'orecchio sono ampiamente smorzati, con tempi di decadimento dell'ordine di 10 millisecondi o anche meno» (Gabor 1947a, p. 593; qui p. 14). Più precisamente, per Gabor lo studio delle soglie di discriminazione uditiva di tempo e di frequenza mostra che «*l'orecchio possiede un'area di soglia di discriminazione di ordine unitario*» (ibid., p. 592; qui p. 12), «almeno nell'ambito di frequenze sotto i 1000 Hz e per durate comprese tra 20 e 250 millisecondi» (Gabor 1946, p. 444). Pur con questi importanti limiti, Gabor può allora collegare il suo modello di analisi all'analisi delle sensazioni uditive, rivelandone il valore empirico. Il diagramma informazionale insomma agisce da schema concettuale del modello formale, ma allo stesso tempo vale come schema che lega il modello analitico all'esperienza uditiva.

C'è però un'altra osservazione importante a proposito della metafora meccanica qui richiamata: i risonatori tradizionalmente usati nell'analisi del suono<sup>12</sup> sono tarati su bande di frequenza fisse, con risposte nel tempo proporzionali, non liberamente determinabili. Le variazioni di periodicità e ampiezza nel corso della durata di un suono, importanti per la percezione uditiva, non sono rilevabili con lo stesso grado di precisione da meccanismi del genere. La consuetudine di pensare l'orecchio come un banco di risonatori rinvia implicitamente ad un'orecchio piuttosto astratto (in modo speculare al modello di Fourier, che rinvia ad uno spettro di frequenza, astraendo dal tempo). In base ad esperimenti attendibili, l'orecchio sembra invece avere un comportamento più sottile, capace peraltro di cogliere con progressiva sottigliezza l'altezza di un suono al crescere della durata. Il che suggerisce che, come sistema di risonatori, esso evidentemente agisce in modo dinamico. Gabor giunge a postulare l'esistenza, nel sistema uditivo, di una «seconda curva di risonanza» (ibid., p. 444), di un «secondo meccanismo nell'orecchio» (Gabor 1947a, p. 593; qui p. 14) di tipo non fisiologico ma (ipoteticamente) relativo alle fibre nervose dell'orecchio interno, capace di attivarsi durante lo stimolo uditivo permettendo in qualche modo ai «risonatori dell'orecchio» di assottigliare progressivamente la banda di frequenza per determinare con crescente precisione la periodicità dello stimolo sonoro (ciò almeno in un ambito di frequenze medio-basse, sotto i 1000 Hz). Come suggerisce Gabor, un simile comportamento dell'orecchio è effettivamente in gioco non tanto nell'ascolto del parlato, quanto nell'ascolto della musica, dove eventi sonori (quasi) periodici hanno mediamente durata maggiore dei suoni del parlato (Gabor 1946, p. 444).

## 2.4 Mach, empirismo, interdipendenza dei fenomeni

A monte di queste tematiche di psicofisiologia dell'udito, in Gabor (1947a) c'è un fugace ma significativo riferimento ad Ernst Mach (1838-1916). Già nel 1865 (a Graz, prima di stabilirsi a Praga dove diventò una delle massime autorità scientifiche del suo tempo), il grande fisico austriaco aveva condotto esperimenti di acustica e di ottica, alcuni dei quali centrati sulla percezione dell'altezza in rapporto alla durata delle oscillazioni e sulle soglie differenziali di percezione della durata (soglie che gli apparvero proporzionali all'estensione di durata considerata, ma in termini non esattamente logaritmici, devianti rispetto alla legge di Weber e Fechner). Tali ricerche si inquadravano in un più vasto interesse che Mach nutriva per il

<sup>12</sup>Come quelli descritti nella parte iniziale del trattato di Helmholtz (1863, pp. 66-68).

«senso del tempo» (*Zeitsinn*) di cui è capace l'udito umano (Mach 1865b)<sup>13</sup>. Pur di passaggio, Gabor si richiama a Mach in modo esplicito attribuendogli la verifica del fatto che «una breve oscillazione sinusoidale sarà percepita come un rumore, e solo oltre una certa durata minima diventa un tono di altezza accertabile» (Gabor 1946, p. 442). La soglia fissata da Mach era di 40 millisecondi. Al di sotto di essa non si avrebbe percezione della durata ma, com'è stato scritto recentemente, di «punti di tempo» (Pöppel, Bao e Zhou 2011, p. 776). Gabor segnala comunque che, in studi degli anni trenta, questa soglia risultava anche più bassa, assestandosi a circa 10 millisecondi almeno nell'ambito di frequenze tra 500 e 1000 Hz. Per intuire il valore di tali misure, si consideri che al giorno d'oggi le neuroscienze indicano in circa 50 millisecondi il tempo complessivo necessario all'elaborazione cerebrale di ciascuno degli stimoli uditi in una sequenza di eventi impulsivi (Chouard 2001, p. 103), il che non a caso coincide col limite inferiore di udibilità in frequenza.

Nel contesto di una prospettiva di ricerca come quella perseguita da Gabor, il nome di Mach può avere un significato particolare. Il fisico era stato apertamente polemico nei confronti dell'atomismo del suo tempo, quale sviluppatosi a partire dalla chimica di vari decenni prima, cioè dell'inizio del XIX secolo. Nel suo empirismo radicale, egli considerava gli atomi della scienza moderna come un residuo neokantiano, una sorta di 'noumeno' inaccertabile sul piano dell'esperienza sensibile, la cui idealizzazione poteva essere tollerata solo come male utile e necessario ad ulteriori avanzamenti scientifici. In tal senso egli ebbe una posizione 'strumentalista' nei confronti dell'atomismo (Oldroyd 1989, pp. 233-234)<sup>14</sup>. Rievocare il nome di Ernst Mach, può voler dire insomma evocare una tradizione epistemologica per la quale la scienza non si limita a dare rappresentazioni formali, ma appunto si sforza di introdurre anche un legame sensibile tra rappresentazione ed esperienza, interpretando in senso intuitivo ed empirico le osservazioni e i dati raccolti.

In fondo, un atteggiamento del genere rispecchia lo spirito che condusse, proprio nel cuore della scienza quantistica, al principio di indeterminazione di Heisenberg, facendo del punto di vista dell'osservatore una variabile ineliminabile e decisiva delle possibilità di conoscenza. Ciò apriva quindi al pensiero di una certa relatività o almeno di una certa mutua relazione nei fenomeni osservati. Lo stesso Mach, ormai anziano, ebbe occasione di indicare come nelle ricerche condotte da giovane su durata e temporalità delle sensazioni uditive, si avesse una qualche evidenza di una più generale «interdipendenza dei fenomeni» (Mach 1893, pp. 541, 544), da cui derivare un modo di fare scienza orientato al «nesso tra il fatto osservato psicologicamente e il processo fisico (fisiologico) corrispondente» (Mach 1914, p. 59)<sup>15</sup>.

<sup>13</sup>Nel contesto di questi studi, Mach si interessò anche al «senso dello spazio» (*Raumsinn*) di cui è capace l'apparato uditivo (Mach 1886), ed intuì forme di adattamento dinamico nel funzionamento dell'apparato uditivo (Mach 1865a). Infine, qualche tempo dopo, egli identificò nel labirinto dell'orecchio medio l'organo 'statocinetico' che rende partecipe l'orecchio della «sensazione del movimento» (Mach 1875). Una ricostruzione complessiva di questi esperimenti di Mach si trova in Debru (2003).

<sup>14</sup>«Mach accetta la teoria atomistica come un male necessario fino a quando il processo scientifico non riuscirà a offrire una migliore sistemazione unitaria dei fenomeni naturali» (Sigismondi 2002, p. 11). Questa posizione critica finì in realtà col fare da stimolo per una generazione di scienziati più giovani, favorendo una trasformazione dell'atomismo e contribuendo dunque costruttivamente alle teorie quantistiche successive. Stando alla ricostruzione storica di Paul Feyerabend, alcuni dei primi lavori di Albert Einstein (nei primi anni del XX secolo), riprendono le obiezioni di Mach allo scopo di sollevarle in maniera produttiva: la fisica che ne emerse introduceva nell'indagine scientifica un punto di vista che Mach avrebbe reputato consono alla sua idea di 'scienza naturale'; cfr. il capitolo intitolato "Mach's theory of research and its relation to Einstein" in Feyerabend (1987).

<sup>15</sup>Per un approfondimento di questi temi nell'opera di Ernst Mach, si vedano Bozzi (1990) e D'Elia (1971). In D'Elia (1971) leggiamo: «Le ricerche intorno alle sensazioni di moto, a quelle della vista e dell'udito hanno portato Mach a questa conclusione: il dato immediato è un complesso di qualità sensibili» (p. 78), la cui sintesi va concepita

In questa prospettiva, si dovrà dire allora che il modello di Gabor si astiene dal giudizio circa l'oggettività delle proprietà che descrive nel fenomeno esaminato. Muovendo dalle limitazioni della teoria di Fourier (come d'altronde altri importanti scienziati suoi contemporanei<sup>16</sup>), Gabor perviene ad una rappresentazione del segnale esplicitamente comprensiva della determinazione temporale: in questo quadro formale, ogni cambiamento di scala (temporale) di osservazione potenzialmente conduce ad evidenziare proprietà diverse nel segnale sottoposto ad osservazione, all'interno di uno spazio di soluzioni piuttosto ampio i cui estremi vanno appunto dall'oscillazione armonica (funzione sinusoidale) alla discontinuità dell'impulso (funzione delta). Certo, in Gabor la nozione di 'scala temporale di osservazione' è del tutto implicita: renderla esplicita toccherà ad alcuni suoi continuatori, tra cui Jean Morlet (1931-2007), che arriveranno a formulare metodi di analisi in cui la determinazione della frequenza è addirittura sospesa e considerata derivativa. Come accenneremo più avanti, si parlerà allora di 'rappresentazioni tempo-scala', per esempio nell'analisi di segnali mediante *wavelet* (Kronland-Martinet e Grossmann 1991; Kronland-Martinet, Guillemain e Ystad 1997).

Probabilmente una certa propensione a stabilire un legame tra modello analitico ed esperienza uditiva spiega anche come mai Gabor utilizzi metafore esplicitamente musicali – quali «nota» o «trillo», usati in Gabor (1947c) in rapporto ovviamente a fenomeni di modulazione di frequenza. Peraltra, nel contesto della fisica dei primi anni del XX secolo, non era raro ricorrere a metafore musicali proprio in rapporto alla centralità del fluire del tempo nel vivo dell'esperienza sensibile. Mach stesso aveva accompagnato la sua 'analisi delle sensazioni uditive', oltre che con esempi relativi al parlato, con veri e propri esempi musicali<sup>17</sup>. Il padre fondatore della cibernetica, Norbert Wiener, in una conferenza tenuta a Gottinga nel 1925 aveva affermato che «la frequenza di una nota e la sua durata interagiscono in maniera molto complicata», e che la precisione nella determinazione dell'una va a svantaggio della precisione nella determinazione dell'altra (Werner Heisenberg era tra gli ascoltatori di quella conferenza)<sup>18</sup>. È da segnalare che il matematico francese Jean Ville (1910-1989) nel 1947 arrivò a delineare una propria forma di analisi quantistica del segnale introducendo come argomento cruciale i risultati paradossali ai quali porta l'applicazione dell'analisi di Fourier a interi passaggi musicali (Ville 1948, p. 64). Gabor sapeva che il suo metodo di scomposizione del segnale era «piuttosto sconveniente» nel senso che i segnali elementari in esso utilizzati in definitiva sono «funzioni [che] non formano un insieme ortogonale» (Gabor 1947a, p. 592; qui p. 12). Il punto venne approfondito successivamente da vari ricercatori e gruppi di ricerca, tuttavia – dopo un primo tentativo a metà degli anni sessanta (Helstrom 1966) – solo negli anni settanta e ottanta questa non-ortogonalità è stata descritta rigorosamente, insieme ad alternative adeguate (Balian 1981; Bastiaans 1980; Morlet, Kronland-Martinet e Grossmann 1987). In particolare, Jean Morlet (inizialmente attivo in ambito geofisico) sostituì i segnali elementari di Gabor con altri segnali pure tempo-finiti ma di durata inversamente proporzionale alla banda di frequenza o scala di osservazione, dando così avvio a metodi di rappresentazioni basati su pacchetti d'onda, o *wavelet* (Meyer 1993; Young 1993)<sup>19</sup>.

come «una attività del sistema nervoso» (p. 50). A tal proposito, vorrei segnalare come questi aspetti del lavoro di Mach siano stati visti come una forma antesignana di «costruttivismo fenomenologico» (Glaserfeld 1975; Riegler 2012).

<sup>16</sup>Come suggerisce Yves Meyer, intorno al 1945 anche personaggi quali Léon Brillouin, Claude Shannon e altri avevano chiare le limitazioni dei metodi basati sul modello analitico di Fourier (Meyer 2008).

<sup>17</sup>Si vedano i capitoli "The sensation of time" e "The sensation of tone" in Mach (1914, pp. 245-261, 262-309).

<sup>18</sup>Si veda Wiener (1964, pp. 105-107), la cui redazione iniziale è della metà degli anni cinquanta. E si veda la ricostruzione data in Conway e Siegelman (2005, pp. 84-86).

<sup>19</sup>Nello sviluppo dell'analisi *wavelet* va segnalato il lavoro portato avanti dal gruppo di lavoro dell'Università di Marsiglia (Morlet, Kronland-Martinet, Grossmann e altri) anche per i risvolti diretti sull'elaborazione di segnali

### 3 Dispositivi

Sarà utile ora segnalare una serie di circostanze tecnologico-musicali all'interno delle quali il modello teorico dei quanta acustici di Gabor ha avuto ripercussioni più o meno dirette.

#### 3.1 Meyer-Eppler, il *visible speech process*, e il *vocoder*

Nell'agosto 1950 si tenne a Darmstadt una serie di incontri dal titolo *Die Klangwelt der elektronischen Musik*. Ne furono protagonisti Werner Meyer-Eppler (1913-1960) e Robert Beyer (1901-1989), che nell'occasione incontrarono il compositore Herbert Eimert (1897-1972) insieme al quale poco più tardi fondarono lo Studio für Elektronische Musik, alla stazione radio WDR di Colonia (Beyer 1953)<sup>20</sup>. Nel corso degli incontri (seguiti tra gli altri dall'anziano Edgard Varèse e da un giovanissimo Karlheinz Stockhausen), Meyer-Eppler tenne un intervento nel quale si soffermò sulla matrice di Gabor (Ungeheuer 1992, p. 104). Fu quasi certamente la prima volta che il lavoro di Gabor venne esposto a musicisti e musicologi.

Non possiamo accettare quando Meyer-Eppler abbia conosciuto il modello di Gabor. In ogni caso, in quanto assistente all'Istituto di Fonetica e Scienze della Comunicazione dell'Università di Bonn, egli aveva certamente accesso a molte delle pubblicazioni dell'epoca in ambito di teoria dell'informazione, oltre che di fonetica (ambito più specifico della sua ricerca). Era proprio all'incrocio di teoria dell'informazione e fonetica che gli scritti di Gabor venivano collocati, in quella fase. In un libro apparso qualche anno dopo, Meyer-Eppler illustra il diagramma informazionale di Gabor (Meyer-Eppler 1959, pp. 25-27) e cita vari articoli del collega anglo-ungherese, tra cui «Theory of communication» (p. 14) e «New possibilities in speech transmission» (pp. 26 e passim). Ma già nel 1955, in un articolo sul primo numero della rivista *Die Rehie* (diretta da Eimert e Stockhausen), egli aveva abbozzato un approccio al suono in termini informazionali, soprattutto al fine di dar conto di aspetti aleatori del suono che ne fanno un fenomeno essenzialmente aperiodico, quindi relativamente consono ad essere affrontato nei termini della teoria di Fourier (Meyer-Eppler 1955). Dall'articolo si deduce anche che Meyer-Eppler conosceva alcuni degli esperimenti di psicofisica citati da Gabor, e anche la precedente intuizione di Stewart (1931) circa la relazione di incertezza  $\Delta t \Delta f$  richiamata in Gabor (1946), di cui abbiamo fatto cenno<sup>21</sup>.

All'Istituto di Fonetica di Bonn, Meyer-Eppler poteva utilizzare uno strumento di recentissima disponibilità chiamato *visible speech process* che permetteva di realizzare rappresentazioni sonografiche (fu usato anche per preparare il sonogramma apparso nel 1955 in apertura del primo numero di *Die Rehie*, illustrante un passaggio del brano elettronico *Glockenspiel* composto da Eimert due anni prima). Tale dispositivo nasceva dal lavoro condotto a partire dal 1941 da Ralph Potter e altri presso i laboratori Bell Telephone, negli Stati Uniti. Gabor (ibid., p. 432) stesso ne fa menzione, citando anche le pubblicazioni in cui Potter ne illustrò i criteri (Gabor 1947a, p. 594; qui p. 12)<sup>22</sup>. Gabor cita anche, e anche più in particolare, un altro ricercatore dei laboratori Bell Telephone, Homer Dudley, che anni prima aveva progettato il

---

musicali. Si veda la ricostruzione storico-scientifica in Farge et al. (2012), e si vedano i contributi di Daniel Arfib, Jean-Claude Risset, Alex Grossman e Richard Kronland-Martinet in De Poli, Piccialli e Roads (1991).

<sup>20</sup> Si veda anche la ricostruzione storica in Morawska-Büngeler (1988, pp. 9 e sgg.).

<sup>21</sup> Si veda nota 9.

<sup>22</sup> Riferendosi a questo tipo di apparecchiature, Gabor parla di «sound spectrograph» (Gabor 1946, p. 432). A partire da (Gabor 1952), egli sostituisce l'espressione col termine «sonogram», più adatto ad evitare confusioni altrimenti possibili, e d'altra parte in uso ancora oggi.

vocoder (Dudley 1936, 1939) definito da Gabor una «ingegnosa macchina di analisi e sintesi della voce parlata» (Gabor 1946, p. 442)<sup>23</sup>.

Quel che interessa maggiormente Gabor, in dispositivi del genere, è l'analogia formale con la sua matrice di quanta di informazione: anche tali dispositivi infatti implicavano una forma di discretizzazione dei domini di rappresentazione. Riferendosi al vocoder, Gabor in particolare pone la questione in termini della massima «compressione» d'informazione possibile capace di preservare l'intellegibilità del parlato (*ibid.*, p. 445). È questa ricerca dell'unità minima d'informazione che ai suoi occhi rende il vocoder di grande interesse, indipendentemente dalla qualità del suono che esso poteva generare, che in effetti gli appare piuttosto innaturale e «machine-made» (Gabor 1952, p. 40), e che Gabor giudica comunque migliorabile aumentando i canali del processo di analisi e sintesi della macchina di Dudley. La terza parte di Gabor (1946) è appunto dedicata ad un metodo per «comprimere» ed «espandere» il suono.

Vi torneremo più avanti. Per ora notiamo che, declinata in questo modo, la problematica posta da Gabor è quella della riduzione dei dati, e naturalmente si inserisce in un quadro assai generale degli sviluppi delle tecniche di comunicazione; per quel che riguarda i temi che ci interessano più da vicino, essa viene intesa da Gabor come ‘riduzione dei dati in base a criteri percettivi’, e come tale riaffiorerà decenni dopo, prima tangenzialmente alla ricerca sulla sintesi digitale del suono – almeno in una fase in cui, in quel settore, le indicazioni della psicoacustica sono apparse decisive (Risset e Wessel 1982a) – e poi soprattutto nello sviluppo di algoritmi di compressione per la trasmissione di segnali audio (e non solo) lungo le reti telematiche (Brandenburg 1999).

### **3.2 Sintesi ed elaborazione granulare del suono**

Ci stiamo dunque avvicinando agli aspetti di carattere più tecnologico e ingegneristico del lavoro di Gabor sui quanta acustici. In appendice alla prima parte di Gabor (1946), viene descritto un dispositivo per la generazione del suono utile a verificare aspetti del modello teorico proposto. Il dispositivo sfrutta mezzi ottici elettromeccanici, in parte basati su precedenti studi sulla riverberazione artificiale del suono (Goldmark e Hendricks 1939). Lo si può pensare come il primo strumento mai ideato per la sintesi granulare del suono (Roads 2001a, p. 302). Va precisato però che, per Gabor (1946, p. 440), non si tratta tanto di un «metodo praticabile» ma di un momento di verifica sperimentale di certi effetti di modulazione (*sidebands*) insiti, in linea di principio, in qualsiasi realizzazione concreta del modello teorico, e simili comunque a fenomeni già noti in ambito di teoria delle comunicazioni.

Dire che Gabor «aveva preconizzato una sintesi dei suoni» a base granulare (Xenakis in Restagno 1988, p. 30) rappresenta, da un certo punto di vista, una leggera forzatura. Se non altro, all'epoca di Gabor il concetto di sintesi del suono non ha un suo statuto definito: da una parte, appare del tutto implicito nella costruzione di strumenti musicali elettronici; dall'altra, vale come strumento di verifica della coerenza di un modello teorico – come appunto nel caso del generatore di suono descritto alla fine della prima parte di Gabor (1946). Come segnala

---

<sup>23</sup>Gabor si trattiene più a lungo sul vocoder di Dudley in Gabor (*ibid.*, pp. 40-41). Per un quadro generale sullo sviluppo del vocoder e le sue ripercussioni anche musicali, si veda Zavagna (2013).

Yves Meyer, l'idea di sintesi del suono era stata in qualche modo preavvertita da Jean Ville (negli stessi anni delle pubblicazioni di Gabor che abbiamo preso in esame) appunto in quanto momento sperimentale complementare alla descrizione di una qualsiasi forma di analisi del segnale. Più in particolare, l'insieme degli atomi tempo/frequenza del modello di Gabor

... costituisce una collezione di segnali elementari che è fin troppo ampia per fornire una rappresentazione univoca [e] ciò porta a dover scegliere quella che risponde meglio rispetto a un certo criterio. Il criterio potrebbe essere quello suggerito da [Jean] Ville: l'analisi o decomposizione di un segnale in atomi tempo/frequenza dovrà essere messa in relazione ad una sintesi (Meyer 1993, p. 64).

Al di fuori dei centri di ricerca scientifica e dei laboratori dei costruttori di strumenti elettronici, l'idea di 'sintesi del suono' come formalismo coerente passibile di implementazioni diverse, a fini musicali o di altro tipo, si delinea solo qualche anno dopo le proposte di Gabor. Da una parte come schema concettuale che guida l'implementazione di processi di generazione del suono nel contesto degli sviluppi della musica elettronica analogica e specialmente informatica. Dall'altra, in campo più strettamente scientifico, come esito di una sempre maggiore specializzazione, e anche di una scissione sempre più netta tra esigenze sperimentali e sviluppi teorici.

In tema di implicazioni pratiche, ciò che preoccupa Gabor (1946, pp. 437-438) molto più della possibilità di una sintesi del suono è la velocità di trasmissione necessaria per trasportare i quanta di informazione attraverso un canale – tema ovviamente cruciale e decisivo nell'ambito della teoria delle comunicazioni. Ragionando su questo punto, Gabor perviene a calcolare valori che, non a caso, implicano una velocità di trasmissione inversamente proporzionale al doppio della larghezza di banda<sup>24</sup>. Date le limitazioni caratteristiche dei mezzi di comunicazione dell'epoca (ivi comprese le trasmissioni radiofoniche e televisive, ancora soggette a fattori troppo aleatori), Gabor suggerisce che l'unica alternativa per pervenire presto a condizioni operative efficienti consiste nel ridurre al minimo l'informazione da trasmettere, escogitando «metodi di compressione-espansione» del segnale (*ibid.*, p. 438) e cioè «condensando il contenuto di informazione ad un minimo» (*ibid.*, p. 445).

Il suo obiettivo diventa allora quello di verificare l'utilità del modello teorico al fine di codificare («condensare») il segnale ad un capo di una catena di trasmissione, per poi decodificarlo («espanderlo») all'altro capo (*ibid.*, p. 445). A tal fine, Gabor propone il progetto di un dispositivo che chiama «convertitore di frequenza», per indicare il fatto che esso può modificare («convertire») la frequenza e allo stesso tempo «non cambiare il ritmo, ovvero l'articolazione temporale, di segnali sonori del parlato o musicali» (*ibid.*, p. 452). L'idea viene così riassunta:

il segnale [da elaborare], registrato su un supporto in modo permanente o in modo temporaneo, passa accanto ad una finestra fissa, dal contorno opportunamente graduato [inviluppo di ampiezza] e viene, durante il tempo di esposizione in questa finestra, letto più volte da rilevatori [*pick-ups*] che si susseguono a velocità differente dalla velocità di scorrimento del supporto di registrazione (*ibid.*, p. 453)<sup>25</sup>.

<sup>24</sup>Ciò andrebbe ricondotto al teorema del campionamento di Shannon, che d'altra parte derivava dal lavoro di Hartley e Nyquist più volte citato da Gabor.

<sup>25</sup>Si noti, in questo passaggio, il termine «finestra» [*window*], che più tardi sarà integrato al vocabolario dell'elaborazione numerica e dell'analisi tempo/frequenza, e che equivale alla durata di esposizione – e quindi all'inviluppo di ampiezza (o 'funzione finestra') – di un certo segmento minimo di segnale.

Per costruire un dispositivo del genere, una volta esclusi supporti facilmente danneggiabili come «dischi grammofonici e relative puntine di lettura», bisogna considerare le «possibilità promettenti» (*ibid.*, p. 453) offerte da tecnologie elettromeccaniche messe a punto qualche anno prima, e ancora ampiamente in corso di perfezionamento nel momento in cui Gabor scrive. Un primo metodo, o «metodo cinematico», si baserà su un proiettore per pellicole cinematografiche da 16 mm (dotate di ‘colonna sonora’) opportunamente modificato mediante un tamburo rotante che provvede a molteplici letture di una medesima ‘finestra’ di esposizione del segnale impresso sulla pellicola. Una descrizione più approfondita di questo primo metodo sarà fornita da Gabor insieme alla richiesta di brevetto nel 1944 (Gabor 1949). Un secondo metodo sfrutta invece i principii della magnetizzazione e può ricorrere cioè alla registrazione magnetica su cavo (*wire recorders*) o su nastro (*tape recorders*).

Sarà proprio modificando il funzionamento di un magnetofono a nastro che l’idea di Gabor prenderà piede. Si tratta di sostituire il normale blocco di lettura del magnetofono (la ‘testina’ nel cui campo magnetico passa il nastro) con un blocco dotato di una o più testine di lettura e capace di girare su sé stesso a velocità diversa (superiore o inferiore) rispetto alla velocità di scorrimento del nastro, e di registrare il segnale rilevato da questo blocco rotante su un secondo magnetofono. Il segnale risultante presenterà allora un numero proporzionale di quanta di informazione dal contenuto in frequenza identico a quello dell’originale: lo si potrà allora riprodurre alla medesima velocità di registrazione, ottenendo un suono di durata diversa (per esempio più lungo, in proporzione al numero di testine sul blocco di lettura e alla velocità di rotazione del blocco stesso) senza alterazioni (o almeno senza deformazioni troppo evidenti) del contenuto in frequenza; oppure lo si potrà riprodurre a velocità diversa dalla velocità di registrazione, regolata in modo da ottenere un suono di durata identica all’originale ma con conseguente alterazione del contenuto in frequenza.

Si tratta della prima forma di trattamento del segnale in cui tempo e frequenza siano modificabili in modo virtualmente indipendente l’uno dall’altro. Pur con tutte le differenze tecnologiche, si tratta dello stesso processo effettuato, vari decenni dopo, da certi algoritmi di *time-stretching* (o di *frequency-stretching*, se si vuole richiamare l’espressione «*frequency converter*» di Gabor) di segnali audio digitali (Arfib 1991; Jones e Parks 1988). Va infine segnalata l’osservazione di Gabor secondo la quale, in linea di principio, l’operazione del suo «convertitore di frequenza» andrebbe effettuata non sull’intero spettro del segnale iniziale ma su «almeno due porzioni dello spettro, diciamo tra 25 e 500 Hz, e tra 500 e 7500 Hz» (Gabor 1946, p. 454). L’osservazione rimase probabilmente senza seguito, tuttavia oggi, nell’ambito delle tecniche di analisi-sintesi del suono e in quello parallelo della compressione di dati audio, è normale procedere con metodi di codifica applicati indipendentemente su (molte più di due) porzioni dello spettro, ottimizzandone il funzionamento in rapporto alle corrispondenti grandezze di tempo e frequenza.

### **3.3 Moles, il *phonogène*, il *time-compressor*, il *tempophon***

Nel 1958 Abraham Moles (1920-1992) parla del fenomeno musicale nei termini di una «atomizzazione dell’oggetto sonoro, proposta dalla teoria dell’informazione, in elementi di durata percettiva» (Moles 1969, p. 173). È con tale orientamento che qualche anno prima, ancora molto giovane, egli si era accostato allo studio degli allora nuovissimi mezzi di produzione musicale elettroacustici per giungere infine a delineare una sua ‘teoria della musica sperimentale’ (Moles 1961a). Verso la fine degli anni quaranta, Moles aveva iniziato una

collaborazione importante con Pierre Schaeffer al GRM di Parigi<sup>26</sup>. Nei suoi scritti degli anni successivi, troviamo citati sia i metodi sonografici di Potter sia il lavoro di Meyer-Eppler, oltre ad altre fonti che ben si sovrappongono ai riferimenti bibliografici di Gabor. Moles non cita Gabor, pur essendo piuttosto verosimile che ne conoscesse la ricerca, se non attraverso gli scritti appunto di Meyer-Eppler, almeno a partire dall'articolo dello stesso Gabor (1948b) apparso in francese su *L'Onde électrique*, cui abbiamo accennato sopra<sup>27</sup>.

L'idea di Moles consiste nel quantizzare i 'diagrammi di uguale intensità sonora' di Fletcher-Munson, cioè le mappe psicoacustiche fondamentali del campo di sensibilità del sistema uditivo umano, per ottenerne dei «quanti di percezione» (Moles 1969, pp. 167-172). L'idea appare già nei diagrammi pubblicati da Moles in Schaeffer (1952) e può essere assimilata effettivamente ad una ripresa del diagramma informazionale gaboriano riformulato in termini psicoacustici. Nei primi mesi del 1950<sup>28</sup> il tecnico del GRM, Jacques Poullin (1920-2014), costruisce per Moles e per Schaeffer il *phonogène*, apparecchiatura che poteva trasformare il suono registrato su nastro magnetico secondo diverse possibilità, in modi virtualmente simili a quelli del «convertitore di frequenza» di Gabor, e ottenendo delle «trasposizioni totali» (Schaeffer 1966, p. 407), in cui spettro e durata sono gestiti indipendentemente. Nel descriverne le possibilità, Poullin (1954) assimila le 'trasposizioni di durata' al *ralenti* del cinema. Il dispositivo, che utilizzava un blocco rotante di quattro testine di lettura (Moles 1961a, pp. 71-76) e che già nel 1952 Moles (1969, p. 310) accostò al vocoder, fu brevettato da Schaeffer e fu oggetto di qualche tentativo di commercializzazione.

Possibilità simili, e per certi versi più direttamente correlate al funzionamento del «convertitore di frequenza» di Gabor, erano quelle del *tempophon* (o *zeitregler*), costruito e commercializzato nel corso della prima metà degli anni cinquanta da Anton Springer, inizialmente illustrato in pubblico nello studio di Hermann Scherchen a Gravesano, in Svizzera (Springer 1955), e quelle del *time-compressor* sviluppato da Grant Fairbanks all'Università di Illinois (Fairbanks, Everitt e Jaeger 1954). Quest'ultimo, ripreso e perfezionato in ambito statunitense nel corso degli anni sessanta, ebbe una qualche notorietà tra gli esperti e gli appassionati di registrazione magnetica, e fu ampiamente adoperato nel contesto della ricerca sulla comprensione del parlato (Foulke 1971)<sup>29</sup>.

Il *tempophon* di Springer andava semplicemente collegato per via puramente meccanica ad un normale magnetofono, come una protesi, il che ne facilitava l'uso e la diffusione tra tecnici e musicisti. Non a caso fu presto disponibile allo Studio di Fonologia Musicale della RAI di Milano. Nel 1956, il fisico e fondatore dello studio milanese, Alfredo Lietti (1919-1998), ne dava questa succinta illustrazione<sup>30</sup>:

Questo dispositivo usato con un magnetofono normale consente di variare la durata del tempo della registrazione senza alterare l'altezza dei suoni registrati [...]. Il principio di

<sup>26</sup>Il capitolo finale in Schaeffer (1952), intitolato "Esquisse d'un solfège concret", fu scritto in collaborazione con Moles. Secondo la ricostruzione di Mathien (2007), Moles conobbe Schaeffer nel 1945, dunque ancor prima delle circostanze che nel 1948 portarono Schaeffer alla ricerca della 'musica concreta'.

<sup>27</sup>Si veda nota 7.

<sup>28</sup>Desumiamo questa datazione da Schaeffer (1952, p. 97).

<sup>29</sup>Qualche anno più tardi, il lavoro di Fairbanks costituì anche la base di alcuni primi tentativi di compressione e dilatazione temporale del parlato realizzati mediante computer (Cramer e Talambiras 1971; Huggins 1974). È significativo inoltre come nel contesto di queste ricerche, e particolarmente in Cramer e Talambiras (1971), venga richiamato il lavoro di Gabor e di Springer.

<sup>30</sup>Prendo la citazione dalla copia personale di un documento interno dello Studio di Fonologia privo di numerazione di pagina. Il documento nel 1956 venne riprodotto quasi alla lettera in un numero della rivista *Elettronica*. Più recentemente è stato ripubblicato in Lietti (1998).

funzionamento dell'apparecchio è il seguente: il nastro viene riprodotto, anziché da una testina fissa, da un gruppo di quattro testine montate a 90° tra loro su un disco rotante. Sia  $V_a$  la velocità di scorrimento del nastro e  $V_k$  la velocità tangenziale del [blocco] porta-testine [...] Si fa in modo che  $V_a + V_k = V_r$ , dove  $V_r$  è la velocità nominale della registrazione [15 o 30 pollici al secondo]. Su un tempo  $T$  la lunghezza  $l$  di nastro utilizzata [per registrare un suono] è  $l = TV_a$ . Agli effetti della corretta riproduzione [...] senza l'alterazione delle frequenze è necessario che il nastro sia riprodotto alla velocità relativa  $V_r$ , che corrisponde come si è visto a quella normale.

Pertanto, riproducendo il nastro lungo  $l$  alla velocità  $V_r$  si ha una durata di riproduzione diversa dalla durata originale, ovvero di  $T' = 1/V_r = 1/(V_a + V_k) \neq T$ . In particolare:

$$\begin{aligned} V_k = 0 &\rightarrow T' = T \\ V_k > 0 &\rightarrow T' < T \\ V_k > 0 &\rightarrow T' > T \end{aligned}$$

Queste relazioni indicano la modifica di durata senza alterazione di frequenza. Ovviamente è possibile anche ottenere il risultato inverso, cioè modificare la frequenza senza alterare la durata.

Il *tempophon* fu utilizzato in produzioni musicali importanti, per esempio a Milano nel 1958, nella realizzazione di materiali sonori utilizzati in *Thema (Omaggio a Joyce)* di Luciano Berio (Di Scipio 2000; Zavagna 1992). A Colonia, tra il 1960 e il 1962, Herbert Eimert ne esplorò le potenzialità nella realizzazione di *Epitaph für Aikichi Kuboyama*, in modo peraltro alquanto singolare, ‘frantumando’ e rendendo deliberatamente irriconoscibili i suoni di una voce preregistrata, e cioè puntando tutto sugli artefatti di modulazione dei processi di dilatazione in tempo e frequenza – artefatti d’altra parte non trascurabili in condizioni tecniche men che ideali. In tempi recenti, anche le corrispondenti forme di *time-stretching* in ambito di elaborazione numerica di segnali musicali sono state adoperate non solo con finalità per così dire ‘corrette’ (cioè per modifiche della durata e/o della frequenza prive di artefatti udibili, per esempio per scopi professionali di *sound design*), ma spesso anche con finalità più creative e del tutto arbitrarie, cioè miranti a non mascherare (o anche ad esibire) gli effetti residuali del processo tecnico, particolarmente evidenti con dilatazioni estreme della durata. Con ciò ritorniamo, in un certo senso, agli esperimenti di Gabor: i quali, come si è visto, miravano certo a verificare la coerenza teorica del modello di analisi, ma anche a studiare appunto gli effetti di modulazione risultanti dall’applicazione pratica del modello teorico.

È vero comunque che, al di là delle particolari finalità perseguitate, dispositivi quali il *tempophon* e il *phonogène* possono essere considerate «macchine analitiche», per usare un’espressione di Moles (1957), cioè dispositivi operanti in base a rigorosi principii di analisi e rappresentazione tempo/frequenza del segnale. Nel caso in questione, erano proprio quei principii ad essere del tutto nuovi: le implicazioni pratiche ed artistiche nel potenziale tecnico che si dischiudeva con strumenti quali il «convertitore di frequenza» di Gabor, inizialmente non potevano che rimanere ampiamente imprevedibili.

## 4 Analogie musicali

### 4.1 Xenakis, sintesi granulare, «nuvole di suono»

Tra il 1958 e il 1960, Iannis Xenakis (1922-2001) traccia le premesse teoriche ed operative della composizione di ‘musica stocastica’, ma contestualmente avanza anche una sua ‘ipotesi granulare’ circa la natura stessa del suono:

Ogni suono è un’integrazione di grani, di particelle sonore elementari, di quanta sonori. Ciascuno di questi grani elementari ha una tripla natura: durata, frequenza e intensità. Ogni suono, ogni variazione sonora [...] è un assemblaggio di grani elementari sufficientemente numerosi e adeguatamente disposti nel tempo (Xenakis 1963, p. 61) e (Xenakis 1992, p. 43).

L’ipotesi viene avanzata nel quadro dei preparativi che portano il compositore a scrivere *Analogique A et B* (per nove archi e nastro magnetico, 1958-1959), e precisamente nel quadro della lavorazione in studio (al GRM di Parigi) della parte su nastro magnetico, inizialmente pensata come brano a sé stante, intitolato appunto *Analogique B*. L’ipotesi granulare di Xenakis rappresenta

... una prima approssimazione delle considerazioni introdotte in teoria dell’informazione da Gabor. Nella cosiddetta matrice di Gabor, un evento sonoro si risolve in segnali elementari acustici di durata effettiva molto corta, e di ampiezza suddivisa in quanti, nel senso della teoria dell’informazione. Questi segnali elementari costituiscono funzioni sinusoidali aventi un inviluppo ‘a campana’, ovvero di tipo gaussiano, ma possono essere rappresentati abbastanza bene anche con onde sinusoidali di durata ridotta aventi inviluppo rettangolare (*ibid.*, p. 373).

Nel 1959, Xenakis dunque intende comporre il suono a partire da «grani sonori del tipo di Gabor» (Xenakis 1963, p. 122) e (Xenakis 1992, p. 103), creando trame sonore in evoluzione che egli assimila a «nuvole di suono»<sup>31</sup>. Nel 1967, quando quella fase di sperimentazione appare ormai superata e Xenakis è immerso in progetti molto diversi, egli insiste tuttavia nel concepire suono e musica «come una nuvola gigantesca di corpuscoli sonori, di granelli sonori» (Xenakis 1982, p. 30). Il concetto viene ribadito in occasioni successive, almeno fino all’inizio degli anni novanta, quando Xenakis realizza i suoi due ultimi brani elettroacustici, interamente computerizzati, *Gendy3* (1991) e *S.709* (1994): pur con tecniche di generazione del suono diverse (e tuttavia operanti esclusivamente nel dominio del tempo), essi confermano che il legame fra procedure stocastiche di composizione e tecnologie di generazione del suono va ritenuto indissolubile e sistematico nella concezione musicale di Xenakis (Di Scipio 2015).

Dobbiamo aggiungere che Xenakis coglie in Gabor soprattutto la critica del modello di Fourier, nella sua dimensione atemporale e astratta rispetto alla dinamicità dei fenomeni acustici, inclusi quelli di interesse musicale. Di questa visione problematica rispetto all’acustica musicale tradizionale (e alle sue ripercussioni sulla teoria musicale), egli farà un tratto ricorrente in tutto il suo lavoro successivo, in particolare nel contesto di composizioni realizzate con mezzi elettroacustici e informatici – si vedano vari spunti in proposito in Xenakis (2003). Secondo Angelo Orcalli (1993, p. 87), «l’adesione di Xenakis alla concezione atomistica del suono riflette l’attenzione del compositore verso le tendenze più analitiche della speculazione

<sup>31</sup>Riguardo a questa metafora per così dire ‘atmosferica’, e riguardo ad *Analogique A et B* nel suo insieme, mi permetto di rinviare a Di Scipio (2003, 2006b).

scientifica [del XX secolo]» e manifesta «un’istanza teoretica che sta a fondamento di tutta la sua produzione musicale». Si potrà notare, però, che se Gabor giunge «fino ai confini tra teoria dei segnali e psicoacustica» collegando «le dimensioni logiche del segnale con le capacità del ricevitore umano» (*ibid.*, pp. 85, 86), ciò risulta particolarmente conforme alla poetica di Xenakis, col suo tentativo estetico di fondere una certa immediatezza o anche ruvidezza dell’esperienza sensibile del suono, e una dimensione di astrazione logico-costruttiva.

In alcune occasioni Xenakis ha rifiutato una diretta associazione tra la sua ipotesi granulare e il modello di Gabor, quasi che l’accostamento potesse in qualche modo sminuire l’originalità del suo lavoro. Per esempio, nel 1992 egli parla dei quanta sonori come di una ricerca del tutto personale, «erroneamente attribuita a Gabor» (Xenakis 1992, p. xiii), e dichiara di non essere stato a conoscenza del modello di Gabor al momento di iniziare la lavorazione di *Analogique B*. La sua ipotesi granulare andrebbe invece collegata con studi svolti da Albert Einstein «fra il 1917 e il 1918» (Xenakis in Restagno 1988, p. 30), nel contesto dei quali sarebbe stato coniato il termine ‘fonone’ (in parallelo con ‘fotone’, cioè quanto di luce)<sup>32</sup>. In ogni caso, dobbiamo dedurne che soltanto dopo o durante la lavorazione di *Analogique B* Xenakis venne a conoscenza della proposta scientifica di Gabor, forse attraverso Meyer-Eppler (1959), e che la vide come un’autorevole conferma della sua personale ipotesi (Orcalli 1993, p. 81). Personalmente direi che l’ipotesi di Xenakis possa anche derivare dalle importanti indicazioni che egli ricavava dalla conoscenza e dallo scambio personale con Abraham Moles (Xenakis 1962)<sup>33</sup>.

Sicuramente anche alcuni dei compositori attivi allo studio di musica elettronica di Colonia furono a conoscenza del modello di Gabor, dato il ruolo decisivo che svolgeva Meyer-Eppler nello studio. Va segnalata peraltro la circostanza che un brano paradigmatico come *Kontakte* di Stockhausen (per pianoforte, percussioni, e nastro magnetico, 1960) – l’opera musicale che più ha introdotto nell’estetica di Colonia l’uso di tessiture sonore complesse e di processi di microframmentazione sonora – fu realizzato proprio mentre veniva preparata la pubblicazione di Meyer-Eppler (1959), dove, come si è detto, compare anche una descrizione della matrice di Gabor. Ciononostante, accostare la ‘modulazione di impulsi’ e altre soluzioni tecniche impiegate da Stockhausen al modello analitico di Gabor significa effettivamente suggerire un’associazione troppo generica e forzosa. In ogni caso, pur nel suo grandissimo rilievo musicale, *Kontakte* non aprì ad approfondimenti in tal senso.

In retrospettiva, il processo di costruzione del suono attuato per la realizzazione di *Analogique B* è stato assimilato ad una forma particolare di ‘sintesi granulare’, a partire da Roads (1978). Tuttavia, come vedremo subito, esso presentava approssimazioni anche molto importanti rispetto al modello di Gabor. Ciò è del tutto ragionevole, visto che, nell’insieme, il valore del contributo di Gabor si fonda naturalmente sulla coerenza formale nel render conto di qualsiasi segnale mediante una costruzione teorica rigorosa, laddove il principio che guida Xenakis è di ordine sintetico, cioè generativo, non analitico, e vale solo per la sua capacità di dar luogo a costruzioni sonore di cui soltanto la musicalità del compositore stabilisce la coerenza. L’ipotesi granulare di Xenakis, pur ‘teorica’<sup>34</sup>, va considerata essenzialmente come

<sup>32</sup>Xenakis non offre alcuna indicazione bibliografica che chiarisca questo suo riferimento ad Einstein.

<sup>33</sup>Tra coloro che operavano nel contesto del GRM nella seconda metà degli anni cinquanta, Moles fu certamente tra i pochi che ebbero forte sintonia con Xenakis. Anni più tardi, Moles divenne prezioso riferimento anche per Jean-Claude Risset, che spinse a trasferirsi nel 1964 negli Stati Uniti per approfondire insieme a Max Mathews, presso i laboratori Bell Telephone, i fondamenti della sintesi numerica del suono (uno snodo importante per l’evoluzione successiva dell’informatica musicale e della composizione al computer).

<sup>34</sup>Si veda l’articolo di Solomos in questo numero di «Musica/Tecnologia».

procedura operativa, come idea che orienta un procedimento tecnico e creativo di composizione del suono (o meglio di ‘microcomposizione’ in questo caso), secondo motivazioni tutt’altro che analitiche (va tenuto presente che, anche per formazione, l’attitudine di Xenakis fu complessivamente più ingegneristica che scientifica, più pratica che teoretica).

Il procedimento di costruzione del suono attuato per *Analogique B* fu piuttosto rudimentale e faticoso, non solo per la relativa praticità dei mezzi elettroacustici allora utilizzabili<sup>35</sup>, naturalmente, ma in fondo anche per le stesse regole che il compositore volle tematicamente seguire, di tipo probabilistico, le quali difficilmente si sarebbero potute rivelare utili nel far emergere, dalla trama discreta di grani, le sonorità fuse e continue che Xenakis invece sperava di conseguire (Di Scipio 1997). Xenakis approssimò i segnali elementari di Gabor con onde sinusoidali di circa 40 millisecondi, aventi inviluppo praticamente rettangolare anziché gaussiano (dettaglio non marginale sia in termini di coerenza analitica sia in termini di pregnanza dei risultati di sintesi). Un’altra differenza non trascurabile sta nel fatto che, mentre il modello di Gabor implica una successione ordinata di quanta acustici, ovvero una disposizione coerente alla sequenzialità temporale del segnale analizzato, Xenakis invece ordina i suoi granelli sonori conformandosi da una parte alle comprensibili costrizioni pratiche, e dall’altra ai criteri di distribuzione probabilistici che si era assegnato: la linearità della successione, che in Gabor è ovvia condizione di coerenza, viene del tutto sospesa e disarticolata dal compositore.

Con una terminologia affermatasi successivamente, la procedura di Xenakis va considerata come una forma di sintesi granulare asincrona. Solo una forma di sintesi granulare sincrona potrebbe, a rigore, essere ricondotta al modello di Gabor. Non a caso, quasi tutte le tecniche granulari di concreto interesse musicale sono state di tipo asincrono, mentre quelle di tipo sincrono sono state più spesso utilizzate per più specifici interessi di ricerca (Cavaliere e Piccialli 1997; De Poli e Piccialli 1991) e possono anche riallacciarsi a certe tecniche di analisi e sintesi della voce, come per esempio nelle procedure di sintesi con funzioni d’onda formantiche, FOF (Rodet 1984).

## 4.2 Computer music e oltre

Dopo aver seguito un seminario di Xenakis alla Indiana University, nel 1972, Curtis Roads per primo si pose il compito di implementare la sintesi granulare al computer (Roads 1978)<sup>36</sup>. I suoi esperimenti riprendono quindi certi aspetti della sperimentazione di Xenakis, ma muovono anche da una maggiore consapevolezza del modello di Gabor. Roads realizza grani della durata fissa di 40 millisecondi, ma con inviluppi di forma gaussiana, e segue regole statistiche per la distribuzione di grani nel tempo e nella frequenza (Roads 2001a, p. 302), ma di più semplice gestione rispetto a quelle di Xenakis. Procedendo con lunghe liste di istruzioni per il sistema di programmazione MUSIC V di Max Mathews, Roads realizza nel 1974 un primo studio musicale della durata di poche decine di secondi, *Klang-1*. Quel che Roads coglie in Gabor sin dall’inizio è la mescolanza di «approfondimento teorico [...] e pratica sperimentale» (Roads 1991, p. 45). Ciò lo porta, differentemente da Xenakis, a valutare attentamente il «convertitore di frequenza» di Gabor. Allo stesso tempo, tuttavia, Roads prende spunto dal modello teorico gaboriano per dedurne, ai suoi fini musicali, tutta una

<sup>35</sup> Il laborioso procedimento di realizzazione in studio di *Analogique B* è ricostruito, fin dove possibile, nel documento inedito (Di Scipio 2006a).

<sup>36</sup> Sembra comunque che Xenakis, al tempo di *Analogique B*, abbia non solo ipotizzato l’utilizzo di sistemi di calcolo (Xenakis 1963, p. 72), ma anche fatto un concreto ma vano tentativo a quel fine (Leprince-Ringuet 1981).

serie di variazioni circa la struttura degli atomi sonori, optando talvolta per soluzioni che si risolvono in complesse forme di modulazione di impulsi (Roads 2001b), dunque in sostanza più prossime all'operatività di Stockhausen in *Kontakte*, che non a Xenakis in *Analogique B*. In ogni caso, nel corso degli anni il suo lavoro ha contribuito come nessun altro a far conoscere nel mondo della musica sia l'ipotesi granulare di Xenakis sia gli elementi essenziali della teoria di Gabor. Il momento culminante di quest'opera di divulgazione è stato certamente la pubblicazione del volume *Microsound* (Roads 2001a).

Un altro compositore nordamericano, il canadese Barry Truax, a partire dalla metà degli anni ottanta ha lavorato intensamente con la sintesi e l'elaborazione granulare al computer. Sua è stata la prima implementazione di sintesi granulare in tempo reale (Truax 1988b), dove i criteri statistici ripresi da Xenakis e da Roads, riconcettualizzati in termini di ‘maschere di tendenza’ (orientamenti statistici entro margini controllati), sono sottoposti a possibili controlli direttamente da terminale. Il primo risultato importante è stato *Riverrun* (1988), un esteso brano di soli suoni generati al computer (in parte anche montati e mixati mediante nastro magnetico quadrifonico), con numerose texture granulari di densità variabile, tra loro sovrapposte. Si tratta forse dell'unico lavoro musicale in cui Truax abbia usato davvero la *sintesi* granulare, dato che tutti i suoi lavori successivi sfruttano piuttosto una varietà di tecniche di *elaborazione* granulare del segnale in tempo reale, alle quali spesso il compositore sottopone registrazioni prese in contesti ambientali diversi (Truax 1988a, 1990).

Dagli anni ottanta in poi, molti altri compositori e ricercatori hanno esplorato forme più o meno personali di sintesi ed elaborazione granulare, con o senza riferimento a modelli analitici sottostanti. Tra gli altri, spicca il lavoro del compositore franco-argentino Horacio Vaggione, i cui lavori acustematici sono quasi sempre basati su sofisticate tecniche di micro-montaggio di suoni prodotti da tradizionali strumenti musicali. In Italia, a parte alcuni esiti iniziali che ho personalmente conseguito con controlli digitali di oscillatori analogici (*Punti di tempo*, 1987) e con la granulazione di sorgenti sonore strumentali (*Plex*, 1991, per contrabbasso e supporto quadrifonico), si segnalano le tecniche approntate da Eugenio Giordani nella forma di complesse modulazioni di ampiezza (*Solaria*, 1988) e varie forme di elaborazione in tempo reale indagate da Michelangelo Lupone (*Varianti di un grano*, 1993).

Nel corso degli anni novanta, innumerevoli sono stati i musicisti che hanno trovato in una dimensione più generale di ‘granularità’ del suono mezzi ed idee consoni alla propria sensibilità, anche nel contesto di tendenze estetiche descritte come *microsound* o *glitch*, accomunate da un certo ‘residualismo tecnologico’ (Cascone 2000; Di Scipio 2001; Thomson 2004). In generale, la quantità di esperienze di microcomposizione, talvolta legate anche ad una recuperata condizione esecutiva e performativa dell'elettronica, anche di tipo improvvisativo, è ormai pari solo all'incredibile varietà di metodi analitici praticabili sulla base di rappresentazioni tempo/frequenza o tempo/scala, e di interesse applicativo ben oltre l'ambito dell'elaborazione di segnali musicali – si veda la rassegna in Meyer (1993), con esempi presi da settori di interesse e di studio assai diversi tra loro, dalla sismologia all'elaborazione numerica delle immagini, e tanti altri.

## 5 Conclusioni

Sotto il profilo tecnico-compositivo, qualsiasi forma di sintesi o di elaborazione granulare è definita in fondo da scelte di progettazione compiute in rapporto a due momenti distinti e determinanti: (1) nell'individuare le *proprietà del grano* e stabilirne i tratti da assoggettare

a controlli musicalmente significativi; e (2) nell'escogitare adeguate *strutture di controllo* – ovvero un *front end processor* (Roads 1991, p. 152) – per la distribuzione dei grani nel tempo e per la determinazione dei parametri in ciascun grano. Va sottolineato che è proprio la struttura dei controlli (più ancora delle proprietà dell'unità sonora minima) a distinguere una particolare forma di sintesi granulare da un'altra. Il che è di grande importanza teorico-musicale e compositiva, perché significa che allora il potenziale espressivo prende forma già nello svolgersi dei processi di sintesi del suono, appunto nel micro-tempo (Di Scipio 1994), a partire da condizioni percettive, relative alla ‘stratificazione’ o alla ‘densità’ (MacKay 1984), costitutive di una nuova sintassi. In un certo senso, parlare di sintesi granulare può apparire riduttivo o inadeguato, se con questo termine si vuole indicare uno specifico metodo di sintesi tra gli altri. Non a caso sono stati suggeriti termini più generici, come ‘microcomposizione’ (o come ‘micro-time sonic design’ e ‘microsound’) che, pur relativamente diffusi, a loro modo appaiono rilevanti nella misura in cui rinviano più direttamente ad una condizione operativa dove ‘comporre il suono’ significa mettere insieme gli elementi minimi della struttura temporale della materia sonora, facendo emergere articolazione e forma dall’interno del suono, piuttosto che imporla dall’esterno.

Per Albert Bregman, studioso della percezione uditiva, la distinzione tra quei due momenti – individuazione delle proprietà delle unità minime e criteri per la loro distribuzione nel tempo – è di rilievo anche per lo studio della percezione di eventi sonori naturali dal carattere articolato e texturale (le foglie mosse dal vento, il battere della pioggia, il rotolare di pietre, il frangersi di masse d’acqua, il frantumarsi di corpi solidi, ecc.) (Bregman 1990, pp. 118-119). In proposito, Bregman menziona le proposte di Gabor, ma prende spunto soprattutto dalle modifiche arbitrarie che ne sono state date in contesti di sperimentazione musicale (egli si riferisce in particolare a Curtis Roads). Ed è notevole che certe ricerche che egli descrive, volte a studiare la percezione uditiva di eventi complessi e articolati – per esempio vetri che si rompono o certi eventi sonori naturali (Schumacher e Gilden 1993; Warren e Verbrugge 1984) – spingano i ricercatori a seguire una prassi di modellizzazione empirica del suono che somiglia molto al *bricolage* elettroacustico-musicale.

Considerando i molteplici sviluppi che abbiamo cercato di seguire, almeno nelle loro linee principali, dobbiamo concludere che le idee avanzate da Gabor tra il 1946 e il 1947 hanno una grande profondità di implicazioni, mentre allo stesso tempo si ricollegano al portato culturale e conoscitivo di alcune delle rivoluzioni scientifiche più importanti della storia moderna e contemporanea, come si è visto nella prima parte del nostro percorso. Gabor introduce nell’analisi del suono due condizioni di rilievo da un punto di vista epistemologico, tra loro separabili in un’ottica strettamente riduzionistica: la considerazione della finitezza temporale a livello delle rappresentazioni fisiche – e ciò evoca anche quella che poi è stata, in tempi più recenti, una profonda riconsiderazione della temporalità e dell’irreversibilità nella fisica dei sistemi dinamici lontani dall’equilibrio, nella teoria del caos, etc. (Prigogine e Stengers 1999) – e un riposizionamento del punto di osservazione sul piano concreto dell’esperienza. Una proposta simile non poteva non insinuarsi e ripercuotersi sul pensiero musicale, finendo allora per articolare un intreccio di scienza fondamentale, di tecnologia e di potenziale espressivo-musicale che si è dipanato nel corso dei decenni. Nel contesto della teoria musicale contemporanea forse i quanta acustici di Gabor sono ancora una sponda lontana e poco praticabile per molti: in un certo senso, non potrebbe essere altrimenti, dato che costituiscono un’eredità impegnativa, e che richiedono un rinnovamento nel modo di pensare il suono e di farne musica. Abbiamo cercato di evidenziare come da tempo tale eredità sia fertile e vitale per il pensiero e la prassi della ricerca scientifica e della sperimentazione musicale del nostro tempo.

## Bibliografia

- Arfib, Daniel (1991). «Analysis, transformation and resynthesis of musical sounds with the help of a time-frequency representation». In: *Representations of musical signals*. A cura di Giovanni De Poli, Aldo Piccialli e Curtis Roads. Cambridge Mass.: MIT Press, pp. 87–118.
- Balian, Roger (1981). «Un principe d'incertitude fort en théorie du signal ou en mécanique quantique». In: *Comptes Rendus de l'Académie des Sciences* 292, pp. 1357–1362.
- Bastiaans, Martin (1980). «Gabor's expansion of a signal into Gaussian elementary signals». In: *Proceedings of the IEEE*. 68, pp. 538–539.
- Beyer, Robert (1953). «Zur Geschichte der elektronischen Musik». In: *Melos* 10, pp. 278–280.
- Bozzi, Paolo (1990). «Mach e i fatti». In: *Nuova civiltà delle macchine* 8.1, pp. 49–54.
- Brandenburg, Karlheinz (1999). «MP3 and AAC explained». In: *Proceedings of the AES 17th International Conference on High-Quality Audio Coding*.
- Bregman, Albert (1990). *Auditory Scene Analysis*. Cambridge, Mass.: MIT Press.
- Brillouin, Léon (1956). *Science and Information Theory*. New York: Academic Press.
- Cascone, Kim (2000). «The Aesthetics of Failure: Post-Digital Tendencies in Contemporary Computer Music». In: *Computer Music Journal* 24.4, pp. 12–18.
- Cavaliere, Sergio e Aldo Piccialli (1997). «Granular synthesis of musical signals». In: *Musical signal processing*. A cura di Curtis Roads et al. London: Routledge, pp. 155–186.
- Chouard, Claude-Henri (2001). *L'oreille musicienne. Les chemins de la musique de l'oreille au cerveau*. Paris: Gallimard.
- Conway, Flo e Jim Siegelman (2005). *L'eroe oscuro dell'informazione. Alla ricerca di Norbert Wiener, il padre della cibernetica*. Torino: Codice Edizioni.
- Cramer, Leslie e Robert Talambiras (1971). «The Braided-Speech Method of time compressing speech». In: *Proceedings of the Second Louisville Conference on Rate and/or Frequency-controlled Speech, 1969*. [Ripreso in *Transactions of the Institute of Radio Engineers (Professional Group on Audio-2)*, pp. 7-12], pp. 206–224.
- De Poli, Giovanni e Aldo Piccialli (1991). «Pitch-synchronous granular synthesis». In: *Representations of musical signals*. A cura di Giovanni De Poli, Aldo Piccialli e Curtis Roads. Cambridge, Mass.: MIT Press, pp. 391–412.
- De Poli, Giovanni, Aldo Piccialli e Curtis Roads, cur. (1991). *Representations of musical signals*. Cambridge, Mass.: MIT Press.
- Debru, Claude (2003). «Ernst Mach et la psychophysiologie du temps». In: *Philosophia Scientiae* 7.2, pp. 59–91. URL: [http://archive.numdam.org/ARCHIVE/PHSC/PHSC\\_2003\\_\\_7\\_2/PHSC\\_2003\\_\\_7\\_2\\_59\\_0/PHSC\\_2003\\_\\_7\\_2\\_59\\_0.pdf](http://archive.numdam.org/ARCHIVE/PHSC/PHSC_2003__7_2/PHSC_2003__7_2_59_0/PHSC_2003__7_2_59_0.pdf) (visitato il 26/11/2015).
- D'Elia, Alfonsina (1971). *Ernst Mach*. Firenze: La Nuova Italia.
- Di Scipio, Agostino (1994). «Micro-time sonic design and the formation of timbre». In: *Contemporary Music Review* 10.2, pp. 135–148.
- (1997). «The problem of 2nd-order sonorities in Xenakis' electroacoustic music». In: *Organised Sound* 2/3, pp. 165–178.
- (1998). «Scienza e musica dei quanti acustici: l'eredità di Gabor». In: *Il Monocordo* 6, pp. 71–76.
- (2000). «Da un'esperienza in ascolto tra phoné e logos. Testo, suono e musica in *Thema (Omaggio a Joyce)* di Berio». In: *Il saggiautore musicale* 7.2, pp. 325–359.
- Di Scipio, Agostino (2001). «Uno scenario post-digitale? Riflessioni ambivalenti sul suono in Internet». In: *Musica/Realità* 65, pp. 121–156.

- (2003). «Métaphore, formalisation et intuition dans *Analogique A et B* de Xenakis». In: *Iannis Xenakis, Gérard Grisey. La métaphore lumineuse*. A cura di M. Solomos. Paris: L'Harmattan, pp. 181–192.
- (2006a). «Analysis of *Analogique B*». Inedito.
- (2006b). «Le nubi di suono ed i loro meccanismi. Uno studio di *Analogique A et B*». In: *Iannis Xenakis. Musicista scienziato architetto*. A cura di Alessandro Melchiorre. Quaderni di Ricerca IRMus 2. Numero speciale. Milano: Edizioni Scuole Civiche di Milano, pp. 129–160.
- (2015). «Stochastics and granular sound in Xenakis' electroacoustic music». In: *Proceedings of the Conference Iannis Xenakis. La musique électroacoustique - 2012*. A cura di M. Solomos. Paris: L'Harmattan, pp. 277–296.
- Dudley, Homer (1936). «Synthesizing Speech». In: *Bell Laboratories Record* 15, pp. 159–177.
- (1939). «Remaking Speech». In: *The Journal of the Acoustical Society of America* 11.2, pp. 169–177.
- Fairbanks, G., W. L. Everitt e R. P. Jaeger (1954). «Method for time or frequency compression-expansion of speech». In: *Convention Record of the Institute of Radio Engineers (part 8: Information Theory)*. [Ripreso in *Transactions of the Institute of Radio Engineers (Professional Group on Audio-2)*, pp. 7-12], pp. 120–124.
- Farge, Marie et al. (2012). «Les ondelettes et le CIRM». In: *Gazette des Mathématiciens* 131, pp. 47–57. URL: <http://wavelets.ens.fr/PUBLICATIONS/ARTICLES/PDF/319.pdf> (visitato il 30/11/2015).
- Feichtinger, Hans e Thomas Strohmer (1998). *Gabor Analysis and Algorithms*. Boston: Birkhäuser.
- Feyerabend, Paul (1987). *Farewell to reason*. London: Verso.
- Foulke, Emerson, cur. (1971). *Proceedings of the Second Louisville Conference on Rate and/or Frequency controlled Speech* (October 22-24, 1969). Louisville, Kentucky.
- 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.
- (1947a). «Acoustical quanta and the theory of hearing». In: *Nature* 159.4044, pp. 591–594.
- (1947b). «New possibilities in speech transmission». In: *Journal of the Institution of Electrical Engineers* 94.32, pp. 369–387.
- (1947c). «Theory of communication». In: *Journal of the Institution of Electrical Engineers* 94.73, p. 58.
- (1948a). «A New Microscopic Principles». In: *Nature* 161.4098, pp. 777–778.
- (1948b). «Recherches sur quelques problèmes de télécommunication et d'acoustique». In: *L'Onde électrique* 257, pp. 433–439.
- (1949). «Improvements in and relating to transmission, recording and reproduction of sound». British Patent 616,156 (London). January 18, 1949. URL: <http://worldwide.espacenet.com/publicationDetails/biblio?CC=GB&NR=616156> (visitato il 05/12/2015).
- (1951). «La théorie des communications et la physique». In: *La cybernétique. Théorie du signal et de l'information*. A cura di Louis de Broglie. Paris: Éditions de la Revue d'Optique, pp. 115–149.
- (1952). *Lectures on communication theory*. Technical Report 238. Cambridge, Mass.: MIT Research Laboratory of Electronics.
- (1979). *Olografia*. In: *Enciclopedia del Novecento*. Vol. 4. Roma: Istituto Encicopedico Italiano Treccani.

- Glaserfeld, Ernst von (1975). «Piaget and the Radical Constructivist Epistemology». In: *Epistemology and Education*. A cura di C. D. Smock e E. von Glaserfeld. Athens Ga.: Follow Through Publications, pp. 1–24.
- Goldmark, Peter e P. S. Hendricks (1939). «Synthetic Reverberation». In: *Proceedings of the Institute of Radio Engineers*. Vol. 27. 12, pp. 747–752.
- Heisenberg, Werner (1927). «Über den anschaulichen Inhalt der quantumtheoretische Kinematik und Mechanik». In: *Zeitschriften für Physik* 43. Trad. it. in (Heisenberg 1991), pp. 172–198.
- (1991). *Indeterminazione e realtà*. Napoli.
- Helmholtz, Hermann L. F. (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Braunschweig: Friedrich Vieweg und Sohn. Google Books: Ve09AAAAcAAJ.
- Helstrom, Carl (1966). «An expansion of a signal in Gaussian elementary signals». In: *IEEE Transactions in Information Theory* 12, pp. 81–82.
- Huggins, A. W. F. (1974). *More Temporally Segmented Speech: Is Duration or Speech Content the Critical Variable in Its Loss of Intelligibility?* Quarterly Progress Report 114. Cambridge, Mass.: MIT Research Laboratory of Electronics, pp. 185–193.
- Jakobson, Roman (1963). *Essais de linguistique générale*. Paris: Minuit.
- (1966). «Linguistica e teoria dell'informazione». In: *Saggi di linguistica generale*. Ed. or. in (Jakobson 1963). Milano, pp. 65–76.
- Jones, Douglas e Thomas Parks (1988). «Generation and Combination of Grains for Music Synthesis». In: *Computer Music Journal* 12.2, pp. 27–34.
- Kronland-Martinet, Richard e Alex Grossmann (1991). «Application of time-frequency and time-scale methods – wavelet transforms – to the analysis, synthesis and transformation of natural sounds». In: *Representations of musical signals*. A cura di Giovanni De Poli, Aldo Piccialli e Curtis Roads. Cambridge, Mass.: MIT Press, pp. 45–85.
- Kronland-Martinet, Richard, Philippe Guillemain e Sølvi Ystad (1997). «Modelling of natural sounds by time-frequency and wavelet representations». In: *Organised Sound* 2.3, pp. 179–191.
- Leprince-Ringuet, Louis (1981). «Sur Iannis Xenakis». In: *Regards sur Iannis Xenakis*. A cura di M. Fleuret. Paris: Stock, pp. 52–54.
- Lietti, Alfredo (1956). «Gli impianti tecnici dello Studio di Fonologia Musicale di Radio Milano». In: *Elettronica* V.3, pp. 116–121.
- (1998). «Gli impianti tecnici dello Studio di Fonologia Musicale di Radio Milano». In: *Elettronica e comunicazioni* 2/3. Ed. or. (Lietti 1956), pp. 71–78. URL: <http://www.crit.rai.it/eletel/Important/Anni1990/1998-2e3.pdf> (visitato il 30/11/2015).
- Lindsay, Robert (1966). «The Story of Acoustics». In: *Journal of the Acoustical Society of America* 39.4, 629644.
- Mach, Ernst (1865a). «Bemerkungen über die Accomodation des Ohres». In: *Sitzungsberichte der Wiener Akademie der Wissenschaften: Mathematisch-Naturwissenschaftliche* 51, pp. 343–346.
- (1865b). «Untersuchungen über den Zeitsinn des Ohres». In: *Sitzungsberichte der Wiener Akademie der Wissenschaften: Mathematisch-Naturwissenschaftliche* 51, pp. 133–150.
- (1875). *Grundlinien der Lehre von den Bewegungsempfindungen*. Lipsia.
- Mach, Ernst (1886). *Beiträge zur Analyse der Empfindungen*. Jena: Verlag von Gustav Fischer. URL: <http://www.archive.org/details/beitrgezuranaly00machgoog>.

- (1893). *Science of Mechanics. A Critical and historical account of its development.* London: Open Court Pub Co.
- (1914). *Analysis of Sensations.* Ed. or. (Mach 1886). London: Open Court Pub Co.
- MacKay, John (1984). «On the Perception of Density and Stratification in Granular Sonic Textures: An Exploratory Study». In: *Interface - Journal of New Music Research* 13.4, pp. 171–186.
- Mathien, Michel (2007). «Abraham Moles: affronter scientifiquement la quotidieté de la communication humaine». In: *Hermès* 48, pp. 101–108. URL: <http://www.cairn.info/revue-hermes-la-revue-2007-2-page-101.htm> (visitato il 04/12/2015).
- Meyer, Yves (1993). *Wavelets. Algorithms and applications.* Philadelphia: SIAM (Society for Industrial e Applied Mathematics).
- (2008). *Tribute to Jean Morlet.* Testo di un intervento presentato alla conferenza «Continuous Wavelet Transform and Morlet's wavelet: International conference in honor of Jean Morlet». URL: <http://citeseerx.ist.psu.edu/viewdoc/download?jsessionid=EA924BFB86A8499B4231EDF8E37F3EBA&doi=10.1.1.397.6175&rep=rep1&type=pdf> (visitato il 30/11/2015).
- Meyer-Eppler, Werner (1955). «Statistische und psychologische Klangprobleme». In: *die Reihe* 1, pp. 22–28.
- (1959). *Grundlagen und Anwendungen der Informationstheorie.* Kommunikation und Kybernetik in Einzeldarstellungen 1. Berlin: Springer-Verlag.
- Moles, Abraham (1953a). «Sur la coloration des bruits blancs en acoustique musicale». In: *L'Onde Électrique*, pp. 285–286.
- (1953b). «Théorie de l'information, électronique et cybernétique». In: *L'Onde Électrique*, pp. 637–651.
- (1957). «Machines à musique: l'apport des machines électroniques et électroacoustiques à la nouvelle sensibilité musicale». In: *La Revue Musicale* 236. Il testo era stato presentato col titolo *Machines à musique: du Phonogène au Vocoder*, nel contesto della "Décade de musique expérimentale", Parigi, 1952, pp. 115–127.
- (1961a). *Les Musiques expérimentales.* Trad. da D. Charles. Zurigo: Cercle d'art contemporain.
- (1961b). *Théorie de l'information et perception esthétique.* Paris: Flammarion.
- (1969). *Teoria dell'informazione e percezione estetica.* Ed. or. (Moles 1961b). Roma: Lerici Editore.
- Morawska-Büngeler, Marietta (1988). *Schwingende Elektronen: eine Dokumentation über das Studio für Elektronische Musik des Westdeutschen Rundfunks in Köln.* Köln: Tonger Verlag.
- Morlet, Jean, Richard Kronland-Martinet e Alex Grossmann (1987). «Analysis of sound pattern through wavelet transforms». In: *International Journal of Pattern Recognition and Artificial Intelligence* 1.2, pp. 273–301.
- Oldroyd, David (1989). *Storia della filosofia della scienza.* Milano: Il Saggiatore.
- Orcalli, Angelo (1993). *Fenomenologia della musica sperimentale.* Potenza: Sonus Edizioni Musicali.
- Poullin, Jacques (1954). «L'apport des techniques d'enregistrement dans la fabrication de matières et formes musicales nouvelles. Applications à la musique concrète». In: *L'Onde Électrique* 324, pp. 282–291.

- Pöppel, Ernst, Yan Bao e Biz Zhou (2011). «Temporal Windows as Logistical Basis for Cognitive Processing». In: *Advances in Psychological Science* 19.6, pp. 775–793.
- Prigogine, Ilya e Isabelle Stengers (1999). *La nuova alleanza. Metamorfosi della scienza*. Torino: Einaudi.
- Restagno, Enzo (1988). «Un'autobiografia dell'autore raccontata da Enzo Restagno». In: *Xenakis*. A cura di Enzo Restagno. Torino: EDT/Musica, pp. 3–70.
- Riegler, Alexander (2012). «Constructivism». In: *Paradigms in Theory Construction*. A cura di L. L'Abate. Springer, pp. 235–255.
- Risset, Jean-Claude e David Wessel (1982a). «Exploration of Timbre by Analysis and Synthesis». In: *Psychology of Music*. A cura di Diana Deutsch. Riedizione 1999, pp. 113–169. Trad. it. (Risset e Wessel 1982b). San Diego, California: Academic Press, pp. 25–58.
- (1982b). «Indagine sul timbro mediante analisi e sintesi». In: *Bollettino del Laboratorio di Informatica Musicale della Biennale di Venezia*. 2. Ed. or. (Risset e Wessel 1982a). Venezia: La Biennale, pp. 12–35.
- Roads, Curtis (1978). «Automated granular synthesis of sound». In: *Computer Music Journal* 2.2, pp. 61–62.
- (1991). «Asynchronous granular synthesis». In: *Representations of musical signals*. A cura di Giovanni De Poli, Aldo Piccialli e Curtis Roads. Cambridge, Mass.: MIT Press, pp. 143–185.
- (2001a). *Microsound*. Cambridge, London: The MIT Press.
- (2001b). «Sound composition with pulsars». In: *Journal of the Audio Engineering Society* 49.3, pp. 134–147.
- Rodet, Xavier (1984). «Time-domain Formant Wave Function Synthesis». In: *Computer Music Journal* 8.3, pp. 9–14.
- Schaeffer, Pierre (1952). *A la recherche d'une musique concrète*. Paris: Éditions du Seuil.
- (1966). *Traité des objets musicaux. Essai interdisciplines*. Paris: Éditions du Seuil.
- Schumacher, Mark e David Gilden (1993). «Auditory Perception of Fractal Contours». In: *Journal of Experimental Psychology* 19.3, pp. 641–660.
- Shannon, Claude e Warren Weaver (1949). *The Mathematical Theory of Communication*. Trad. it. (Shannon e Weaver 1971). Urbana-Champaign: University of Illinois.
- (1971). *La teoria matematica delle comunicazioni*. Milano: Etas Kompass.
- Sigismondi, Renato (2002). *La teoria della conoscenza di Ernst Mach*. Chieti: Tabula Fati Editore.
- Springer, Anton (1955). «Ein akustischer Zeitregler». In: *Gravesaner Blätter* 1, pp. 32–37.
- Stewart, George Walter (1931). «Problems suggested by an uncertainty principle in acoustics». In: *Journal of the Acoustical Society of America* 2, pp. 325–329.
- Thomson, Phil (2004). «Atoms and errors: towards a history and aesthetics of microsound». In: *Organised Sound* 9.2, pp. 207–218.
- Truax, Barry (1988a). «Discovering inner complexity: Time-shifting and transposition with a real-time granulation technique». In: *Computer Music Journal* 18.2, pp. 38–48.
- (1988b). «Real-time granular synthesis with a digital signal processor». In: *Computer Music Journal* 12.2, pp. 14–26.
- (1990). «Composing with Real-Time Granular Sound». In: *Perspectives of New Music* 28.2, pp. 120–134. JSTOR: 833014.
- Ungeheuer, Elena (1992). *Wie die elektronische Musik »erfunden« wurde...* Mainz: Schott.

- Ville, Jean (1948). «Théorie et applications de la notion de signal analytique». In: *Cables et Transmissions* 2.1, pp. 61–74. URL: <https://archive.org/details/VilleSigAnalytiqueCablesEtTrans1948Fr> (visitato il 30/11/2015).
- Warren, William e Robert Verbrugge (1984). «Auditory perception of breaking and bouncing events. A case in ecological acoustics». In: *Journal of Experimental Psychology* 10.3, pp. 704–712.
- Wiener, Norbert (1964). *I Am a Mathematician: The Later life of a Prodigy*. Cambridge, Mass.: MIT Press.
- Xenakis, Iannis (1962). «Stochastic Music». In: *Gravesaner Blätter* 23-24, pp. 169–184.
- (1963). *Musiques formelles*. Revue Musicale 253-254. Numero monografico. Paris.
- (1971). *Formalized Music*. Trad. da C. Butchers, G.H. Hopkins e J. Challifour. Blooming-ton: University Press.
- (1976). *Musique. Architecture*. Tournai: Casterman.
- (1982). *Musica. Architettura*. Ed. or. (Xenakis 1976). Milano: Spirali.
- (1985). «Music composition treks». In: *Composers and the computer*. A cura di Curtis Roads. Los Altos CA: Kaufmann, pp. 184–185.
- (1992). *Formalized Music*. A cura di Sharon Kanach. Versione rivista e ampliata di (Xenakis 1971). Stuyvesant (New York): Pendragon Press.
- (2003). *Universi del suono. Scritti e interventi 1955-1994*. A cura di Agostino Di Scipio. Lucca/Milano: LIM/Le Sfere.
- Young, Randy (1993). *Wavelet theory and its application*. Norwell Mass.: Kluwer Academic.
- Zavagna, Paolo (1992). «*Thema (Omaggio a Joyce)* di Luciano Berio: un'analisi». In: *I quaderni della Civica Scuola di Musica* 21-22, pp. 58–64.
- (2013). «La voce senz'anima: origine e storia del Vocoder». In: *Musica/Tecnologia* 7, pp. 27–63. URL: <http://www.fupress.net/index.php/mt/article/view/13206>.

# Xenakis e il “circolo granulare”

Makis Solomos\*

(Originale del 2006, traduzione ricevuta il 7 ottobre 2015<sup>1</sup>)

## Sommario

Molte sono le strade che partono da Iannis Xenakis e che ne attraversano l’opera. Una di esse è quella che indichiamo, in questo articolo, come “paradigma granulare”<sup>2</sup>. Inizierò appunto da Xenakis, per suggerire che, nella sua estetica, l’approccio granulare ha lo statuto di una vera e propria “teoria”, nel senso antico della parola, e per verificare quali siano gli elementi che ne fanno appunto una “visione”. Proverò poi ad inserire il paradigma granulare in una dimensione di storicità musicale, per verificarne il divenire in altre esperienze estetiche, in particolare nella musica dei compositori Horacio Vaggione e Agostino Di Scipio.

## 1 Xenakis e il paradigma granulare

### 1.1 Una “teoria”

Xenakis espone il paradigma granulare come una «ipotesi di base (lemma)» (Xenakis 1960a, p. 84)<sup>3</sup> nella prima parte dell’articolo intitolato «Elements of Stochastic Music», che finisce di scrivere nel 1959 e che pubblica nel 1960, noto anche come parte integrante del suo libro *Musiques formelles* (Xenakis 1963, Cap. 2) e quindi anche della sua traduzione inglese, *Formalized Music* (1971 e 1992)<sup>4</sup>. Mi permetto di citare per esteso tale «lemma», riprendendolo dalla versione inglese iniziale (*ibid.*, pp. 86-87), ma segnalando in parentesi quadre le differenze più importanti che si trovano in Xenakis (1992, pp. 43-44), e riportando

---

\*Université Paris 8.

<sup>1</sup>Il presente testo è la traduzione dell’intervento tenuto a «The Creative and Scientific Legacies of Iannis Xenakis International Symposium», organizzato da James Harley (University of Guelph), Michael Duschenes (Perimeter Institute for Theoretical Physics), Thomas Salisbury (Fields Institute for Research in Mathematical Sciences), nel giugno 2006. La sezione 2.2 è ulteriormente sviluppata in Solomos (2005a) e la sezione 2.3 è sviluppata in Solomos (2005d), poi ripresa in traduzione italiana in Solomos (2013).

<sup>2</sup>Userò alternativamente i termini “paradigma” e “approccio”, seguendo l’esempio di Vaggione (2005).

<sup>3</sup>In Xenakis (1963, p. 61), si parla di «hypothèse de base (Lemme)». In Xenakis (1992, p. 43), si parla di «basic temporary hypothesis (lemma)».

<sup>4</sup>Il Cap. 2 di *Musiques formelles* e di *Formalized Music* è composto da quattro articoli inizialmente apparsi in inglese e tedesco nella rivista «Gravesaner Blätter» (Xenakis 1960a,b, 1961a,b). La versione inglese reca il titolo complessivo «Elements of Stochastic Music» (quella tedesca si intitola «Grundlagen einer stochastischen Musik»). Rispetto ai quattro articoli originali, il capitolo 2 di *Musiques formelles* presenta differenze significative. Alcune parti sono del tutto nuove, e tra queste la più importante è certamente la parte finale (Xenakis 1963, pp. 108 e sgg.) e (Xenakis 1992, pp. 89 e sgg.), che probabilmente Xenakis avrebbe voluto pubblicare in un numero successivo di «Gravesaner Blätter», quando però ormai l’uscita di *Musiques formelles* era già pianificata. Purtroppo negli Archives Xenakis presso la Bibliothèque Nationale de France, a Parigi, non sembra reperibile alcuna versione originale francese, ammesso che esista, tranne che per il «Prologue» (Xenakis 1960a, pp. 84-86).

anche, in corsivo, i termini di *Musiques formelles* che appaiono divergenti rispetto a *Formalized Music* (Xenakis 1963, p. 61)<sup>5</sup>:

All sounds represent an integration of corpuscles [grains], of elementary acoustic particles, of sound quanta. Each of these elementary particles possesses a double nature: the frequency and the intensity (the life-time of each corpuscle being minimum and invariable) [Each of these elementary grains has a threefold nature: duration, frequency, and intensity]. Every sound, every even continuous variation of sound is to be understood as an assembly of a sufficient number [of a large number / suffisamment nombreux] of elementary particles [grains] being disposed adequately within the time level. Thus any sound complex can be analyzed into series of pure sinusoidal tones, even if the variations of these latter are infinitely close together, of short duration and complex. During the attack of a complex sound, its full chord [body] and its decay, thousands of pure tones appear within a very short interval  $\Delta t$  [in a more or less short interval of time,  $\Delta t$  / dans un intervalle de temps  $\Delta t$  assez court]. Hecatombs of pure sounds are necessary for the creation of a complex sound. This one should be imagined as to be a display of fireworks sparkling in all colours, each luminous point of which appearing and disappearing instantaneously on the background of the black sky. But there would be so many luminous spots in this firework and they would be organized in such a way that their rapid and swarming succession creates forms, slowly uncoiling volutes<sup>6</sup> or in the contrary short explosions inflaming the whole sky. A sufficient number of instantaneously appearing and disappearing spots would form a luminous line.

Alla fine della seconda frase c’è un rinvio a nota che è abbastanza importante segnalare:

This description of the micro-structure of acoustic signals [of the elementary structure of the sonic symbols / signaux] is used as the starting-point of the musical realization and must be understood rather as to be an intuitive representation than of scientific consistence [and is consequently only a hypothesis, rather than an established scientific fact / et n'est par conséquent qu'une image plutôt qu'un fait scientifiquement fondé]. But it can be considered as to be a first approach [approximation / approche] towards the ideas introduced into the theory of information by Gabor...

Torneremo su Gabor più tardi. Per il momento, mettiamo l’accento sul fatto che, in un primo tempo, l’approccio granulare per Xenakis costituisce dunque una «rappresentazione intuitiva» o (nei termini di *Musiques formelles*) una «immagine», e che solo nel 1971 diventa una «ipotesi». Penso che dalla differente formulazione si possa derivare che il modo stesso di presentare il paradigma granulare fosse, per Xenakis come forse per altri compositori, di grande importanza. Intendo dire, per limitarci ora a Xenakis, che parlare di una «rappresentazione intuitiva» o di una «immagine», pur non implicando certo solo una suggestione “poetica”, allo stesso tempo è ben più di un tecnicismo per indicare una forma di sperimentazione. Non dico che vi sia netta opposizione tra una dimensione intuitiva e una di razionalità pratica (o astratta). Dico, al contrario, che «rappresentazione intuitiva» presenta entrambi gli aspetti allo

<sup>5</sup> Vi sono altre differenze, ma possiamo ritenerle dovute solo all’uso delle due diverse lingue, o anche dovute al fatto che, assai probabilmente, il traduttore di *Formalized Music* ha operato direttamente a partire da *Musiques formelles*. Pertanto, la sequenza cronologica delle circostanze sarebbe stata: a) probabile manoscritto francese dei quattro articoli; b) traduzioni inglesi e tedesche nei quattro numeri di «Gravesaner Blätter»; c) ripresa del probabile originale francese per *Musiques formelles*, con importanti modifiche; d) traduzione inglese a partire da *Musiques formelles*, con ulteriori modifiche.

<sup>6</sup>L'espressione «forms and spirals, slowly unfolding» (in *Formalized Music*) è probabilmente una cattiva traduzione del francese «des formes, des volutes à déroulement lent» (*Musiques formelles*).

stesso tempo e finisce per agganciare anche altre dimensioni, musicali, scientifiche, perfino filosofiche, e ancora altre esperienze umane. Si tratta in fondo di una visione del mondo. In altri termini, di una teoria nel senso antico della parola *θεωρία* (dal greco *θεωρεῖν* cioè “vedere”, “contemplare”, “pensare”)<sup>7</sup>.

Nella produzione musicale di Xenakis, c’è solo una composizione che possa essere riferita con certezza all’ipotesi granulare, cioè *Analogique A et B*, opera che è infatti al centro del discorso formulato in «Elements of Stochastic Music», e che in pratica risulta dalla sovrapposizione di due lavori, cioè *Analogique A* (1958), per nove archi impegnati costantemente nella produzione di suoni brevi, pizzicati e battuti col legno (Fig. 1) e *Analogique B* (1959), per soli suoni elettronici, idealmente sinusoidali. È probabile che per Xenakis si trattasse di un

**Figura 1:** Iannis Xenakis, *Analogique A* (1958), battute 0-2 (© Éditions Salabert).

momento di pura sperimentazione, visto che, nel capitolo 2 di *Musiques formelles*, scrive a proposito di *Analogique A*: «the hypothesis [...] cannot, therefore, be confirmed or invalidated under these conditions» (Xenakis 1992, p. 103) (Xenakis 1963, p. 122)<sup>8</sup>. Forse fu proprio in base a tale constatazione negativa che egli non riprovò l’esperimento<sup>9</sup>. Si cita spesso in

<sup>7</sup>Si veda anche Solomos (2005b), dove si tratta dello statuto “teorico” degli automi cellulari usati da Xenakis.

<sup>8</sup>Questo passaggio, presente nel libro di Xenakis, non appare nei suoi articoli su «Gravesaner Blätter».

<sup>9</sup>C’è stato un certo dibattito sul “fallimento” di *Analogique A et B*. Per esempio, Agostino Di Scipio ha scritto: «...*Analogique* può ragionevolmente essere ritenuto un brano problematico. Alcuni lo considerano il meno riuscito che Xenakis abbia mai composto. Di solito la circostanza viene spiegata col marcato accento che egli mise sui dettagli matematici e costruttivi, insomma con un eccessivo sovraccarico di premesse teoriche e tecniche recante esiti musicali piuttosto poveri. In breve, in tal senso il brano andrebbe visto come un esperimento insoddisfacente. Il che, in un certo senso, è semplicemente vero, prova ne sia che Xenakis in effetti non è mai tornato a seguire un simile approccio in lavori successivi. Ma tutto ciò dice assai poco. Direi piuttosto che le problematiche affrontate da Xenakis in *Analogique A et B*, e che certamente lasciò senza soluzioni definitive, diano a questa musica un carattere peculiare molto palpabile nell’esperienza d’ascolto [...], esse hanno a che fare esattamente con aspetti del procedimento compositivo che rimanevano (forse necessariamente) non “formalizzati”. Infatti Xenakis si sentì indotto a introdurre un certo numero di aggiustamenti manuali, cioè non previsti dai processi che aveva formalizzato. All’ascolto, il risultato costituisce non tanto un lavoro musicalmente insoddisfacente, quanto il segno espressivo di una dialettica

proposito la musica concreta di *Concret PH*, basata sul suono di braci ardenti. Essa però fu composta subito prima di *Analogique A et B*, nel 1958, e probabilmente la sua lavorazione era in corso nel 1957<sup>10</sup>. *Concret PH* inoltre va considerato nel più ampio contesto dell’architettura concepita da Xenakis per il famoso padiglione Philips dell’Expo di Bruxelles nel 1958<sup>11</sup>.

Tuttavia, ad indicare meglio come l’approccio granulare costituisca per Xenakis una visione teorica più ampia, e non solo una forma di sperimentazione, c’è l’evidente dimensione di granulosità caratteristica del suono di molti suoi lavori, almeno in determinate sequenze musicali: insomma, la sua musica manifesta indubbiamente una “sensibilità granulare” (accanto ad altre)<sup>12</sup>. Sarebbe troppo lungo elencare tutti i lavori che lo dimostrano. Diciamo solo che questa dimensione granulare così spesso caratteristica della sua musica rinvia a grandi masse di suoni brevi, a «nuvole di suoni» (per usare un’espressione dello stesso Xenakis, che peraltro si è assai diffusa)<sup>13</sup>. In Fig. 2 e in Fig. 3 propongo due esempi di musica strumentale che all’ascolto conseguono evidenti effetti di granularità, presi da *Herma* (per pianoforte, 1960-61) e da *Eridanos* (per orchestra, 1972).

Le prime sequenze di suono granulare in Xenakis le troviamo già in *Pithoprakta* (per orchestra d’archi, due tromboni, percussione, del 1955-56), dunque già prima di *Concret PH* e di *Analogique A et B*. Molti conoscono le battute 52-59 di *Pithoprakta*, nelle quali oltre mille pizzicati-glissandi degli archi vengono eseguiti in poco meno di venti secondi<sup>14</sup>. Xenakis calcolò i valori (di altezza iniziale, di durata, di velocità di glissando) secondo leggi di probabilità e ottenne un grafico della loro distribuzione (Fig. 4)<sup>15</sup>.

vitale e intricata tra formalizzazione e intuizione» (Di Scipio 2005b, p. 96).

<sup>10</sup>In una lettera che Xenakis invia a Hermann Scherchen il 10 Luglio 1958 (conservata negli Archives Xenakis, Dossier (Euvre 3/4), leggiamo: «a dispetto di tutte le difficoltà materiali del ‘sacro club’ di Schaeffer, sto provando a terminare i pochi minuti che avevo iniziato l’anno scorso». Segnalo anche una seconda lettera, assai divertente, inviata il 30 Novembre 1957 da Louis Kalff (responsabile tecnico della realizzazione del padiglione Philips) a Xenakis, in cui leggiamo: «Sappiamo che avete cercato un rumore come il crepitio che fanno i carboni del legno che si raffreddano. Abbiamo ascoltato qualche settimana fa un concerto elettronico ad Eindhoven. Per uno dei pezzi, la maggior parte del pubblico seguiva un testo di qualche pagina, stampato su carta sottile ma dura. Il rumore fatto da queste centinaia di pagine, voltate nello stesso momento, era molto curioso e assomigliava molto a della musica. Forse potete trarne qualche elemento di interesse» (Archives Xenakis, Dossiers Architecture X(A) 4/1).

<sup>11</sup>In un documento inedito presente negli Archive Xenakis, si trova questa nota su *Concret PH*: «opera commissionata da Philips per il suo padiglione all’Esposizione Universale di Bruxelles, 1958. *Concret P.H.* doveva preparare psicologicamente il pubblico allo spettacolo elaborato da Le Corbusier e accompagnato da musica di Varèse. I 400 altoparlanti che tappezzavano l’interno del padiglione dovevano riempire lo spazio con questo scintillio sonoro, come una emanazione comune di architettura e musica, concepite come una sola cosa: in effetti, la rugosità del cemento e il suo coefficiente di attrito trovavano in qualche modo una eco nel timbro di tale scintillio sonoro. La continuità statistica degli sviluppi formali di queste sonorità granulari e la stereofonia cinematica dovevano stabilire un’omologia completa tra il suono e le forme plastiche del padiglione, di cui avevo io stesso concepito interamente l’architettura secondo superfici regolate come ‘paraboloidi iperbolici (P.H.)».

<sup>12</sup>L’espressione «sensibilità granulare» è di Horacio Vaggione, che la usa di solito in relazione alla propria musica (Vaggione 2005, p. 348).

<sup>13</sup>Sulla nozione di “sonorità” nella musica di Xenakis, e in particolare sulle tre principali sonorità che vi si ritrovano (suoni glissati, suoni statici e prolungati, e masse di punti sonori), si veda Solomos (1993), Solomos (1996, Cap. 5), e Solomos (2001).

<sup>14</sup>Secondo quanto riportò inizialmente il compositore, si dovrebbe trattare esattamente di 1142 brevi eventi sonori (Xenakis 1956, p. 31). Secondo quanto riportato in *Musiques formelles* ve ne sarebbero 1148 (Xenakis 1963, p. 30) (Xenakis 1992, p. 15). La mia personale verifica ne conta 1146.

<sup>15</sup>Xenakis propose due grafici di questa sezione di *Pithoprakta* (Gibson 1994). Quello ripreso qui, tratto da Xenakis (1963, p. 31), è il secondo e definitivo.



**Figura 2:** Iannis Xenakis, *Herma* (1960-61), pagina 1 (© Boosey and Hawkes).

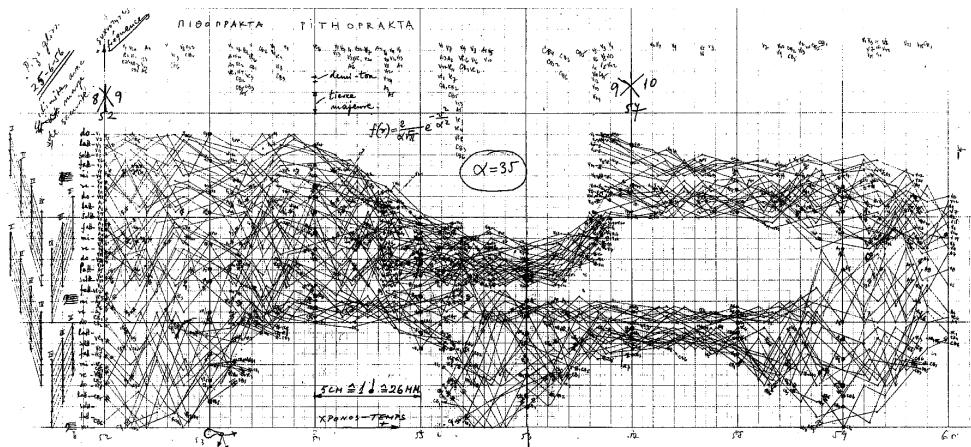


**Figura 3:** Iannis Xenakis, *Eridanos* (1972), battute 37-40 della parte degli ottoni (© Éditions Salabert).

## 1.2 Elementi costitutivi della teoria

Cerchiamo ora di capire quali siano gli elementi costitutivi della “teoria” (o visione) che porta Xenakis al paradigma granulare. Con “elementi costitutivi” intendo fattori che abbiano contribuito alla sua nascita, ma anche elementi che ne abbiano determinato le finalità o che ne costituiscano uno sfondo filosofico in senso ampio. Considereremo sei elementi in particolare:

1. Dualità onda-corpuscolo. Xenakis era a conoscenza del dibattito della fisica del tempo. Allo stesso tempo, voleva che la musica si “sintonizzasse” col dibattito scientifico.



**Figura 4:** Iannis Xenakis, *Pithoprakta* (1955-56), grafico manoscritto della distribuzione di suoni pizzicati-glissati nelle battute 52-59. Da Xenakis (1963, p. 31).

Sembra quindi piuttosto naturale che abbia pensato di trasporre l’idea di “corpuscolo” in musica<sup>16</sup>. Una delle prime manifestazioni (o forse proprio la prima) di questa idea è nell’articolo del 1955 intitolato «La crisi della musica seriale», laddove, parlando della natura discreta delle divisioni seriali, Xenakis scrive:

Pourquoi la continuité du spectre des fréquences? du spectre des timbres? du spectre des intensités et des durées? Mais laissons de côté la question de la continuité (elle sera d’ailleurs dans peu de temps, pour la recherche musicale, le pendant de l’état ondulatoire du corpuscule-onde de la matière) [...] (Xenakis 1955, p. 3)<sup>17</sup>.

2. Nozioni quali «massa», «nuvola di suoni», e «probabilità» sono tra i più importanti elementi che si debbano considerare in questo contesto. La sezione granulare di *Pithoprakta*, sopra citata, è una massa probabilistica di numerosi brevi eventi sonori. L’idea in effetti è già inscritta nell’uso che Xenakis fa delle probabilità nei primi anni del suo percorso, cioè nella “parabola” (Xenakis 1958-59) che lo condusse ad usare le leggi probabilistiche nella composizione:

Identifichiamo dunque dei suoni molto brevi – per esempio dei suoni in pizzicato – con le molecole: otteniamo una trasformazione omomorfica dal campo della fisica al campo dei suoni. Il movimento individuale dei singoli suoni non conta più (ibid., p. 19).

<sup>16</sup>Quando si mise ad illustrare l’evoluzione parallela di musica e scienza – negli anni settanta (Xenakis 1976, pp. 192-196) – il compositore non fece menzione a quel che qui stiamo chiamando “paradigma granulare”. Ciò può apparire strano, però va considerato che, in quel parallelo tra musica e scienza egli in realtà prese in considerazione, per quanto riguarda l’ambito scientifico, i soli sviluppi della matematica.

<sup>17</sup>Nella traduzione italiana (Xenakis 2003, p. 28), il passaggio è così reso: «Perché [non pensare piuttosto alla] continuità dello spettro delle frequenze? dello spettro dei timbri? dello spettro delle intensità e delle durate? Ma lasciamo ora da parte la questione della continuità (che diventerà presto, per la ricerca musicale, il corrispettivo dello stato ondulatorio nel dualismo corpuscolo/onda della materia)».

Xenakis descrive *Concret PH* come una «nuvola di polveri sonore» [*nuage de poussières de sons*], o anche come un «gas sonoro» (Delalande 1997, p. 115). Ciò in effetti vale come definizione delle masse probabilistiche di brevi suoni che si trovano in tanti altri suoi brani. Le questioni che vengono sollevate da una concezione probabilistica – densità, ordine e disordine, continuità e discontinuità – rientrano tutte a pieno titolo nell’approccio granulare.

3. Pluralità di livelli. Ciò che è molto specifico del paradigma granulare è l’idea di poter passare da un “livello” ad un altro della struttura sonora e musicale. Naturalmente, è proprio questa idea che porterà alla “sintesi granulare” in anni successivi, come realizzazione dell’approccio storicamente delineato nelle prime esperienze di Xenakis. In «Elements of Stochastic Music», l’idea viene espressa con queste due formulazioni generali:

- creare «complessi sonori» [*sound complexes*], o «suoni complessi» [*complex sounds*];
- creare sonorità di ordine superiore.

A questo proposito, conviene riprendere di nuovo il confronto tra il testo inglese originale (Xenakis 1960a, p. 90) e le sue modifiche successive (Xenakis 1992, p. 47):

Supposing that every point of these clouds represents not only a pure frequency and its satellite intensity but already a structure of elementary particles [grains] being arranged a priori. We think that a sonority of second order and even third order etc. [second, third, or higher order] could be created in this way<sup>18</sup>.

Vale la pena notare che il compositore vide in questa idea un’effettiva tecnica di sintesi del suono, anche se non usò una tale terminologia. Lo si vede anche dalla seguente citazione, presa in prossimità del passaggio appena citato, che mostra l’utopia di una sintesi del suono valida per ogni suono possibile, e con ciò anche l’impulso utopico caratteristico di molti pionieri della sintesi del suono:

Within human limits and with the manipulations of all sort of these particle clouds [grain clusters], we can hope to produce not only the sounds of classical instruments, of elastic bodies and, generally spoken, of those ones utilized with predilection by the concrete music, but also acoustic emotions [sonic perturbations / *ébranlements sonores*<sup>19</sup>] with evolutions which have been unprecedented and unimaginable up to now.

Forse è proprio qui che appare per la prima volta la distinzione tra micro- e macrostruttura (e anche tra micro- e macrosuono<sup>20</sup>), anche se la sua teorizzazione non viene sviluppata negli scritti successivi di Xenakis<sup>21</sup>.

<sup>18</sup>L’ultima modifica segnalata («second, third, or higher orders») non appare nella versione francese (Xenakis 1963, p. 65).

<sup>19</sup>Il termine inglese «perturbation» non è una buona traduzione del francese «ébranlement» (*ibid.*, p. 65), che è più vicino al termine usato nella versione originale: «émotion». Si tratta insomma di movimento, di “motion”, pur senza implicazioni psicologiche.

<sup>20</sup>Troviamo il termine «microsound» in tutte le tre versioni del testo, cioè in Xenakis (1960a, p. 93), e poi anche in Xenakis (1992, p. 50) e Xenakis (1963, p. 68)

<sup>21</sup>Per esempio, in Xenakis (1985, p. 127) leggiamo: «La musica contiene forme di ogni sorta, a livello microscopico, a livello macroscopico, a livello del singolo campione digitale» [riprendiamo la traduzione dalla traduzione italiana contenuta in *Universi del suono*, p. 96 (cfr. nota 17)].

4. Un altro elemento piuttosto specifico sta anche nell'idea che l'ipotesi granulare possa valere non solo per la natura del suono ma anche per il senso dell'udito. Nel loro contesto originale, le ultime citazioni fatte qui sopra figurano in un paragrafo intitolato «Psychophysiology», dove si fa riferimento – fra le altre cose – al ben noto grafico di Fletcher e Munson (che mette in correlazione intensità e frequenza per illustrare i limiti di estensione dell'udito umano) oltre che a due importanti lavori di psicoacoustica<sup>22</sup>.
5. La formulazione granulare di Xenakis contiene una chiara correlazione di udito e vista. Riprendiamo la parte conclusiva della sua «basic hypothesis»:

Hecatombs of pure sounds are necessary for the creation of a complex sound. This one should be imagined as to be a display of fireworks sparkling in all colours, each luminous point of which appearing and disappearing instantaneously on the background of the black sky. But there would be so many luminous spots in this firework and they would be organized in such a way that their rapid and swarming succession creates forms, slowly uncoiling volutes or in the contrary short explosions inflaming the whole sky. A sufficient number of instantaneously appearing and disappearing spots would form a luminous line.

L'articolo «Elements of Stochastic Music» presenta un'introduzione, poi fatta cadere nella stesura di *Musiques formelles*. Xenakis vi manifesta l'intenzione di scrivere un libro<sup>23</sup> destinato in particolare ad una nuova categoria di ricercatori, cioè di "compositori sperimentali" che potrebbero essere musicisti, certo, ma anche

... film-producers, painters, sculptors etc. provided that they dispose of a certain mathematical luggage and particularly of a spirit of adventure. For I am convinced that the mathematics, physics and psychology, a sort of new Trivium, must be taught by the conservatoires within one generation if not the Scientific Research will run the risk of founding a new department: the "Audio-visual Research" which would absorb the studies about composition brought about by the actual conservatoires (Xenakis 1960a, pp. 85-86).

Xenakis non ha mai dato spiegazione del titolo *Analogique*. Si può ipotizzare che si riferisse all'analogia di udito e vista. Nouritza Matossian ha notato che l'ipotesi corpuscolare di Gabor

esercita un'influenza profonda, come immagine, sulla concezione che Xenakis ha dell'opera, perché le sue annotazioni personali presentano delle fotografie di cellule o

<sup>22</sup>Si tratta di Stevens e Davis (1938) e Winckel (1960). Si veda (Xenakis, 1960, p.90), (Xenakis 1963, p. 65) e Xenakis (1992, p. 47). Come è stato rilevato, nell'edizione del 1992 di *Formalized Music* manca la bibliografia finale (pur essendovi, nel corpo principale del testo, i rimandi bibliografici)! Negli appunti presi dal compositore tra il 1957 e il 1959 (Archives Xenakis, Carnet 19) troviamo dei riferimenti in particolare al primo dei due libri qui menzionati. Va notato che nel secondo, quello di Winckel, il compositore poté trovare un'elaborazione piuttosto avanzata del concetto di "timbro" ed altre indicazioni preziose, tra cui: «l'insieme delle constatazioni fatte nelle pagine precedenti ci mostra che, nella sua architettura complessiva, la musica è movimento; anzi, che il materiale stesso, il suono, è soltanto un elemento in movimento» (Winckel 1960, p. 119). Inoltre: «si può dunque affermare con certezza che i suoni stazionari, come quelli che sono di solito rappresentati nella notazione musicale, in effetti non hanno alcuna realtà sonora: la musica non vive che grazie a regimi transitori nel corso dei quali i suoni possiedono un timbro continuamente variabile e differente dal timbro neutro del regime stazionario» (ibid., p. 129).

<sup>23</sup>È probabile che l'idea di realizzare un libro fosse animata dalla benevolenza di Hermann Scherchen. Secondo Kanach (2001, p. 203), si tratterebbe di una versione ancora in nuce di *Musiques formelles* dal titolo *Mécanisme d'une musique*.

grani ripartiti attorno ad un nocciolo dalla densità più concentrata, raggruppati insieme in modo particolare [...] Come può un’immagine trasformarsi in un procedimento di composizione? (Matossian 1981, p. 158)<sup>24</sup>.

6. Nel 1991, preparando la Prefazione della seconda edizione di *Formalized Music*, Xenakis scrive:

An important task of the research program at CEMAMu is to develop synthesis through quantified sounds but with up-to-date tools capable of involving autosimilitudes, symmetries or deterministic chaos, or stochastics within a dynamic evolution of amplitude frequency frames where each pixel corresponds to a sound quantum of ‘phonon’, as already imagined by Einstein in the 1910s. This research, which I started in 1958 and wrongly attributed to Gabor, can now be pursued with much more powerful and modern means. Some surprises can be expected! (Xenakis 1992, p. xiii).

Qui Xenakis riporta la sua teoria granulare non al lavoro di Dennis Gabor, ma a quello di Albert Einstein. In una nota di «Elements of Stochastic Music» (l’abbiamo già citata sopra, ma incompleta), leggiamo:

... it can be considered as to be a first approach towards the ideas introduced into the theory of information by Gabor (see Meyer-Eppler, p. 21).

Dunque, al tempo, egli forse non ha ancora letto gli articoli di Gabor risalenti a qualche anno prima – cioè Gabor (1946, 1947) – e si limita a citarli attraverso il libro di Werner Meyer-Eppler del 1959, *Grundlagen und Anwendungen der Informationstheorie*. Tuttavia Xenakis inizia gli esperimenti relativi ad *Analogique A et B* almeno nel 1958. (Certo, egli avrebbe potuto seguire qualche conferenza di Meyer-Eppler su questi temi<sup>25</sup>, o avrebbe potuto sentire di Gabor da parte di Abraham Moles<sup>26</sup> o da altri). L’equivoco – se si può parlare di equivoco – è dovuto forse al fatto che il riferimento a Meyer-Eppler non compare in *Formalized Music* (la versione del 1992, la più conosciuta da compositori e ricercatori) a causa di un errore compiuto

<sup>24</sup>Negli Archives Xenakis (Dossiers œuvres 5/10) sono effettivamente conservate, tra gli appunti del compositore, due fotografie del genere.

<sup>25</sup>«Nella sua conferenza *Metamorphose des Klangelemente* – presentata nel 1955 in varie sedi tra cui a Gravesano, in Svizzera, presso lo studio di Hermann Scherchen – Meyer-Eppler descriveva la matrice di Gabor come una specie di partitura che si presta ad essere composta mediante una tecnica a mosaico [Mosaiktechnik]» (Roads 2001, p. 62). Xenakis potrebbe essere stato presente a tale conferenza.

<sup>26</sup>Questa è l’ipotesi avanzata in Di Scipio (1998, p. 70) [Nella revisione dell’articolo di Di Scipio presentata in questo numero di «Musica/Tecnologia» il riferimento è a p. 32]. Dopo una personale verifica, tendo a confermare tale ipotesi. In un testo apparso nel 1962, Xenakis scriveva: «[l’]ipotesi atomica è corroborata dalle teoria del segnale acustico elementare introdotta nel contesto della teoria dell’informazione da Gabor [...] Inoltre, A. Moles da anni sottolinea la natura quantistica della percezione uditiva [...]. Questi due fatti mi hanno indotto ad esprimere una teoria quantistica della composizione musicale» (Xenakis 1962, p. 182). Il libro di Moles, che Xenakis cita per esteso (*Théorie de l’information et Perception Esthétique*, Flammarion, Parigi) era apparso nel 1958. Xenakis riprese tali osservazioni più volte tra il 1961 e il 1962, aggiungendo ad esempio: «per altro, A. Moles ha messo in rilievo, *da molti anni*, la quantificazione della percezione» (Xenakis 1961c, p.313; corsivo nostro). Solo in una circostanza, curiosamente, dimentica di menzionare Moles in proposito (Xenakis 1961d, p. 139). Egli conosceva bene Abraham Moles già da qualche anno. Negli Archives Xenakis, c’è una copia posseduta da Xenakis di un articolo di Moles intitolato «Some Basic Aspects of an Information Theory of Music» (apparso in «Journal of the Audio Engineering Society», 6/3, 1958, pp.184-186) e recante una dedica proprio a Xenakis. Nel 1960, i due costituirono un gruppo di lavoro informale, indicato come MYAM, insieme ad Alain de Chambure e Michel Philippon (Moles, Yannis, Alain, Michel).

dalla casa editrice<sup>27</sup>. In occasione di varie interviste, Xenakis ha ripetuto di aver avanzato la ipotesi granulare senza conoscere il lavoro di Gabor, ma tenendo in considerazione appunto uno spunto di Einstein<sup>28</sup>. In una di queste interviste (Varga 1996, p. 197) egli afferma:

Negli anni 1950 ho proposto una teoria circa la sintesi del suono basata su quanti acustici. I quanti acustici rimontano ad una teoria di Einstein intorno al 1917, che parlava di «fononi»<sup>29</sup>. È una questione connessa con l’osservazione che la trasmissione del calore e del suono attraverso molecole di atomi avviene mediante salti di *energia* [corsivo nostro], proprio come nel caso dei fotoni. Sono controllate dalla stessa equazione, l’equazione di Planck. Io ho sviluppato la mia teoria in modo puramente intuitivo, ma ho capito più tardi che era stata proposta in fisica.

*Energia*: la parola appare anche nel quaderno di annotazioni personali di Xenakis datato novembre 1958, probabilmente in relazione a questioni “granulari” appunto [riportiamo dal carnet di Xenakis, lasciando il testo in francese]:

Une musique est un ensemble de transformations énergétiques.

I. Transformations ordonnées ou simplement dénombrables ou quelconques atomiques (sons pures, sinusoïdaux)

de fréquence continues (vitesse), discrètes (log).

de dynamique.

Cest-à-dire deux paramètres variables indépendantes fréquence cycle/sec ou rapport de fréquences = intervalles et niveau sonores dB attribué à chaque valeur de a). Les transformations peuvent être indépendantes ou liées entre ces deux variables et contrôlées par des représentations canoniques ou par des protocoles (matrices). Déterminées ou pas. Ces transformations me définissent la collection et sa structure interne.

II. Sur un plan plus général.

Transformations entre groupes de transformations type TI.

Sur ce plan on rencontre les timbres car on introduit la simultanéité.

Come sappiamo, il concetto di “energia” è centrale nell’estetica di Xenakis – uno dei suoi ultimi articoli, intitolato «Sur le temps» (Xenakis 1988)<sup>30</sup>, è molto chiaro in proposito. Il suo interesse per il paradigma granulare andrebbe messo in relazione anche con questo aspetto. In un certo senso, il paradigma granulare coincide del tutto con la sua teoria (o visione) circa la musica come fenomeno energetico<sup>31</sup>.

<sup>27</sup>Come detto sopra, in quella edizione di *Formalized Music* è stata erroneamente omessa la bibliografia finale. Occorre quindi riprendere la bibliografia inclusa ovviamente in *Musiques formelles* (1963).

<sup>28</sup>Si veda Restagno (1988, p. 30), Albèra (1989, p. 80), Varga (1996, p. 197). Brigitte Robindoré ha scritto: «Alla fine degli anni quaranta, Xenakis lesse un articolo [...] nella rivista inglese «Nature». L’articolo, firmato dal fisico Dennis Gabor, premio Nobel, presentava una teoria sulla natura particolare dei fenomeni acustici. Non contento dei contenuti matematici della teoria, Gabor riuscì anche a costruire un proprio “granulatore” capace di comprimere o espandere la scala temporale di segnali sonori registrati [...]. Xenakis considerò l’applicazione compositiva di tale scoperta» (Robindoré 1996, p. 11). Tuttavia l’autrice non fornisce alcuna evidenza per confermare simili indicazioni, né offre a Xenakis, nel corso della sua intervista col compositore, la possibilità di contraddirle o confermarle!

<sup>29</sup>Per il momento negli Archives Xenakis non è stata rinvenuta alcuna possibile fonte in riferimento ad Einstein.

<sup>30</sup>Trad. it. «Sul tempo», in Restagno (1988, pp. 271-281).

<sup>31</sup>Pensare la musica in termini di “energia”, di “movimento”, di “moto”, ha naturalmente una lunga tradizione. La troviamo già in Platone, nel *Timeo* – non a caso citato da Xenakis in *Musiques formelles*: «i movimenti dei suoni, che causano corrispondenti movimenti dentro di noi, procurano un piacere molto comune in coloro che non sanno

## 2 Qualche elemento per una storia dell'approccio granulare

### 2.1 Storicità del paradigma granulare

Ci fermiamo qui per quanto riguarda Xenakis<sup>32</sup>. Vorrei ora inserire l'approccio granulare in un possibile contesto storico-musicale “prima” e “dopo” Xenakis, anche per ribadire come ad esso si possa riconoscere il ruolo di un’ampia visione teorica, piuttosto che quello di una forma di sperimentazione o di un particolare contesto tecnico. Si noterà che il contesto storico di cui parliamo è difficilmente riconducibile ad una linea di continuità storica, e che non a caso uso l'espressione “circolo granulare” [nel testo inglese: *granular connection*] nel titolo di questo articolo.

Se la storia degli aspetti scientifici del paradigma granulare è certamente molto lunga – Roads (2001, p. 52) la fa risalire a Isaac Beekman (1588-1637) – quella dei suoi sviluppi musicali appare ben più corta. Difficile trovare sonorità granulose prima di Xenakis. Potremmo ricordare la “liquefazione” delle figure musicali in Debussy (Fig. 5) – e, in un certo senso, seguendo quest’idea di “scioglimento” del materiale musicale, potremmo dire che, se il modello dell’acqua ha in Debussy un ruolo assai importante, Xenakis si spinge oltre fino a lasciare che il liquido evapori diventando gas! Ma ovviamente l’aggancio con Debussy appare problematico. Il puntillismo weberniano potrebbe essere una via più interessante, ma anch’essa difficile da seguire. E ciò vale anche per certe soluzioni texturali ottenute da Bartók con suoni pizzicati. In Varèse, certe sonorità hanno certamente una dimensione granulare, per esempio in *Ionisation*<sup>33</sup> (Fig. 6).

Per quanto riguarda i contemporanei di Xenakis, il serialismo puntillista degli anni 1950, laddove sviluppato a partire già dalla microstruttura musicale, effettivamente raggiunge talvolta dimensioni granulari, specialmente nella musica di Stockhausen. Un buon esempio è *Gruppen* (per quattro gruppi orchestrali, 1955-57). In Fig. 7 si vede una battuta in cui, grazie alla struttura per «formanti ritmici» (Stockhausen 1957), la densità di suoni è abbastanza elevata da

---

*ragionare; coloro che lo sanno fare, ne ottengono una gioia della ragione per via dell’imitazione, pur in movimenti destinati a finire, dell’armonia divina* – le parole in corsivo sono riferite a Platone (Xenakis 1963, p. 212), (Xenakis 1992, p. 179). A tal proposito, pensiamo all’ammirazione di Xenakis per Brahms, amico di quel Eduard Hanslick che aveva scritto il famoso *Del bello musicale* (1854) proprio sulla base di una visione del genere. Hanslick fu detto “formalista” perché, come Brahms, si oppose al contenutismo della musica a programma (Liszt); con occhi xenakisiiani, potremmo leggerlo oggi come propugnatore di un’estetica energetica, antitetica ad una estetica espressiva o linguistica. In effetti, una concezione della musica come energia può forse spiegare perché Xenakis fosse contrario ad assimilare la musica ad un linguaggio. Si veda Solomos (2004).

<sup>32</sup> Per approfondimenti sulla dimensione granulare in Xenakis, e in particolare su *Concret PH* e *Analogique A et B*, rinviamo ad alcuni scritti di Agostino Di Scipio (si veda bibliografia) oltre che a Bridoux-Michel (2005), Harley (2002) e Meric (2005). Per specifici aspetti tecnici rinviamo a Roads (2001, pp. 64-68 e passim). Per approfondimenti su *Analogique* rinviamo anche a Orcalli (1993, Cap. 4) e Bokesoy (2004).

<sup>33</sup> In Solomos (1995) ho svolto un’analisi di *Ionisation* in termini di tre “sonorità” principali, che vengono prima esposte, e poi servono da materiale per lo sviluppo dell’intero brano. A proposito di Varèse, si può segnalare quanto suggerisce Curtis Roads (2001, p. 54) subito dopo aver discusso dei «fononi» (o quanti acustici) di Einstein: «A suo modo, Edgar Varèse nella sua capacità visionaria riconobbe [in un testo del 1940] il significato della scoperta [dei fononi]: ‘Ogni suono [tone] è un’entità complessa fatta di elementi ordinati in vari modi [...] In altri termini, ogni suono [tone] è una molecola di musica, e allora può essere dissezionata in atomi sonori componenti [...] i quali possono essere pensati come nient’altro che onde di energia sonora che si irradia nell’universo, come i raggi cosmici scoperti di recente, che Milliken chiama, in modo interessante, vagiti degli elementi semplici: elio, ossigeno, silicio, e acciaio’». Mettendo in evidenza la parola “energia”, troviamo qui uno dei molti punti di passaggio da Varèse a Xenakis. Si veda Solomos (2006).



**Figura 5:** Claude Debussy, *Jeux* (1913). L'inizio del numero 31 presenta una superficie liscia costruita con cinque figure distribuite tra varie parti dell'orchestra (© Éditions Durand).

potersi tradurre, all'ascolto, in un fenomeno di sintesi: per ogni mezza battuta, potenzialmente vi sono 28 note tra loro non sovrapposte (Tab. 1). Negli anni 1950 e 1960, Xenakis usava sovrapposizioni ritmiche del genere, ma limitandosi di solito a suddivisioni più semplici (per 2, 3 e 5), proprio come all'inizio di *Analogique A* (Fig. 1), ottenendo pertanto un troppo esiguo numero di eventi nell'unità temporale. Naturalmente, in Xenakis c'è un'intenzione esplicita di ottenere una dimensione granulare, mentre in Stockhausen si tratta piuttosto di una possibilità derivata. La parte elettronica di *Kontakte* (1959-60) è forse più interessante in tal senso, dato che quasi tutti i suoni elettronici furono prodotti da Stockhausen mediante generatore di impulsi (la velocità di impulso poteva essere variata con continuità tra 16 al secondo e 0.125 al secondo, mentre la durata di ciascun impulso poteva variare tra 0.0001" e 0.9") (Stockhausen 1996, p. 174). Ma anche qui, l'idea di granularità non è elaborata come tale. Alcune delle soluzioni "metronomiche" che Gyorgy Ligeti escogitò negli anni sessanta

**Figura 6:** Edgar Varèse, battute 18-20 di *Ionisation* (1929-31), riduzione dell'orchestra di percussione.

4

p.      p.      p.

a. fl.      *mf*

tpt.      *p*

wood drums      *p*

drums      *mf*

mar.      *weich*      *mf*      *fp*

glock./cel.      *mf*      *pp*      *mf*

harp      *mf*

I + II      *pizz.*      *p*      *mf*

vns.      (arco)      *s*

III + IV      *mf*

va.      *p*      *mf*

vc.      *pizz.*      *p*      *mf*

**Figura 7:** Stockhausen, *Gruppen*: «formante ritmico» costituito da una «fondamentale» (minima alle viole) e da nove «armonici» (seminimina ai violoncelli, terzina all’arpa, ecc.) (© Universal Edition).

**Tabella 1:** Sequenza temporale ottenuta dalla divisione armonica di una unità ritmica per 1, 2, 3...9.

---

|   |
|---|
| 0; 0.111; 0.125; 0.143; 0.167; 0.2; 0.222; 0.25; 0.286; 0.333; 0.375; 0.4; 0.429; 0.444; 0.5; 0.555; 0.571; 0.6; 0.625; 0.667; 0.714; 0.75; 0.778; 0.8; 0.833; 0.857; 0.875; 0.889. |
|---|

---

raggiungono una maggiore densità.

Dopo Xenakis, inizialmente vi sono stati tentativi isolati, come quelli di Horacio Vaggione all'inizio degli anni 1970, per esempio in *Modelos de Universo* (1971) e in *Movimiento continuo* (1972), dove usò un programma per la sintesi digitale del suono (chiamato Papova, eseguito su un computer mainframe IBM 7090) che generava fino a 20 suoni al secondo in ciascuna delle quattro voci della polifonia di cui era capace (Vaggione 1983). Un altro esempio può essere dato da certi brani di Bernard Parmegiani, come *Matières induites* (un movimento del ciclo *De natura sonorum*, 1975). Il paradigma granulare diventa poi vera e propria “sintesi granulare” alla fine degli anni 1970, nel lavoro di Curtis Roads (1978) e poi di Barry Truax (1988). È appunto a partire dagli anni 1980 che si forma una sorta di “circolo granulare”, in cui rientrano sia Roads, sia Truax, sia Vaggione. Preferisco parlare di un “circolo” [*connection*], piuttosto che di una “linea” di sviluppo (che partirebbe da Xenakis) perché questi compositori manifestano in effetti estetiche tra loro del tutto differenti: Vaggione procede secondo un approccio attento ai differenti livelli temporali della struttura sonora (lo vedremo più avanti); Roads sembra più interessato ai microsuoni stessi; Truax, muovendo da un retroterra di composizione con suoni del paesaggio [*soundscape composition*], sfrutta la granularità come lente di ingrandimento per “scendere nel suono” con effetti di riverberazione non solo acustici ma anche “mnemonici”<sup>34</sup>.

Qui vorrei soffermarmi sull'approccio di due compositori di generazione diversa, Horacio Vaggione (nato nel 1943) e Agostino Di Scipio (nato nel 1962), sia per illustrare la varietà delle posizioni estetiche nel contesto del paradigma granulare, sia perché, benché in modo diverso da Xenakis, anche la loro capacità di teoria – la loro “visione” – si traduce in musica di alta qualità.

## 2.2 Horacio Vaggione e l'approccio “multi-scala” alla struttura temporale

Come abbiamo detto, Vaggione inizia ad elaborare un approccio granulare già all'inizio degli anni 1970. Lo si potrebbe collegare al puntillismo post-weberniano caratteristico della sua primissima produzione, negli anni sessanta (Laliberté 2007, p. 352). Ma è a partire dagli anni 1980 che Vaggione viene considerato un maestro del paradigma granulare, in considerazione di lavori come *Thema* (1985, per sax basso e suoni generati da computer), *Tar* (1987, per clarinetto basso e suoni generati da computer), *Ash* (1989-90, musica elettroacustica),

<sup>34</sup> «La tecnica che trovo più sorprendente, per il modo in cui permette di muoversi dentro il suono, è la granulazione in tempo reale di suoni campionati [...] Essa permette trasformazioni di grande efficacia, grazie al processo di stretching temporale [...] usato non solo per creare sonorità ferme e prolungate [drones], ma anche per far emergere all'ascolto il dettaglio interno della struttura timbrica [*innertimbral character of sound*], come se si lavorasse al microscopio [...] Applicata alla composizione con suoni del paesaggio», questa dilatazione delle durate [...] permette al suono di riverberare nella memoria dell'ascoltatore, lasciando che possano affiorare in superficie ricordi di lungo periodo e associazioni di idee» (Truax 1988, p. 96). Si veda anche in questo numero l'articolo di Truax, *passim*.

*Kitab* (per clarinetto basso, pianoforte, contrabasso e computer), *Schall* (1994, musica elettroacustica), *Myr-S* (1996, per violoncello e dispositivi elettroacustici), *Préludes Suspendus II* (2000, musica elettroacustica), *Atem* (2002, per corno, clarinetto basso, pianoforte, contrabasso e dispositivi elettroacustici) e *Taléas* (2002-2004, per flauto dolce Paetzold e dispositivi elettroacustici)<sup>35</sup>. La questione importante, nel suo approccio compositivo, è quella della strutturazione del tempo.

La musica di Vaggione in realtà presenta una molteplicità di temi relativi al tempo. Iniziamo prima di tutto col notare che egli integra nel suo lavoro una concezione “moderna” del tempo, considerato nella sua irreversibilità – seguendo in ciò Xenakis e procedendo in parallelo coi compositori della scuola spettralista francese. In questo contesto, il suono non è più concepito in termini di periodicità e ripetizione, come nel modello “classico” dell’acustica di Helmholtz, ma come fenomeno dinamico ed energetico. Uno dei punti di riferimento più importanti è il lavoro del fisico Ilya Prigogine, il cui nome è legato alla «teoria delle strutture dissipative». Vaggione vi fa riferimento parlando di «strutture dissipative di energia sonora» (Vaggione 2003, p. 102).

In secondo luogo, Vaggione è interessato a lavorare nella dimensione infinitesimale del suono, cioè all’interno di ciò che è stata definita spesso la «vita interna» del suono – questione cruciale per una lunga genia di compositori che parte certamente da Varèse e continua con Stockhausen fino a certa musica odierna di forte impatto “immersivo”. Tuttavia, Vaggione ha una sua collocazione assai specifica e non assimilabile ad altri esempi. Se non altro, egli non sembra uniformarsi agli aspetti che, in quel tipo di sensibilità musicale, vengono più spesso associati allo spazio che al tempo. Piuttosto che di «vita interna» del suono, in effetti egli preferisce parlare di una “discesa” nel micro-tempo. A parte una certa fascinazione per fenomeni sonori che accadono a scala micro-temporale, la questione cruciale per lui è quella della loro articolazione, cioè di come comporre tali fenomeni. Il titolo di un suo scritto mette insieme proprio questi due aspetti: «Articulating Micro-Time» (1996). Inoltre la problematica della «vita interna» del suono rinvia spesso ad un’ipotetica matrice unitaria originale – dando vita a derive misticheggianti, per esempio in compositori come Scelsi, Stockhausen, Harvey, e anche Grisey – mentre Vaggione invece vede in questa discesa nel micro-tempo un’opportunità di pluralismo:

per un musicista, scendere nel micro-tempo è un mezzo per scoprire fenomeni di cui si è inconsapevoli quando ci si limita ad ascoltare il movimento della superficie sonora senza considerare il suo sostrato [...] Come diceva Bachelard, la nostra intuizione del tempo è ancora molto misera, limitata all’intuizione di un inizio assoluto e di una durata continua. Per Bachelard, dobbiamo invece trovare il pluralismo delle parti nascoste dietro l’identità del tutto, rompere l’identità ben oltre l’esperienza immediata che troppo spesso tendiamo a condensare in apparenza unitaria esteriore (Vaggione 1998b).

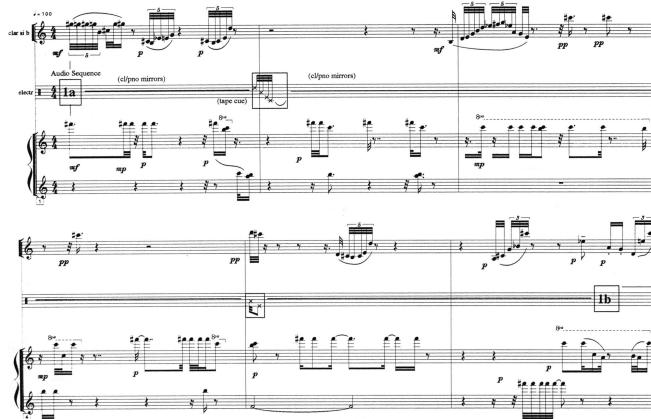
L’interesse di Vaggione per la sintesi granulare nasce allora dal tentativo di trovare una “pluralità” (grani) dietro quella che appare come “identità” di superficie (suono). Inoltre, Vaggione sottolinea che la descrizione corpuscolare, diversamente da quella ondulatoria, rinvia all’irreversibilità del tempo. L’approccio granulare «ci permette di lavorare con morfologie sonore complesse in uno spazio-tempo in cui regna l’irreversibilità: con strutture ‘dissipative’ che emergono dentro uno spazio-tempo direzionale, invece che con continuità simmetriche e lisce» (Vaggione 2005, p. 341). In tal senso, nella sua musica l’approccio granulare è allora ben

<sup>35</sup>Uso qui la stessa terminologia usata da Vaggione per indicare le parti elettroniche di questi lavori. Si veda il volume collettivo curato da Solomos (2005c, pp. 413-418).

più di una tecnica di sintesi. Si tratta non solo di articolare il suono a livello micro-temporale (sintesi), ma di pensare in modo granulare anche il macro-tempo (il che avviene nelle parti strumentali dei suoi lavori). Peraltra Vaggione parla di “granulazione” in modo piuttosto generale, e con piacere cita Lavoisier nel trattare di questi temi (Vaggione 1996, p. 34).

In terzo luogo, alcuni dei testi di Vaggione (dalla seconda metà degli anni 1980 fino alla fine degli anni 1990) portano avanti l’idea di un approccio multi-scala. Di per sé, l’idea non è del tutto nuova – si pensi a quanto detto sopra circa Xenakis, e si pensi alle “tre temporalità” di Grisey (Baillet 2001, p. 25). Tuttavia nuovi sono gli specifici sviluppi che Vaggione riesce a perseguire. Egli inizia dalla constatazione pragmatica che, tanto nella tradizione musicale quanto nella percezione umana, esistono soglie a partire dalle quali tendiamo a delimitare due ordini di scala temporale, genericamente individuate come “micro-tempo” e “macro-tempo”. Nella tradizione musicale, è la “nota” che viene considerata come unità elementare: parleremo allora di macro-tempo per indicare «tutte le scale temporali possibili» al di sopra di quella unità elementare, e di micro-tempo per indicare quelle al di sotto (Vaggione 1998a, p. 172). Le ripercussioni di una formulazione apparentemente innocua come questa sono enormi: esse permettono di riformulare la separazione tra musica strumentale ed elettroacustica, e allo stesso tempo di mitigarne il contrasto. In effetti, da questa angolazione, il gap tra le due non sta in una differenza di “natura” (cioè di materiale) ma in un cambio di scala (temporale) di osservazione. Questo modo di pensare è praticabile grazie all’elettronica digitale, avendo essa reso possibile comporre al livello micro-temporale. Dobbiamo allora considerare entrambi i campi (il micro- e il macro-tempo) all’insegna del “componibile”, ovvero possibili di articolazione – senza peraltro abolire la soglia che porta a distinguere il cambio di scala. In termini di percezione umana, la soglia notoriamente va situata tra i 50 e i 100 millisecondi. Nel contesto del modello granulare, ciò significa che, con meno di dieci o venti suoni al secondo, l’orecchio coglie i singoli grani sonori come entità separate, e che quando ve ne sono di più, ciò che l’orecchio coglie è invece un suono unitario, una texture globale. Applicando questo modo di pensare tanto a livello di sintesi del suono quanto al livello di scrittura strumentale, si possono tenere unite le due dimensioni senza abolirne le differenze. Nel caso della sintesi, percepiamo la granularità del suono, cioè una pluralità interna che non impedisce tuttavia di parlare di “un” suono. Invece a livello strumentale, anche giungendo molto in prossimità della soglia del micro-tempo – come accade spesso nelle partiture di Vaggione, che arrivano a richiedere quintine di biscrome sulla seminimima a 100 MM (Fig. 8) – rimaniamo comunque nel contesto di un flusso di elementi separati. In effetti, è ovvio, Vaggione non si aspetta certo una sorta di «sintesi granulare strumentale» (Vaggione 2005, pp. 341 e sgg.).

Insisto sul fatto che, laddove egli pone la distinzione tra dimensioni micro- e macroscopiche del suono come questione di diverse scale temporali, Vaggione non mira a mettere da parte la soglia che divide i due livelli, piuttosto ne fa elemento di raccordo che li tiene insieme pur preservandone le differenze. In altri termini, non possiamo spostarci da un livello all’altro per semplice trasposizione. O, come dice il compositore, non possiamo trascurare la nonlinearità che esiste tra livelli temporali diversi, la loro mutua irriducibilità. Questo è la vera originalità del suo contributo teorico-musicale, dato che molti compositori interessati alla questione, prima di Vaggione o negli stessi suoi anni, sono apparsi inclini a seguire invece un principio di trasposizione. È proprio il caso di Xenakis, per esempio nei brani realizzati col programma informatico GENDYN [all’inizio degli anni 1990], dove tutto è dedotto automaticamente dal livello della forma d’onda. Ed è anche il caso di Grisey, laddove questi proietta la stessa forma d’onda a scale temporali diverse, per esempio in *Vortex temporum* [per pianoforte e cinque



**Figura 8:** Horacio Vaggione, *Phases* (2001), pagina 1.

strumenti, 1996]. Potremmo citare anche certi usi della geometria frattale in musica<sup>36</sup>. Per Vaggione, queste nonlinearità tra scale di tempo diverse non solo esistono concretamente, ma sono musicalmente produttive, come ha segnalato in un'intervista del 2000:

riconoscere la realtà delle mancate coincidenze [tra scale di tempo diverse] non ci porta alla paralisi, al contrario, ci offre la possibilità di esplorare il passaggio tra differenti dimensioni, di articolare queste dimensioni in una rete di legami sintattici che copre l'intero spettro delle relazioni componibili (Budò 2000, p. 15).

### 2.3 Agostino Di Scipio e la questione dell'emergenza

In un articolo di qualche anno fa, Vaggione scriveva:

[...] si può pensare che il vero motivo di una riuscita integrazione dell'approccio granulare nella composizione musicale debba essere visto nel transfert categoriale dal livello del segnale verso il livello simbolico. I grani manipolati musicalmente presi singolarmente non “spiegano” nulla a proposito della natura elementare del suono: nel nostro caso, essi sono – già loro – delle morfologie musicali, certo situate a livello del micro-tempo, ma tuttavia frutto di composizione. L'interesse compositivo dell'approccio granulare consiste allora nel trattamento simbolico di elementi presenti a scala micro-temporiale. In questo senso, i grani costituiscono elementi duttili coi quali si possono elaborare entità morfologiche, agglutinandoli e proiettandoli in tutto lo spazio componibile, a ogni

<sup>36</sup>Non c'è dubbio che, in una fase iniziale del proprio percorso, anche Vaggione si sia interessato a questioni di autosimilitudine (Vaggione 1989). Egli segnala però che lavori più recenti, quali *Tahil* (1992), *Kitab* (1993), *Schall* (1994) e *Rechant* (1995), esplorano invece «condizioni dinamiche di interazione, di convoluzione, di aliasing, e anche il movimento tra stati laminari e turbolenti della materia sonora, da cui nascono figure che non operano allo stesso modo a differenti scale di tempo». Dopo aver letto alcune mie osservazioni in proposito (Solomos 2005a), Vaggione mi ha detto di essersi interessato alla nonlinearità già a partire dalla metà degli anni 1980, in lavori quali *Thema* (1985), *Ash* (1989-90) e *Till* (1991). È probabile che sia giunto gradualmente ad una precisa consapevolezza della questione, a partire dalla constatazione degli effetti nonlineari dei tentativi di transposizione del timbro (Vaggione 2003, p. 106). Da quel momento in poi, la nonlinearità delle scale temporali è divenuta per lui una problematica centrale come generatore di singolarità.

scala possibile. Partendo da un approccio al sonoro nel segno della discontinuità e della frammentazione, stiamo già creando, per così dire direttamente, oggetti musicali, oggetti composti (Vaggione 2006).

Questo passaggio può costituire una buona introduzione (magari in forma di un confronto critico) all’approccio granulare caratteristico di Agostino Di Scipio, che a sua volta si concentra sul concetto di emergenza quale sviluppato nelle recenti scienze cognitive (Varela 1996). Iniziamo però da quel che Di Scipio ha scritto a proposito di Xenakis (si vedano i vari riferimenti già segnalati sopra). Nella sua analisi di *Analogique A et B*, commentando l’ipotesi dell’(auto-)creazione di sonorità di ordine superiore, Di Scipio opera un piccolo scivolamento concettuale: «oggi – egli osserva – scienziati cognitivisti ed epistemologi probabilmente descriverebbero l’ipotesi di sonorità di secondo ordine in termini di proprietà emergenti della struttura sonora» (Di Scipio 2001, p. 72). La questione può essere formulata così:

in questo caso [in riferimento ad *Analogique B*], è difficile fare una netta distinzione tra articolazione musicale e progettazione del suono, nella misura in cui le azioni del compositore mirano a far in modo che la struttura musicale (macro-livello) emerga dal suono stesso e dalla sua organizzazione interna (micro-livello) (Di Scipio 1997, p. 165).

Di Scipio ritiene che il “fallimento” di *Analogique* – che per lui non sta nella mancata fusione dei grani, come invece forse riteneva Xenakis, ma in un mancato fenomeno di emergenza – sia da far risalire agli strumenti matematici usati da Xenakis: «ci si può chiedere se le leggi stocastiche possano davvero fornire la possibilità che, dall’organizzazione di unità sonore minimali, emergano sonorità di ordine superiore» (Di Scipio 2001, pp. 73 e 79). Ecco perché, all’inizio del suo percorso compositivo, egli decide di adoperare modelli di sistemi aventi comportamento dinamico complesso, più adeguati a gestire la materia granulare in modo da farne emergere delle forme superiori, come si evince da un passaggio di un’intervista di qualche anno dopo:

il caos deterministico e la dinamica di sistemi complessi, almeno in quanto accessibili con semplici modelli numerici iterati, rappresentò per me un modo di comporre configurazioni di piccole unità sonore in modo tale che una sonorità di livello superiore si manifestasse nello svolgersi stesso del processo (Anderson 2005).

Tali modelli lo mettevano in grado di «esplorare [...] un vasto] insieme di configurazioni granulari, da textures più randomiche e disordinate, a textures di struttura più regolare, attraverso vari stati intermedi» (ibid.). In questo modo, e diversamente da Xenakis e da Vaggione, Di Scipio tende a disfarsi di tutto ciò che sia precomposto o composto direttamente in quanto forma o in quanto progetto macrostrutturale. Non a caso, in molta sua musica, non ci sono gesti intenzionalmente drammatici. In uno dei suoi primi scritti, Di Scipio (1994) formula una «teoria dell’emergenza sonologica» in cui la forma musicale (la macrostruttura) è concepita come «processo di formazione timbrica» (Di Scipio 1994, p. 205).

Nella posizione teorica di Di Scipio, l’emergenza è possibile grazie al fatto che il compositore progetta sistemi (nel senso cibernetico del termine), in particolare sistemi simili ai sistemi viventi, ovvero caratterizzati da capacità di auto-organizzazione:

il passaggio di un sistema o di un processo da una data configurazione strutturale ad un nuovo stato di ordine comprensibile come funzione delle proprietà qualitative del precedente, è ciò che in questo contesto chiamo un fenomeno di emergenza [...] Fenomeni del genere

possono seguire criteri di morfostasi (conservazione di coerenza, identità) o di morfogenesi (comportamento dinamico, cambiamento) i quali, presi insieme, portano alla caratteristica peculiare che accomuna sistemi viventi e sistemi sociali: l'auto-organizzazione (ibid., p. 206).

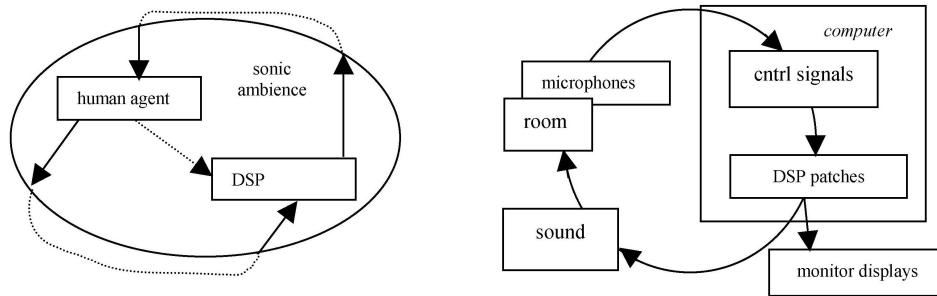
Ad assicurare questa capacità di auto-organizzazione provvede un criterio di «causalità circolare» (Anderson 2005), inteso come estensione dell'idea di feedback. Un esempio lo si trova in un lavoro (che Di Scipio ha dedicato proprio a Vaggione) intitolato *Due di Uno* (per violino, flauto dolce e processi adattivi di elaborazione del suono, 2002), dove il suono strumentale viene elaborato secondo tecniche granulari ma allo stesso tempo costituisce anche la sola sorgente dei controlli esercitati su quei processi di elaborazione (Di Scipio 2005a).

In base a questo criterio di causalità circolare, Di Scipio ripensa anche la nozione, tipica della musica live electronics, di “interazione” (Di Scipio 2003), allargando così il contesto sul quale si ripercuote il suo iniziale approccio granulare. Di solito si concepisce l’interazione come un’operazione svolta su un flusso di informazione: per esempio, un flusso sonoro avente un suo decorso, viene attivamente modificato da un operatore. Ma ciò in effetti non costituisce un’effettiva condizione interattiva, manca una vera reciprocità di azione. Ragionando in questo modo, per Di Scipio la composizione diventa in fin dei conti l’azione di comporre interazioni:

si tratta innanzitutto di creare un sistema dinamico che esibisca un comportamento adattivo alle condizioni esterne circostanti, e capace di interferire con esse [...] ottenendo così la base per un più ampio processo di auto-organizzazione [...] Qui, ‘interazione’ vale come elemento strutturale affinché possa darsi qualcosa che sia effettivamente un ‘sistema’ [...] Le interazioni sistemiche sono dunque solo indirettamente implementate, diventano cioè esiti collaterali di interdipendenze sistemiche aperte ma attentamente progettate [...] Il che costituisce un passo deciso dalla «composizione di musica interattiva» alla «composizione di interazioni musicali», e più precisamente dalla produzione di suoni desiderati, con mezzi interattivi alla produzione di interazioni desiderabili, che poi lasciano le proprie tracce udibili (ibid., p. 271).

Potremmo dire che per Di Scipio il processo di queste interazioni diventa più importante del risultato sonoro – e perfino dell’origine del suono. Diventa importante dunque una certa concezione «ecosistemica»: le interazioni di cui parliamo accadono nella struttura granulare della materia sonora, ma anche in rapporto all’ambiente acustico circostante: nel ciclo di brani intitolati *Ecosistemico Udibile* (2002-2005) – lavori la cui esecuzione richiede quasi sempre solo risorse di live electronics – il concetto di ecosistema è esplicitato come dominio di interazioni triangolari tra l'esecutore, i mezzi tecnici, e l'ambiente in cui avviene l'esecuzione (Fig. 9) (ibid., pp. 272-275). Peraltra, in tale prospettiva assume un ruolo decisivo il rumore: semplificando, si può dire che in questo contesto “rumore” non vale in quanto “disturbo” (come sarebbe per qualsiasi musica tradizionalmente concepita), e nemmeno soltanto in quanto “materiale sonoro” da manipolare (come in molta musica moderna e contemporanea), e vale piuttosto come agente ambientale partecipe del dominio di interazioni ecosistemiche, come presenza prodotta dallo spazio in cui l’interazione ha luogo: esso, in definitiva, è una parte integrante del sistema. In *Ecosistemico Udibile*,

il suolo del “rumore” è cruciale [...] Esso in definitiva è il medium in cui è situato il dispositivo di produzione del suono, il suo stesso ambiente in senso stretto. In più, il rumore è qui energia minima e necessaria grazie alla quale il sistema può, in base alla sua auto-organizzazione, sostenere sé stesso e svilupparsi (ibid., p. 271).



**Figura 9:** Agostino Di Scipio, *Ecosistemico udibile* (2002-2005): a sinistra, schema generico della rete di interazioni; a destra, processo ricorsivo fondamentale. Riprodotto a partire da Di Scipio (2003, p. 272).

Un ultimo ma importante aspetto su cui voglio soffermarmi consiste nella prospettiva sub-simbolica che Di Scipio deriva dalle sue iniziali strategie [granulari]. L'emergenza di sonorità di livello superiore ha luogo a partire da "grani" o anche da "campioni" che in sé stessi non sono simboli, e che operano cioè ad un livello che precede ogni possibilità di simbolizzazione musicale (Di Scipio 1994, p. 207). La «composizione di interazioni» porta allora l'interazione direttamente a livello del segnale: nel lavoro di Di Scipio, ogni scambio di informazione avviene nel suono, ha natura sonora (Di Scipio 2003, p. 272). Si potrebbe fare un parallelo col modello di emergenza proprio dei recenti sviluppi delle scienze cognitive [come in Solomos (2013)]. Alla domanda cos'è la cognizione? il computazionalismo risponde: «un processo di elaborazione dati, la manipolazione di simboli sulla base di regole» (Varela 1996, p. 42). La prospettiva dell'emergenza, invece, risponde: il formarsi di «stati globali complessi in una rete di componenti semplici» (*ibid.*, p. 77). Per tornare alla musica, si tratta in pratica di ciò: se vogliamo che un livello strutturale superiore (la forma musicale, la macroforma) si manifesti come emergenza dal basso, e non come costruzione indipendentemente e astrattamente determinata, dobbiamo operare appunto a basso livello, abbandonando ogni livello temporale intermedio e i relativi processi di carattere simbolico.

Processi, «interazioni composte», «ecosistemi», strategia «sub-simbolica»: tutti questi elementi convergono. *Cos'è la musica?* chiede Di Scipio. Il risultato sonoro di determinate azioni compositive? No, perché quel che egli compone è un processo, non l'esito risultante. Un gesto deliberato di comunicazione tra esseri umani (che riguarda cioè il compositore, gli esecutori, gli ascoltatori)? Non esclusivamente, perché l'ambiente ne è parte integrante. Un linguaggio (dove la mediazione di simboli crea la dicotomia di materia e significato)? No... Per Di Scipio la musica

...non è qualcosa che possa esistere al di fuori del momento in cui si agisce per farla, ma qualcosa la cui esistenza resta sempre da realizzare, da rinnovare ogni volta; non è mai qualcosa di dato, di delimitato come forma ideale o virtuale, come qualche cosa che si presti ad essere rappresentata [ri-presentata], re-incarnata. In breve, personalmente non compongo la musica in quanto tale, ma le condizioni favorevoli che potranno eventualmente far nascere della musica (la mia musica). La responsabilità delle azioni da compiere (nel comporre, nel suonare, nell'ascoltare) ha altrettanta importanza degli oggetti da produrre (da comporre, da suonare, da ascoltare) (Di Scipio 2007, p. 296).

### 3 Conclusione

In questo articolo ho sviluppato due linee di ricerca. La prima riguarda Xenakis, che ha introdotto un modo di pensare la musica che abbiamo indicato come “paradigma granulare”. Si tratta di una ricerca filologica, perché prova ad individuare le fonti di questo pensiero e l’evoluzione nel modo stesso in cui esso è stato presentato; ma ha anche un carattere teorico ed estetico, nel senso che illustra i vari aspetti di quel modo di pensare la musica. La seconda linea di ricerca mostra che c’è una sorta di filiazione all’interno del paradigma granulare, e che potremmo scrivere una storia della “composizione granulare”. Qui mi sono concentrato su due nomi importanti, quelli di Horacio Vaggione e di Agostino Di Scipio, ma naturalmente c’è ancora molto lavoro da fare. In effetti, a partire dagli anni novanta, l’approccio granulare diventa sempre più diffuso, e viene sviluppato per esempio da compositori come Ludger Brümmer, Manuel Rocha (1999), Ramon Gonzalez-Arroyo (2005), Eduardo Miranda (2002), Damian Keller, Mara Helmuth (Teresa Davis 2004), e da molti altri, tra cui anche i musicisti della scena *electronica*.

### Bibliografia

- Albèra, Philippe (1989). «Entretien avec Iannis Xenakis». In: *Musique en création. Le festival d’automne à Paris*. Paris: Contrechamps, pp. 79–86.
- Anderson, Christine (2005). «Dynamical Networks of Sonic Interactions. An Interview with Agostino Di Scipio». In: *Computer Music Journal* 29.3, pp. 11–28. URL: <http://www.mitpressjournals.org/doi/pdf/10.1162/0148926054798142> (visitato il 30/12/2015).
- Baillet, Jérôme (2001). *Gérard Grisey. Fondements d’une écriture*. Paris: L’Harmattan.
- Bokesoy, Sinan (2004). «Macro Son: situation et évaluation. Essai de définition du macro-son et création dune méthodologie le synthétisant pour usage compositionnel». PhD. Paris: Université Paris 8.
- Bridoux-Michel, Séverine (2005). «Autour de Concret PH». In: *International Symposium Iannis Xenakis*. A cura di A. Georgaki e M. Solomos. University of Athens, pp. 109–112.
- Budòn, Osvaldo (2000). «Composing with Objets, Networks and Time Scales: An Interview with Horacio Vaggione». In: *Computer Music Journal* 24.3, pp. 9–22.
- Delalande, François (1997). «Il faut être constamment un immigré». *Entretiens avec Xenakis*. Paris: Buchet-Chastel/INA-GRM.
- Di Scipio, Agostino (1994). «Formal Processes of Timbre Composition. Challenging the Dualistic Paradigm of Computer Music». In: *Proceedings of the 1994 International Computer Music Conference*. Vol. 17, pp. 202–208.
- (1997). «The problem of 2nd-order sonorities in Xenakis’ electroacoustic music». In: *Organised Sound* 2/3, pp. 165–178.
- (1998). «Scienza e musica dei quanti acustici: l’eredità di Gabor». In: *Il Monocordo* 6, pp. 71–76.
- (2001). «Clarification on Xenakis: the Cybernetics of Stochastic Music». In: *Présences de / Presences of Iannis Xenakis*. A cura di M. Solomos. Paris: CDMC, pp. 71–84.
- (2003). «Sound is the interface: from *interactive* to *ecosystemic* signal processing». In: *Organised Sound* 8.3, pp. 269–277.
- (2005a). «Due di uno: a composition dedicated to Horacio Vaggione». In: *Contemporary Music Review* 24.4-5, pp. 383–397.

- Di Scipio, Agostino (2005b). «Formalization and intuition in *Analogique A et B* (with some remarks on the historical-mathematical sources of Xenakis)». In: *International Symposium Iannis Xenakis*. A cura di A. Georgaki e M. Solomos, pp. 95–108.
- (2007). «*Due di uno*. Une composition dédiée à Horacio Vaggione». In: *Espaces composables*. A cura di Makis Solomos. Paris: L’Harmattan, pp. 293–315.
- 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.
- (1947). «Acoustical quanta and the theory of hearing». In: *Nature* 159.4044, pp. 591–594.
- Gibson, Benoît (1994). «La théorie et l’œuvre chez Xenakis: éléments pour une réflexion». In: *Circuits* 5.2, pp. 41–54.
- Gonzalez-Arroyo, Ramon (2005). «Le concept de son en synthèse numérique». PhD. Paris, pp. 153–159.
- Harley, James (2002). «The Electroacoustic Music of Iannis Xenakis». In: *Computer Music Journal* 26.1, pp. 35–38.
- Kanach, Sharon (2001). «A propos de Musiques formelles». In: *Portrait(s) de Iannis Xenakis*. A cura di F.B. Mâche. Paris: Bibliothèque Nationale de France, pp. 201–213.
- Laliberté, Martin (2007). «An Analytic Approach to Horacio Vaggione’s *Till*». In: *Espaces composables. Essais sur la musique et la pensée d’Horacio Vaggione*. A cura di M. Solomos. Paris: L’Harmattan.
- Matossian, Nouritza (1981). *Iannis Xenakis*. Paris: Fayard.
- Meric, Renaud (2005). «Concret PH, un espace mouvant». In: *Actes des Journées d’Informatique Musicale (JIM)*. A cura di A. Sedes e H. Vaggione. Paris: Université Paris 8 CICM MSH Paris Nord Afim, pp. 147–155.
- Meyer-Eppler, Werner (1959). *Grundlagen und Anwendungen der Informationstheorie. Kommunikation und Kybernetik in Einzeldarstellungen 1*. Berlin: Springer-Verlag.
- Miranda, Eduardo Reck (2002). *Computer Sound Design: Synthesis Techniques and Programming. Pauline Oliveros and Lesbian Musicality*. Oxford: Focal Press.
- Orcallì, Angelo (1993). *Fenomenologia della musica sperimentale*. Potenza: Sonus Edizioni Musicali.
- Restagno, Enzo, cur. (1988). *Xenakis*. Torino: EDT/Musica.
- Roads, Curtis (2001). *Microsound*. Cambridge, London: The MIT Press.
- Robindoré, Brigitte (1996). «Eskhaté Ereuna: Extending the Limits of Musical Thought Comments On and By Iannis Xenakis». In: *Computer Music Journal* 20.4, pp. 11–16.
- Rocha, Manuel (1999). «Les techniques granulaires dans la synthèse sonore». PhD. Paris: Université Paris 8.
- Solomos, Makis (1993). «À propos des premières (1953–1969) de Iannis Xenakis. Pour une approche historique de l’émergence du phénomène du son». PhD. Paris: Université Paris 4.
- (1995). «Lectures d’*Ionisation*». In: *Percussions* 40, pp. 11–27.
- (1996). *Iannis Xenakis*. Revisione col titolo *Introduction à la musique de Iannis Xenakis*, 2005. Mercuès: P.O. Éditions. URL: <http://www.univ-montp3.fr/~solomos> (visitato il 09/10/2015).
- (2001). «Sculpt le son». In: *Portrait(s) de Iannis Xenakis*. A cura di F. B. Mâche. Paris: Bibliothèque Nationale de France, pp. 133–142.
- (2004). «Apollon and Dionysos. The writings of Xenakis». In: *Journal of New Music Research* 33.2, pp. 125–136.

- (2005a). «An Introduction to Horacio Vaggione's Musical and Theoretical Thought». In: *Contemporary Music Review* 24.4-5, pp. 311–326.
- (2005b). «Cellular automata in Xenakis's music. Theory and practice». In: *International Symposium Iannis Xenakis*. A cura di A. Georgaki e M. Solomos, pp. 120–137.
- Horacio Vaggione: Composition Theory* (2005c). In: *Contemporary Music Review* vol. 24.4-5. A cura di Makis Solomos. Numero monografico dedicato a Horacio Vaggione.
- Solomos, Makis (2005d). «Notes sur Agostino Di Scipio et sur la notion d'émergence». In: *Actes des Journées d'Informatique Musicale (JIM)*. A cura di A. Sedes e H. Vaggione. Paris: Université Paris 8 CICM MSH Paris Nord Afim, pp. 101–109.
- (2006). «Xenakis-Varèse et la question de la filiation». In: *Edgar Varèse. Du son organisé aux arts audio*. A cura di T. Horodyski e P. Lalitte. Paris: L'Harmattan, pp. 139–170.
- (2013). «Note sul concetto di 'emergenza' e su Agostino Di Scipio». In: *Agostino Di Scipio. Polveri sonore: una prospettiva ecosistemica della composizione*. A cura di Andrea Semerano. Roma: La Camera verde, pp. 107–116.
- Stevens, Stanley Smith e Hallowell Davis (1938). *Hearing: Its Psychology and Physiology*. New York: John Wiley & Sons.
- Stockhausen, Karlheinz (1957). «die Zeit vergeht...» In: *Die Reihe* 3, pp. 10–43.
- (1996). *Libretto del CD Elektronische Musik 1952-1960*. Stockhausen Gesamtausgabe.
- Teresa Davis, Mara Helmuth e (2004). «Rock Music: Granular and Stochastic Synthesis based on the Matanuska Glacier». In: *Proceedings of the 2004 International Computer Music Conference*. San Francisco: International Computer Music Association, pp. 619–622.
- Truax, Barry (1988). «Real-time granular synthesis with a digital signal processor». In: *Computer Music Journal* 12.2, pp. 14–26.
- Vaggione, Horacio (1983). «Poly-phonies: critique des relations et processus compositionnels». PhD. Paris: Université Paris 8.
- (1989). «Dimensions fractionnaires en composition musicale».
- (1996). «Articulating Micro-Time». In: *Computer Music Journal* 20.1, pp. 33–38.
- (1998a). «Son, temps, objet, syntaxe. Vers une approche multi-échelle dans la composition assistée par ordinateur». In: *Cahiers de philosophie du langage* 3, pp. 169–202.
- (1998b). «Transformations morphologiques: quelques exemples». In: *Actes des Journées d'Informatique Musicale (JIM) 1998*. Marseille: LMA-CNRS, G1–1–G1–10.
- (2003). «Composition musicale et moyens informatiques: questions d'approche». In: *Formel-informel: musique-philosophie*. A cura di M. Solomos, A. Soulez e H. Vaggione. Paris: L'Harmattan, pp. 91–116.
- (2005). «Notes on Atem». In: *Contemporary Music Review* 24.4-5, pp. 339–350.
- (2006). «Symboles, signaux, opérations musicales».
- Varela, Francisco J. (1996). *Invitation aux sciences cognitives*. Paris: Seuil.
- Varga, Bálint A. (1996). *Conversations with Iannis Xenakis. For amplified ensemble*. London: Faber e Faber.
- Winckel, Fritz (1960). *Vues nouvelles sur le monde des sons*. Trad. da A. Moles e J. Lequeux. Paris: Dunod.
- Xenakis, Iannis (1955). «La crise de la musique sérielle». In: *Gravesaner Blätter* 1, pp. 2–24.
- (1956). «Wahrscheinlichkeitstheorie und Musik». In: *Gravesaner Blätter* 6, pp. 28–34.
- (1958-59). «De tre parablerna». In: *Nutida Musik* 4, pp. 9–11.
- (1960a). «Elements of Stochastic Music (1)». In: *Gravesaner Blätter* 18, pp. 84–105.
- (1960b). «Elements of Stochastic Music (2)». In: *Gravesaner Blätter* 19-20, pp. 140–150.

- Xenakis, Iannis (1961a). «Elements of Stochastic Music (3)». In: *Gravesaner Blätter* 21, pp. 113–121.
- (1961b). «Elements of Stochastic Music (4)». In: *Gravesaner Blätter* 22, pp. 144–145.
- (1962). «Stochastic Music». In: *Gravesaner Blätter* 23-24, pp. 169–184.
- (1963). *Musiques formelles*. Revue Musicale 253-254. Numero monografico. Paris.
- (1971). *Formalized Music*. Trad. da C. Butchers, G.H. Hopkins e J. Challifour. Blooming-ton: University Press.
- (1976). *Musique. Architecture*. Tournai: Casterman.
- (1985). «Condition du musicien». In: *France Forum* 223-224. Ripreso in (Xenakis 1994, pp. 121-128).
- (1988). «Sur le temps». In: *Redécouvrir le temps (tome II)*. Bruxelles: Université de Bruxelles, pp. 193–200.
- (1992). *Formalized Music*. A cura di Sharon Kanach. Versione rivista e ampliata di (Xenakis 1971). Stuyvesant (New York): Pendragon Press.
- (1994). *Kéleütha*. Paris: L’Arche.
- (2003). *Universi del suono. Scritti e interventi 1955-1994*. A cura di Agostino Di Scipio. Lucca/Milano: LIM/Le Sfere.

## *Interagire con la complessità sonora interna ed esterna: dal microsuono alla composizione con suoni del paesaggio*

Barry Truax\*

(Versione rivista ricevuta l'11 settembre 2015<sup>1</sup>)

### Sommario

È possibile pensare ai due estremi del mondo sonoro come al dominio interno del microsuono (inferiore a 50 ms), da una parte, dove frequenza e tempo sono interdipendenti, e al mondo esterno della complessità sonora, ossia il paesaggio sonoro, dall'altra. In termini di progettazione sonora, l'elaboratore fornisce sempre maggiori strumenti per affrontare ognuno di questi domini, attraverso pratiche quali la sintesi granulare e la composizione multicanale con suoni del paesaggio.

## 1 Microsuono

Uno dei più sorprendenti sviluppi degli ultimi decenni è stato spostare le frontiere dei modelli del suono e della musica al micro livello; è diventato abbastanza comune riferirsi a questo fenomeno con il termine «microsuono» (Roads 2001). A questo livello, i concetti di frequenza e di tempo sono collegati da relazioni quantistiche, con un principio di indeterminazione che le rende interdipendenti in maniera precisamente analoga al più famoso principio di indeterminazione della fisica quantistica. Dennis Gabor formulò questo principio quantistico del suono nel 1947<sup>2</sup> nella sua critica al teorema di Fourier in quanto rappresentazione «senza tempo» [*timeless*], vale a dire priva di determinazione della variabile temporale.

Gabor illustrò il quanto acustico come un'area rettangolare nel dominio del tempo e della frequenza, in modo tale che mentre la durata del suono diminuisce il suo spettro nel dominio della frequenza si allarga. In altre parole, un suono sinusoidale la cui durata è inferiore a 50 ms diventa sempre più un click a banda larga mentre la sua durata diminuisce. Al contrario, per restringere l'indeterminazione nell'analisi della frequenza, è richiesta una “finestra” di durata maggiore, sia per l'analisi sia per la sintesi. Il sistema uditivo bilancia la sua risoluzione temporale e in frequenza in maniera significativa per la percezione dei fonemi linguistici, dove il riconoscimento simultaneo di forme sia spettrali sia temporali gioca un ruolo cruciale nella rapida identificazione del parlato. L'analogia con il principio di indeterminazione di Heisenberg della meccanica quantistica non è metaforica ma esatta, poiché come la velocità è il tasso di cambiamento della posizione (da cui l'accuratezza nel determinare l'una è legata

---

\* Simon Fraser University – Vancouver.

<sup>1</sup> Il presente articolo è stato pubblicato online nel dicembre 2014 in lingua inglese al seguente url: [http://econtact.ca/16\\_3/truax\\_soniccomplexity.html](http://econtact.ca/16_3/truax_soniccomplexity.html). La versione fornita dall'autore presenta minime differenze rispetto all'originale pubblicato, senza cambiarne tuttavia il senso. Traduzione italiana e annotazioni fra parentesi quadre a cura di Paolo Zavagna.

<sup>2</sup> Trad. it. in questo numero, pp. 9-16. N. d. T.

alla mancanza di accuratezza nel determinare l'altra) così la frequenza può essere pensata come il tasso di cambiamento della fase.

I modelli tempo-frequenza – una classe di metodi di sintesi del suono e di elaborazione del segnale che è emersa negli ultimi due decenni – hanno le loro basi in questo livello quantistico, in modo tale che cambiamenti del segnale nel dominio del tempo provocano alterazioni spettrali e viceversa. Tra questi, sono ben conosciuti i metodi di sintesi granulare e di elaborazione granulare, che producono i loro risultati per mezzo della generazione di un'alta densità di quanti acustici chiamati grani. Questi grani sono composti da forme d'onda inviluppate, normalmente più brevi di 50 ms (che significa un tasso di ripetizione maggiore di 20 Hz), tali che una sequenza di grani si fonde in un suono continuo, così come la percezione dell'altezza emerge dagli impulsi che si ripetono a tassi superiori ai 20 Hz. I cosiddetti “grani di Gabor” hanno la frequenza della forma d'onda indipendente dalla durata del grano, mentre le “wavelets” mantengono una relazione inversa fra frequenza e durata, e per questo sono utili nei modelli di analisi e risintesi (Roads 1996).

Tuttavia, diversi altri metodi di sintesi riconosciuti sono oggi considerati modelli tempo-frequenza, ad esempio il VOSIM [VOice SIMulaton] e il FOF [Fonction d’Onde Formantique], entrambi originariamente progettati per la simulazione del parlato. Ciascuno di essi è basato su una forma d'onda con inviluppo ripetuta – un impulso sinusoidale riquadrato con una componente DC nel caso del VOSIM, e un'onda sinusoidale asimmetrica con inviluppo nel caso del FOF. Inoltre, sono i parametri utilizzati nel dominio del tempo in ciascun modello che controllano la larghezza di banda del risultato, solitamente concepita per modellare le formanti delle vocali simulate. Michael Clarke (1996) ha compreso subito il rapporto fra il metodo FOF e la sintesi granulare, e ha proposto una versione ibrida chiamata FOG [FOF Granular]. Nel suo lavoro, un suono basato sulla fusione di un formante può disintegrarsi in uno schema ritmico o in una texture granulare e ritornare al suono originale, pur mantenendo nel processo la coerenza della fase.

Nel mio lavoro, il concetto di granularità ha informato molte delle manipolazioni del suono campionario, la più sorprendente delle quali è la possibilità di allungare [*stretching*] il suono nel tempo senza necessariamente cambiarne l'altezza (Truax 1990, 1992a, 1994). È una rivelazione paradossale che collegando tempo e frequenza al micro livello li si possa manipolare indipendentemente al macro livello. In effetti, tutti i metodi attuali per allungare il suono si basano su qualche forma di operazione di finestratura, normalmente con inviluppi che si sovrappongono e la cui forma e frequenza di ripetizione sono controllabili. L'effetto percettivo dell'allungamento temporale è anche molto suggestivo. Nel momento in cui la forma temporale di un suono si dilata, sia di una piccola percentuale sia di un valore molto grande, l'attenzione si sposta verso le componenti spettrali del suono, siano frequenze discrete, armoniche o inarmoniche, siano regioni risonanti o texture a banda larga. Spesso mi riferisco a questo processo come all'ascolto “dentro” il suono, e tipicamente collogo le altezze che emergono dallo spettro con quelle utilizzate da esecutori dal vivo (come nelle mie opere miste *Dominion* [1991], *Powers of Two: The Artist* [1995], o nel recente lavoro per pianoforte e tracce audio che si basa sull'analisi spettrale, *From the Unseen World* [2012]). In altri casi, le risonanze espansse anche di un semplice parlato o di un suono ambientale suggeriscono un'esaltazione delle immagini e delle possibili associazioni mentali, come nel mio lavoro *Basilica* [1992], dove le risonanze dilatate [*stretched*] della campana suggeriscono di essere dentro il vasto volume della chiesa stessa.

## 2 Convoluzione: collegare i dominii di tempo e frequenza

Per tornare al dominio del microsuono, possiamo notare che un altro principio fondamentale che collega i dominii di tempo e frequenza è illustrato dalla tecnica della convoluzione. La convoluzione è stata per un po' di tempo un tipico argomento in ingegneria e informatica, ma soltanto dall'inizio degli anni novanta è stata ampiamente disponibile ai compositori di musica digitale, in gran parte grazie alle descrizioni teoriche di Curtis Roads (1996) e al programma *SoundHack* di Tom Erbe, che ha reso questa tecnica accessibile.

Effettuare la convoluzione fra due forme d'onda nel dominio del tempo significa moltiplicarne i rispettivi spettri (cioè il contenuto in frequenza) nel dominio della frequenza. Moltiplicare gli spettri significa che ogni frequenza forte in *entrambi* i segnali diventerà molto forte nel segnale convoluto, e al contrario ogni frequenza debole in entrambi i segnali in ingresso sarà debole nel segnale in uscita. In pratica, un'applicazione relativamente semplice della convoluzione è data dalla risposta all'impulso di uno spazio. Questa si ottiene registrando un segnale a banda larga molto breve, esplosivo, mentre viene elaborato dalle caratteristiche di riverberazione dello spazio in cui viene prodotto. Quando effettuiamo la convoluzione di un qualsiasi segnale anecoico con la risposta all'impulso di uno spazio, il risultato è che il suono sembra essere stato registrato in quello spazio. In altre parole, è stato elaborato dalla risposta in frequenza dello spazio nello stesso modo in cui sarebbe stato elaborato nello spazio reale. Infatti, la convoluzione, in questo esempio, è semplicemente una descrizione matematica di quello che accade quando un qualsiasi suono viene colorato dallo spazio acustico in cui viene emesso, ciò che accade realmente a tutti i suoni in qualsiasi spazio fuorché in camera anecoica. Il suono convoluto sembrerà inoltre collocarsi alla stessa distanza dell'impulso della registrazione originale. Se effettuiamo la convoluzione di un suono per due volte con la stessa risposta all'impulso, la sua distanza apparente sembrerà doppia. Ad esempio, il mio lavoro *Temple* [2002] si basa sull'elaborazione di voci cantate con la risposta all'impulso di una cattedrale italiana, quella di San Bartolomeo a Busseto, disponibile nell'effetto "Impulseverb" del software *Peak*. Data l'invasione del colpo di pistola, per registrare la risposta all'impulso di uno spazio è più accettabile lo scoppio di un palloncino.

Inoltre si può effettuare la convoluzione di un qualsiasi suono con un qualsiasi altro suono, non soltanto tramite risposta all'impulso. In questo caso accade che il primo suono viene filtrato attraverso lo spettro del secondo, in modo tale che le frequenze comuni a entrambi i suoni verranno rinforzate. Un caso particolare è quando effettuiamo la convoluzione di un suono con se stesso, chiamata auto-convoluzione, garantendo in questo modo la massima correlazione fra le due sorgenti. In questo caso, le frequenze più evidenti verranno esagerate e le frequenze con poca energia verranno attenuate. In ogni caso, la durata del segnale in uscita è la somma delle durate dei due ingressi. Con la riverberazione ci aspettiamo che il segnale riverberato sia più lungo dell'originale, ma questo prolungamento e la risultante sbavatura dei transitori d'attacco avvengono anche quando effettuiamo la convoluzione di un suono con se stesso, o con un altro suono. I transitori sono smorzati, e il suono complessivo è allungato (di un fattore due nel caso dell'auto-convoluzione). Quando effettuiamo la convoluzione di questa versione allungata con la risposta all'impulso di uno spazio, il risultato sembra una via di mezzo fra l'originale e la versione riverberante, una versione fantasmatica del suono, per così dire. In *Temple*, sia il suono originale sia la versione convoluta con se stessa sono convolute con la risposta all'impulso della cattedrale e sincronizzate per cominciare insieme, producendo così un'immagine residua nella coda di suono riverberante.

Un altro approccio utilizzato nel mio pezzo *Chalice Well* [2009] è consistito nell'effettuare

la convoluzione fra diversi tipi di suoni sgranati [*textured*], di qualità particellare [*particulate*], principalmente gocce d'acqua, spruzzi, ruscelli e gocciolii, ma anche vetri che si rompono, bolle, colpi di serrature e porte, consonanti maschili trasposte, e varie densità di trame [*textures*] di sintesi granulare (come in *Riverrun* [1986]) sia come sorgenti sia come impulsi. Normalmente, quando si effettua la convoluzione di due suoni continui, il risultato è una trama sbavata e piuttosto sfocata, ma poiché questi suoni erano composti da numerosi impulsi, il risultato della convoluzione era altamente dettagliato e ben definito. Inoltre, le qualità spaziali dei file sonori originali andavano da secco a riverberante, e quindi ogni combinazione produceva un senso dello spazio ben definito: secco convoluto con secco produce una trama in primo piano, riverberante con riverberante una trama distante, di sottofondo, e secco con riverberante una collocazione del suono a metà dello spazio. Ogni combinazione (ne sono state utilizzate circa 200) potrebbe essere descritta come un suono ibrido, situato in qualche luogo fra il sintetico e l'elaborato e che incorpora elementi di entrambi i genitori. Sono state create e documentate intere famiglie di trame di suoni, dato che ogni possibile permutazione dei sorgenti ha fornito un risultato diverso ma correlato con essi.

Mentre sperimentavo questa tecnica, ho cominciato rapidamente a percepire il tipo di paesaggio sonoro immaginario, suggeritomi dai risultati ottenuti, come una caverna sotterranea nella quale fossero presenti vari tipi di acqua corrente. Una serie di spruzzi e gocciolii d'acqua è stata infatti registrata in un pozzo risonante (ed era già stata utilizzata nella terza sezione del mio brano *Island* [2000]) e questi suoni mi fanno pensare all'aura che circonda pozzi e caverne in generale, e il pozzo di Chalice a Glastonbury in particolare. È sorta allora l'idea narrativa di una composizione con suoni del paesaggio che potesse simulare un viaggio giù nel pozzo, in caverne leggendarie e altamente simboliche (la cui effettiva esistenza non è mai stata provata).

La dislocazione spaziale delle tracce venne ottenuta con quattro insiemi di otto tracce ciascuno, ovvero quattro coppie stereo distribuite attorno all'ascoltatore. Grazie alla natura di famiglia di ogni insieme di otto tracce, e delle loro individuali qualità spaziali in termini di distanza apparente, è stato relativamente semplice creare un'illusione di uno spazio coerente di grandi dimensioni nonché di presenza. Le 32 tracce sono state mixate in otto in un formato circolare, anche se, quando eseguito in sale dotate di altoparlanti a vari livelli verticali (come ad esempio la prima esecuzione al Sonic Arts Research Centre a Belfast nel marzo 2009), risulta molto efficace raddoppiare le otto tracce su più livelli di altoparlanti, sostenendo l'illusione di trovarsi in una caverna sotterranea.

### 3 Composizione con suoni del paesaggio

La composizione con suoni del paesaggio, sullo sfondo concettuale e interdisciplinare degli studi sul paesaggio sonoro e della comunicazione acustica, nonché dei mezzi tecnici del *time-stretching* granulare (Truax 1988, 1990, 1998) e della diffusione multi canale (Truax 1998), tutti concetti sviluppati alla Simon Fraser University negli ultimi quarant'anni, fornisce un modello maturo dell'uso musicale di suoni ambientali. Altrove ho descritto l'equilibrio ideale che dovrebbe essere raggiunto in questo tipo di lavori come l'incontro della complessità interna dell'organizzazione sonora con la complessità esterna delle relazioni nel mondo reale, senza che l'una sia subordinata all'altra (Truax 1992b, 1994). Infatti, queste fonti di complessità interna ed esterna rispecchiano l'interpretazione quotidiana dei suoni da parte degli ascoltatori attraverso le loro abilità nel riconoscere ed interpretare schemi sonori a diverse scale temporali,

e collegarli alle conoscenze contestuali acquisite dalla loro esperienza nel mondo reale. Il ruolo del compositore del paesaggio sonoro riguardo a quest'ultimo aspetto può essere pensato come il miglioramento del funzionamento interno dei suoni in modo da riflettere i loro significati contestuali, invece che considerarli come entità astratte in relazione soltanto fra loro (Truax 2008, 2013).

Ho anche suggerito (Truax 1996, 2002) quali sono le caratteristiche principali della composizione con suoni del paesaggio come si sono sviluppate nella pratica:

1. il mantenimento della riconoscibilità della sorgente sonora da parte dell'ascoltatore, anche se viene successivamente soggetta a trasformazione;
2. la conoscenza da parte dell'ascoltatore del contesto ambientale e psicologico del materiale del paesaggio sonoro è evocata e incoraggiata per completare la rete di significati ascrivibili alla musica;
3. la conoscenza da parte del compositore del contesto ambientale e psicologico del materiale del paesaggio sonoro consente di influenzare la forma della composizione ad ogni livello, e in definitiva la composizione è inseparabile da alcuni o da tutti questi aspetti della realtà; e idealmente:
4. l'opera migliora la nostra conoscenza del mondo, e la sua influenza si ripercuote nelle abitudini percettive quotidiane.

Quindi, il vero obiettivo della composizione con suoni del paesaggio è la re-integrazione dell'ascoltatore con l'ambiente in un'*'equilibrata relazione ecologica'*.

Dati questi propositi, i vari tipi di elaborazione digitale e le tecniche di diffusione controllate dall'elaboratore oggi disponibili formano un potente strumento per il compositore per creare paesaggi sonori sia mimetici sia extrapolati<sup>3</sup>, aiutati dalla diffusione multicanale immersiva, paesaggi che vanno dall'imitativo al completamente virtuale come quello che si può trovare nel sopra descritto *Chalice Well*. La caratteristica implicita nel termine «extrapolata» è che la progettazione sonora del compositore attraverso l'elaborazione fa emergere in qualche modo gli aspetti interni dei suoni utilizzati, esaltandoli, piuttosto che cancellando l'identità del materiale sorgente e/o imponendo elaborazioni artefatte senza nessuna relazione con i suoni elaborati. La distinzione è spesso sottile, poiché tecniche di elaborazione simili possono essere utilizzate sia per risultati astratti sia per risultati extrapolati. A volte è una questione di grado di elaborazione, altre volte (nella composizione con suoni del paesaggio) di relazione del suono elaborato con l'originale che può non essere chiara all'ascoltatore, né deve esserlo. Comunque, anche una relazione "nascosta" può rivelare alcune connessioni con il mondo reale più profonde o più intuitive. Ad esempio, la versione elaborata delle onde e del fiume nelle prime due sezioni del mio pezzo *Island* suona deliberatamente irrealistica se confrontata alle registrazioni ravvicinate altamente realistiche dei suoni originali. I suoni di onde sono elaborati con filtri passa-basso e passa-alto con un bel po' di *feedback* nel risuonatore utilizzato, in modo che la versione filtrata con il passa-basso suoni come un suono continuo [*drone*], ma il materiale originale entrando nel risuonatore mantiene ancora la regolarità ritmica delle onde originali. Per me, il *drone* pulsante che si sente in contrappunto con le onde reali suggerisce all'ascoltatore di essere in visita in un'isola straniera o addirittura magica, aperta alla nostra

---

<sup>3</sup>Per quanto riguarda i termini mimetico (*mimetic*) e extrapolato (*abstracted*) si veda Emmerson (1986). Trad. it. Emmerson (1995). N. d. T.

immaginazione. Allo stesso modo, i suoni del fiume allungati e risonanti creano ulteriori formanti che ricordano il canto corale, come se voci di Sirene immaginarie ci invitassero a continuare il nostro viaggio lungo il fiume. In altri punti del pezzo, suoni risonanti e allungati semplicemente migliorano il materiale originale e sono spesso miscelati perfettamente con esso.

L'auto-convoluzione come descritta sopra fornisce un tipo complementare di elaborazione che raddoppia la durata del suono ed enfatizza le sue frequenze dominanti, mentre attenua quelle più deboli. Spesso assomiglia a un decadimento riverberante e persistente, come in *Temple*. Con i miei recenti lavori *Aeolian Voices* [2013] e *Earth and Steel* [2013], l'iterazione della convoluzione non solo prolunga i suoni originali (una macchina di passaggio o un suono percussivo attivato dal vento che fischia attraverso un capannone nel primo pezzo, o i suoni delle costruzioni navali in acciaio nel secondo), ma, se utilizzata da sola, senza originale, suggerisce, almeno per me, un ricordo sfocato ed esteso di quei suoni. Questo effetto è particolarmente toccante in *Earth and Steel* a causa del contesto storico del pezzo, che tentava la ri-creazione di un'epoca passata della cantieristica navale, che è stata intensificata alla prima esecuzione (2013) avvenuta in un edificio reale presso il sito dei cantieri navali della Royal Navy in Kent, Regno Unito, ora sede di un museo.

Durante i miei ultimi venticinque anni di lavoro col microsuono o con metodi tempo-frequenza, mi sono affidato principalmente a queste tre tecniche (allungamento temporale tramite granulazione, risuonatori e convoluzione) perché ognuna di esse mette in evidenza le qualità interne dei suoni ambientali o dei suoni strumentali e vocali registrati in studio. Alle volte mi riferisco alla natura di questi processi come a qualcosa che mi permette di comporre “attraverso” il suono, non soltanto “con” esso.

## 4 Conclusionе

La lunga storia della scienza al servizio dell'arte sta chiaramente continuando oggi, attraverso uno dei sentieri più intriganti, che coinvolge il livello quantistico del microsuono, ciò che possiamo chiamare “la frontiera finale” dell’acustica e della ricerca musicale. L’arte al servizio della scienza ha contributi minori, ma ancora significativi, come quello della visualizzazione artistica e della sonificazione musicale delle basi di dati. L’elaboratore è al centro di entrambi i tipi di processo. Tuttavia, come in ogni stretta relazione, entrambi i partner possono cambiare grazie al reciproco incontro. In effetti, potremmo analizzare l’esperienza artista-macchina lungo due percorsi: uno in cui la tecnologia gioca un ruolo passivo limitandosi a fare da assistente al processo produttivo (ciò che io chiamo composizione realizzata mediante elaboratore), l’altro, in cui essa vi partecipa in una maniera tale da cambiare il vocabolario artistico, il processo e il risultato finale. Quest’ultimo tipo di processo spazia dalla collaborazione interattiva, che potremmo chiamare composizione assistita dall’elaboratore, fino a un tipo di lavoro completamente automatizzato, basato su regole, composto dall’elaboratore. Ho suggerito in queste pagine che, da quando l’elaboratore è utilizzato per controllare la complessità che risulta da nuove forme, il ruolo dell’artista appare profondamente cambiato secondo vari scenari possibili: guida, sperimentatore, progettista, visionario, poeta, solo per nominarne alcuni. Nel mio lavoro, ho sperimentato un po’ tutti questi ruoli, ma ciò che li tiene insieme è il mio mettere in relazione la complessità interna del micro dominio del mondo sonoro, con la complessità esterna del mondo reale in tutte le sue dimensioni naturali, umane e sociali. È un viaggio che mi ispira particolarmente.

## Bibliografia

- Clarke, Michael (1996). «Composing at the intersection of time and frequency». In: *Organised Sound* 1.2, pp. 107–117.
- Emmerson, Simon (1986). «The Relation of Language to Materials». In: *The Language of Electroacoustic Music*. A cura di Simon Emmerson. Trad. it. (Emmerson 1995). Hounds mills e London: The Macmillian Press Ltd., pp. 18–39.
- (1995). «La relazione tra linguaggio e materiali nella musica elettroacustica». In: *Teoria e prassi della musica nell'era dell'informatica*. A cura di Agostino Di Scipio. Bari: Giuseppe Laterza Editore, pp. 53–77.
- Gabor, Dennis (1947). «Acoustical quanta and the theory of hearing». In: *Nature* 159.4044, pp. 591–594.
- Roads, Curtis (1996). *The Computer Music Tutorial*. Cambridge, London: The MIT Press.
- (2001). *Microsound*. Cambridge, London: The MIT Press.
- Truax, Barry (1988). «Discovering inner complexity: Time-shifting and transposition with a real-time granulation technique». In: *Computer Music Journal* 18.2, pp. 38–48.
- (1990). «Composing with Real-Time Granular Sound». In: *Perspectives of New Music* 28.2, pp. 120–134. JSTOR: 833014.
- (1992a). «Composing with time-shifted environmental sound». In: *Leonardo Music Journal* 2.1, pp. 37–40.
- (1992b). «Electroacoustic music and the soundscape: the inner and outer world». In: *Companion to Contemporary Musical Thought*. A cura di J. Paynter et al. London: Routledge, pp. 374–98.
- (1994). «The Inner and Outer Complexity of Music». In: *Perspectives of New Music* 32.1, pp. 176–193.
- (1996). «Soundscape, acoustic communication & environmental sound composition». In: *Contemporary Music Review* 15.1, pp. 49–65.
- (1998). «Composition and diffusion: Space in sound in space». In: *Organised Sound* 3.2, pp. 141–6.
- (2002). «Genres and techniques of soundscape composition as developed at Simon Fraser University». In: *Organised Sound* 7.1, pp. 5–14.
- (2008). «Soundscape Composition as Global Music: Electroacoustic Music as Soundscape». In: *Organised Sound* 13.2, pp. 103–109.
- (2013). *From Epistemology to Creativity: A Personal View*. URL: <http://journal.sonicstudies.org/vol04/nr01/a02> (visitato il 07/07/2015).



## *Elaborazione di segmenti temporali*

Pierre Dutilleux, Giovanni De Poli, Adrian von dem Knesebeck e Udo Zölzer\*

(Traduzione ricevuta il 7 gennaio 2016<sup>1</sup>)

### Sommario

Il testo si occupa delle elaborazioni del suono nel dominio del tempo discreto, considerando una fase di segmentazione, una successiva di analisi, ed una strategia di ricomposizione del segnale temporale al fine di ottenere trasformazioni sull'audio risintetizzato. I vari algoritmi sono presentati illustrando le tecniche analogiche storiche oltre a sviluppi e implementazioni mediante le tecniche dell'audio digitale. Sono altresì citati brani antologici dove compositori e musicisti utilizzano tali elaborazioni, contribuendo ad illustrare l'evoluzione delle tecniche della musica elettronica e dell'elaborazione sonora.

## 1 Introduzione

In questo capitolo discutiamo diversi algoritmi nel dominio del tempo in quanto combinazioni di blocchi di elaborazione più piccoli, come modulatori di ampiezza/fase, filtri e linee di ritardo. Questi effetti influenzano principalmente l'altezza e la durata temporale del segnale audio. Prima introduciamo alcuni effetti di base come la riproduzione a velocità variabile ed il ricampionamento con controllo dell'altezza. Questi effetti sono basati sulla modulazione mediante linee di ritardo e sulla modulazione di ampiezza. Poi parleremo di due approcci per la dilatazione temporale (scalatura temporale) di segnali audio. Essi si basano su una fase di analisi, in cui il segnale di ingresso è diviso in segmenti (blocchi) di lunghezza fissa o variabile, e una fase di sintesi in cui i blocchi della fase di analisi vengono ricombinati da procedure di sovrapposizione e somma. Queste tecniche di dilatazione temporale ottengono una scalatura temporale senza modificare l'altezza del segnale. La quarta sezione si concentra sulla modifica dell'altezza, e introduce tre tecniche: elaborazione del blocco in base alla dilatazione temporale e ricampionamento, modulazione della linea di ritardo ed elaborazione del blocco sincronizzata con il periodo. L'elaborazione del blocco basato sulla modulazione della linea di ritardo ottiene la modifica dell'altezza scalando l'inviluppo spettrale di ogni blocco. L'elaborazione del blocco sincronizzata con il periodo ottiene la modifica dell'altezza dal ricampionamento dell'inviluppo spettrale di ogni blocco, preservando l'inviluppo spettrale originale. L'ultima sezione presenta un uso più creativo della elaborazione nel tempo dei segmenti quali la ricombinazione nel tempo e la granulazione. Brevi segmenti del segnale di ingresso sono liberamente assemblati e posizionati nell'asse temporale del segnale di uscita.

\*Energy Engineering GmbH, DEI – Università di Padova, Department of Signal Processing and Communication – Helmut-Schmidt University di Hamburg, Department of Signal Processing and Communication – Helmut-Schmidt University di Hamburg.

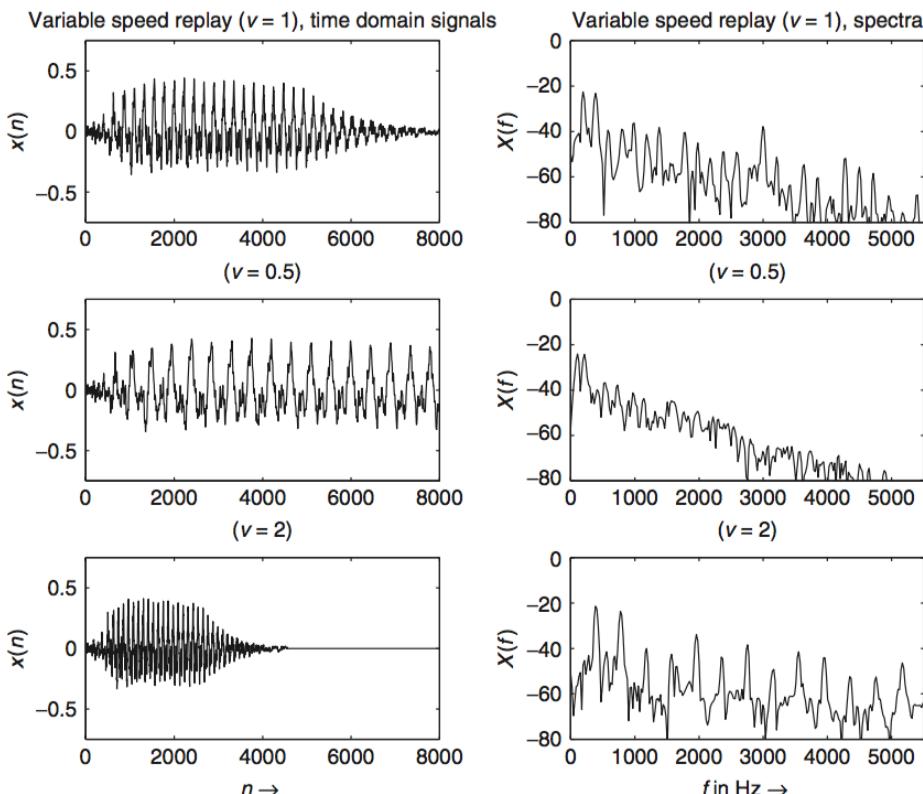
<sup>1</sup>Il presente articolo è la traduzione, a cura di Roberto Neri, di Pierre Dutilleux et al. (2011). «Time-segment processing». In: *Digital Audio Effects*. A cura di Udo Zölzer. 2<sup>a</sup> ed. John Wiley & Sons, Ltd. Cap. 6, pp. 185–217. Revisione della traduzione di Giovanni De Poli ricevuta il 31 marzo 2016.

In questo caso il suono in ingresso può essere molto meno riconoscibile in quello di uscita. L'ampia scelta di strategie di organizzazione dei segmenti temporali implica un atteggiamento alla composizione sonora da parte dell'utente.

## 2 Riproduzione a velocità variabile

### Introduzione

I registratori audio analogici permettono la riproduzione con un ampio intervallo [range] di velocità. In particolare, l'avvolgimento veloce in avanti o indietro, se il monitoraggio del segnale audio è possibile, può essere utilizzato per localizzare un suono. Durante la riproduzione a maggiore velocità, l'altezza del suono cresce e durante la riproduzione a minore velocità, l'altezza del suono diminuisce. Con questa tecnica la durata del suono è più lunga se il regista riproduce più lentamente ed è più corta se la velocità del regista viene aumentata. La Fig. 1 illustra un segmento di suono che è stato allungato ed accorciato, e gli spettri corrispondenti.



**Figura 1:** Pitch shifting: La riproduzione a velocità variabile comporta una compressione/espansione temporale e compressione ed espansione dell'inviluppo spettrale.

## Elaborazione del segnale

La frase “riproduzione a velocità variabile” denota che quello che è successo inizialmente durante l’istante  $nT_{s,in}$  sta adesso accadendo nel tempo

$$nT_{s,replay} = nT_{s,in}/v \quad (1)$$

alla velocità relativa  $v$ , dove  $nT_{s,in}$  e  $nT_{s,replay}$  sono i periodi di campionamento iniziale e di riproduzione. L’espansione temporale corrisponde a  $v < 1$ . Un semplice metodo per implementare la riproduzione a velocità variabile è perciò la modifica della frequenza di campionamento durante la riproduzione del suono in accordo con

$$f_{s,replay} = f_{s,in} \cdot v \quad (2)$$

dove  $f_{s,in}$  and  $f_{s,replay}$  sono le frequenze di campionamento iniziale e di riproduzione. Si dovrebbe distinguere il caso in cui l’uscita sia digitale o analogica. Se l’uscita è analogica, allora il reale metodo effettivo è quello di modificare la frequenza di campionamento del DAC di uscita. Lo spettro del segnale è scalato di  $v$  ed il filtro analogico di ricostruzione dovrebbe essere realizzato per rimuovere le immagini spettrali dopo la conversione (Gaskell 1987; Massie 1998).

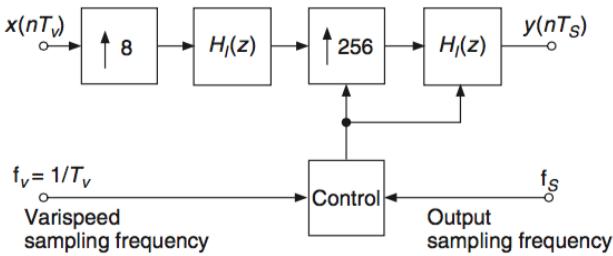
Se è richiesta una uscita digitale, allora la frequenza di conversione deve essere regolata tra la frequenza di riproduzione desiderata  $f_{s,replay}$  e la frequenza di campionamento di uscita  $f_{s,out}$ , che di solito è uguale a  $f_{s,in}$ .

Se  $v < 1$  (espansione temporale) allora  $f_{s,in} > f_{s,replay} < f_{s,out}$  e servono più campioni di uscita di quelli disponibili dal segnale di ingresso. Il segnale di uscita è una versione interpolata (sovra-campionata) di un fattore  $1/v$  del segnale di ingresso. Se  $v > 1$  (compressione temporale) allora  $f_{s,in} < f_{s,replay} > f_{s,out}$  e sono necessari meno campioni di uscita di quelli disponibili dal segnale di ingresso. Il segnale di ingresso è decimato di un fattore  $v$ . Prima della decimazione la larghezza di banda del segnale in ingresso deve essere ridotta a  $f_{s,replay}/2$  da un filtro passa basso digitale (McNally 1984). La qualità della conversione della frequenza di campionamento dipende fortemente dal filtro di interpolazione usato. Un metodo molto diffuso è l’interpolazione lineare tra due campioni adiacenti. Una disamina sui metodi di interpolazione può essere trovata in Crochiere e Rabiner (1983) e Massie (1998).

Una implementazione in tempo discreto può essere ottenuta incrementando o decrementando il tasso di trasferimento al DAC di un segnale audio digitale registrato, in modo da cambiare la frequenza di campionamento di uscita rispetto alla frequenza di campionamento di registrazione. Se il segnale di uscita dovrà essere ancora in formato digitale, dovremmo ricampionare il segnale analogico a velocità variata con la corrispondente frequenza di campionamento. Una implementazione tempo-discreto senza una conversione DA e una nuova conversione AD è stata proposta in McNally (1984) ed è mostrata nella Fig. 2. Questa fa uso di tecniche multi frequenza di elaborazione del segnale e si avvale di una approssimazione della conversione DA/AD. Un ulteriore algoritmo di elaborazione del segnale per ottenere il risultato acustico di una riproduzione a velocità variabile è una modulazione di una linea di ritardo con un cambiamento costante di altezza, che sarà discusso nella Sezione 4.3.

## Applicazioni musicali e controllo

Così come per l’editing audio nei magnetofoni, anche nei sistemi audio digitali la riproduzione con velocità variabile è molto usata. Vedi Wishart (1994, ID 2.9 e 2.10) per una chiara



**Figura 2:** Schema per la riproduzione a velocità variabile.

dimostrazione dell’effetto su un segnale vocale.

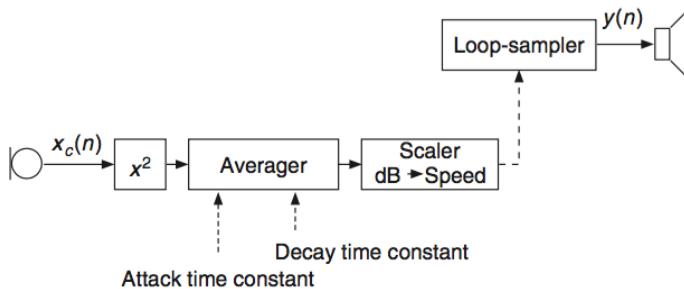
L’effetto di trasposizione tramite la velocità del magnetofono è stata usata da Les Paul nel brano intitolato *Whispering* nel 1951 (Lee 1972). Questo metodo è stato usato molto frequentemente nella musica elettroacustica quando l’altezza di un suono concreto non può essere controllata durante la registrazione. P. Schaeffer progetta il *Phonogène chromatique* per trasporre in ciascuno dei 12 gradi della scala temperata. I pinch rollers potevano essere controllati con una tastiera simile ad un pianoforte. Un ingranaggio aggiuntivo estendeva il range delle operazioni a due ottave (Geslin 2000; Moles 1961; Roads 1996). Jacques Poullin aveva sviluppata un’altra versione, il *Phonogène à coulisse*, che permetteva una modifica di velocità continua. Una coppia di coni, con una ruota di frizione in mezzo, costituiva un collegamento meccanico a rapporto variabile tra il motore ed il capstan del magnetofono. La posizione della frizione, e quindi la velocità di riproduzione, era controllata a livello meccanico. Stockhausen, in *Hymnen* traspose il suono dell’orchestra per conferirle un carattere travolgente ed apocalittico (Chion 1982).

Anche nella computer music, la riproduzione a velocità variabile fornisce un efficace sistema di trasposizione. J. C. Risset dice: «miscelando un suono con la trasposizione di se stesso con piccole differenze di frequenza (diciamo un ventesimo di Hertz) si possono convertire toni a periodo costante in un pattern dove gli armonici del tono crescono e cambiano con tassi differenti, proporzionalmente al grado dell’armonico» (Risset 1987, 1998). In *The gates of H* Ludger Brümmer esplora la proprietà che la riproduzione a velocità variabile ha di modificare sia l’altezza sia la durata di un campione (Brümmer 1993, 14’40”-17’25”). Sette copie della stessa frase suonate simultaneamente alle velocità di 7.56, 4.49, 2.24, 1.41, 0.94, 0.67, 0.42, 0.31 vengono sovrapposte. Il suono risultante inizia con una struttura complessa ed uno spettro esteso. Al proseguire del suono, le copie più veloci svaniscono e le versioni più lente emergono l’una dopo l’altra. La struttura del suono si semplifica progressivamente e si evolve verso i registri più bassi.

Il carattere dei suoni trasposti è modificato in quanto tutti gli aspetti dello spettro scalano simultaneamente. Le formanti sono scalate in alto, arrivando ad un effetto ‘Topolino’, o sono scalate verso il basso, come se i suoni fossero prodotti da oggetti ingigantiti. La struttura del tempo è altresì modificata. I transitori sono allungati o contratti. Un vibrato iniziale nel suono perderà il suo carattere ed apparirà come una modulazione o più lenta o più veloce. I suoni possono essere anche riprodotti a velocità negativa. Una velocità -1 produce un suono con lo stesso spettro medio, anche se suona in maniera molto diversa. Pensate ad esempio ad un parlato o ad un suono percussivo riprodotti al contrario. Altri schemi di trasposizione che sono

liberi da questi ostacoli possono essere ottenuti da metodi più sofisticati<sup>2</sup>.

Una particolare applicazione fu richiesta dal compositore Kiyoshi Furukawa. Lui volle un campionatore per il quale la velocità fosse controllata dall'ampiezza di uno strumento acustico. Un suono è memorizzato in un campionatore ed è suonato in loop. Nel frattempo, l'RMS dell'ampiezza di un segnale di controllo in ingresso è calcolato e mediato nel tempo con costanti di tempo di attacco e decadimento indipendenti. L'ampiezza è convertita in  $dB$  e scalata prima di controllare la velocità del campionatore. I parametri devono essere accordati in una maniera tale che la velocità rimanga all'interno di un intervallo valido e le variazioni di velocità siano intimamente correlate al volume ed alla dinamica dello strumento (vedi Fig. 3).



**Figura 3:** Un loop-sampler controllato da un segnale acustico.

Questo effetto è controllato da un clarinetto in *Swim*, *swan* e da una viola in *Den Ungebo-rennen Gotten* (Furukawa 1993, 1997). L'altezza dello strumento acustico seleziona parole da un insieme predefinito, mentre il volume controlla la velocità di riproduzione di queste parole.

### 3 Time Stretching [Allungamento o accorciamento temporale]

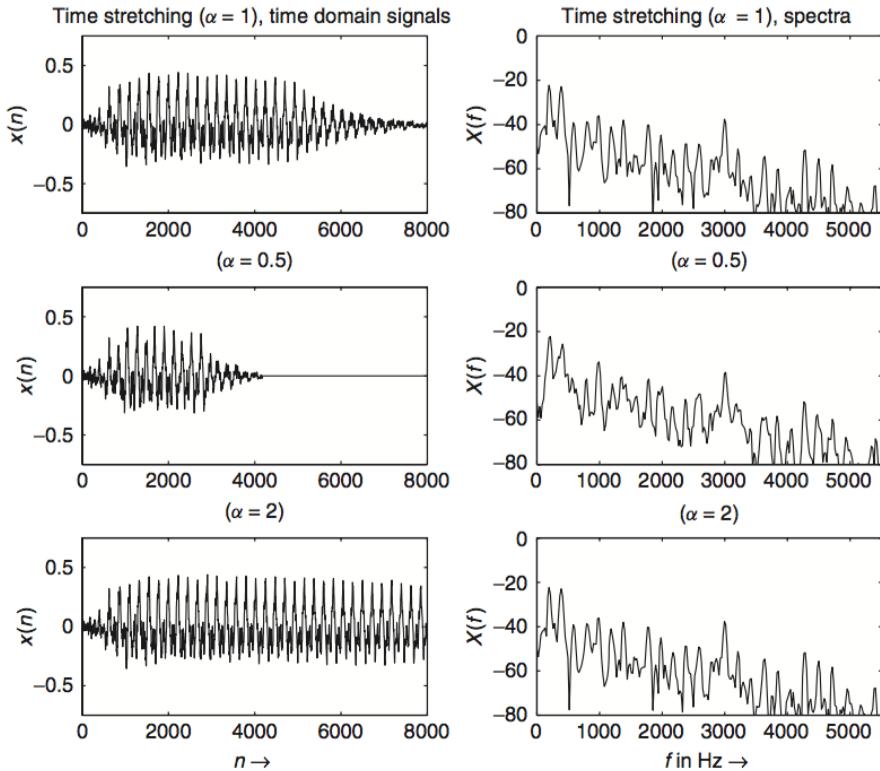
#### Introduzione

Al fine di comprendere il time stretching, prendiamo l'esempio di un segnale la cui durata non combacia con la porzione temporale che è stata riservata per il suo uso. Pensate ad uno speaker che ha già registrato 33 secondi di parlato, ma il cui contributo ad una pubblicità non dovrebbe essere superiore a 30 secondi. Se non vuole registrare nuovamente il suo discorso, il fonico può artificialmente contrarre la durata del suo parlato del 10%. Con il termine “time stretching” ci riferiamo alla contrazione o espansione della durata di un segnale audio. Abbiamo visto nel paragrafo precedente un metodo che altera la durata del suono, la riproduzione a velocità variabile ma questo ha l'inconveniente di trasporre contestualmente il suono. L'*Harmonizer* potrebbe essere usato per trasporre il suono in direzioni opposte e la combinazione di entrambi i metodi porta ad un algoritmo di time stretching.

L'obiettivo principale degli algoritmi di time stretching è di accorciare o allungare il file sonoro di  $M$  campioni in una nuova lunghezza desiderata  $M' = \alpha \cdot M$ , dove  $\alpha$  è il fattore di

<sup>2</sup>Tali metodi sono descritti nelle sezioni successive del libro da cui è tratto il presente testo (Zölzer 2011). N. d. T.

scala. Per realizzare algoritmi di time stretching il file sonoro deve essere disponibile in un supporto di memorizzazione come ad esempio un campionatore, un DAT o un Hard Disk. Il time stretching di una sequenza di campioni audio è mostrata in Fig. 4. Il segnale originale è mostrato nella figura in alto.



**Figura 4:** Time stretching con fattore  $\alpha = 0.5, 2$ .

Il grafico centrale mostra la sequenza ridotta di un fattore  $\alpha = 0.5$  e il grafico più in basso mostra la sequenza secondo un fattore  $\alpha = 2$ .

## Elaborazione del segnale

La contrazione temporale desiderata non corrisponde alla contrazione temporale matematica come è stata realizzata tramite la variazione di velocità. Piuttosto abbiamo bisogno di scalare gli attributi temporali percepiti, come ad esempio la velocità del parlato, senza perturbare gli attributi frequenziali percepiti, come ad esempio l'altezza. Si può dire che vogliamo che si percepisca la versione scalata temporalmente di un segnale acustico con la stessa sequenza degli eventi acustici come se il segnale originale fosse riprodotto secondo un modello temporale scalato. Gli algoritmi di time stretching non dovrebbero interferire con l'altezza o con il contenuto frequenziale dei segnali processati. Questo è dimostrato dagli spettri corrispondenti (primi 2000 campioni) del segnale a tempo discreto nella Fig. 4. Per confronto,

la tecnica tradizionale di time stretching basata sulla riproduzione a velocità variabile introduce un’alterazione dell’altezza (vedi Sezione 2 e Fig. 1). L’idea di base del time stretching tramite l’elaborazione di segmenti temporali è quella di dividere il suono in ingresso in segmenti. Quindi se poi il suono dovrà essere allungato, alcuni segmenti dovranno essere ripetuti; al contrario se il suono deve essere accorciato, alcuni segmenti dovranno essere scartati. Un possibile problema è la discontinuità dell’ampiezza e della fase agli estremi dei segmenti. Le discontinuità di ampiezza saranno evitate da sovrapposizioni parziali dei blocchi, quelle di fasi saranno evitate da un appropriato allineamento dei blocchi. Due strategie differenti saranno presentate nelle Sezioni 3.2 e 3.3.

## Applicazioni

Macchine speciali come il *Phonogène universel* di Pierre Schaeffer o il *Tempophon* utilizzati da Herbert Eimert consentono di modificare la durata così come l’altezza dei suoni. Il *Phonogène* ha trovato molte applicazioni nella *musique concrète* come un “regolatore di tempo”. Nella composizione *Epitaph für Aikichi Kuboyama*, Eimert (1962), utilizza il *Tempophon* per iterare frammenti parlati. Il dispositivo ha permesso la scansione di sillabe, vocali e plosive, per renderle più brevi, più lunghe o iterarle a piacimento (Haller 1995, p. 13).

Come accennato nell’introduzione, l’allungamento dei segnali può essere utilizzato per abbinare la loro durata ad uno slot di tempo riservato. Nella musica Techno, diversi brani musicali vengono riprodotti uno dopo l’altro come un flusso continuo. Si suppone che questo flusso abbia solamente transizioni molto sfumate di tempo o bpm (battiti per minuto), anche se i brani musicali di solito non hanno lo stesso tempo. Per regolare i tempi tra loro, il disc jockey modifica la velocità di riproduzione al passaggio da un brano all’altro. Questo metodo porta a modificazioni di altezza temporanee che potrebbero essere discutibili. L’uso di metodi di time stretching potrebbe eliminare questo problema.

Dopo una breve presentazione della tecnologia del *Phonogène*, le seguenti sezioni discutono due tecniche di elaborazione del segnale che svolgono time stretching senza modificare l’altezza.

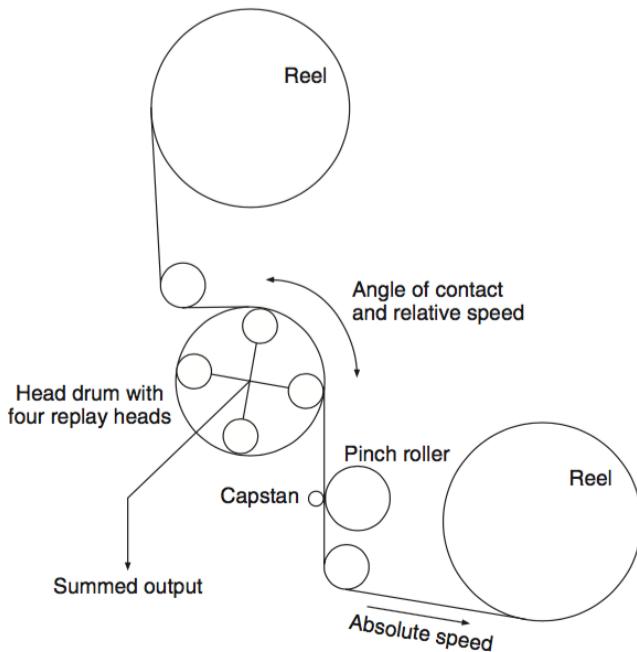
### 3.1 Metodi storici – *Phonogène*

Fairbanks, Everitt e Jaeger parlano già nel 1954 di un registratore a nastro modificato per compressione-espansione temporale o frequenziale del parlato (Laroche 1998; Lee 1972)<sup>3</sup>. Springer sviluppa una macchina simile (Springer 1955; Springer 1959) e Pierre Schaeffer vanta una macchina chiamata *Phonogène universel* che è stata progettata come una combinazione del *Phonogène chromatique* e del *Phonogène à coulisse* di cui abbiamo parlato precedentemente, con il tamburo a testina rotante di Springer (Moles 1961, pp. 71-76), (Schaeffer 1973, pp. 47-48), (Schaeffer e Reibel 1998, CD2, ID. 50-52), (Geslin 2000; Poullin e Sinclair 1957; Poullin 1954).

Il registratore modificato ha diverse testine di riproduzione montate su un tamburo a testina rotante. La velocità assoluta del nastro al capstan determina la durata del suono, mentre la velocità relativa delle testine rispetto a quella del nastro determina l’entità della trasposizione. Elettricamente, con la somma delle uscite delle diverse testine, si ottiene un suono continuo (Fig. 5).

---

<sup>3</sup>Si veda anche, in questo numero, Di Scipio, p. 30. N. d. T.

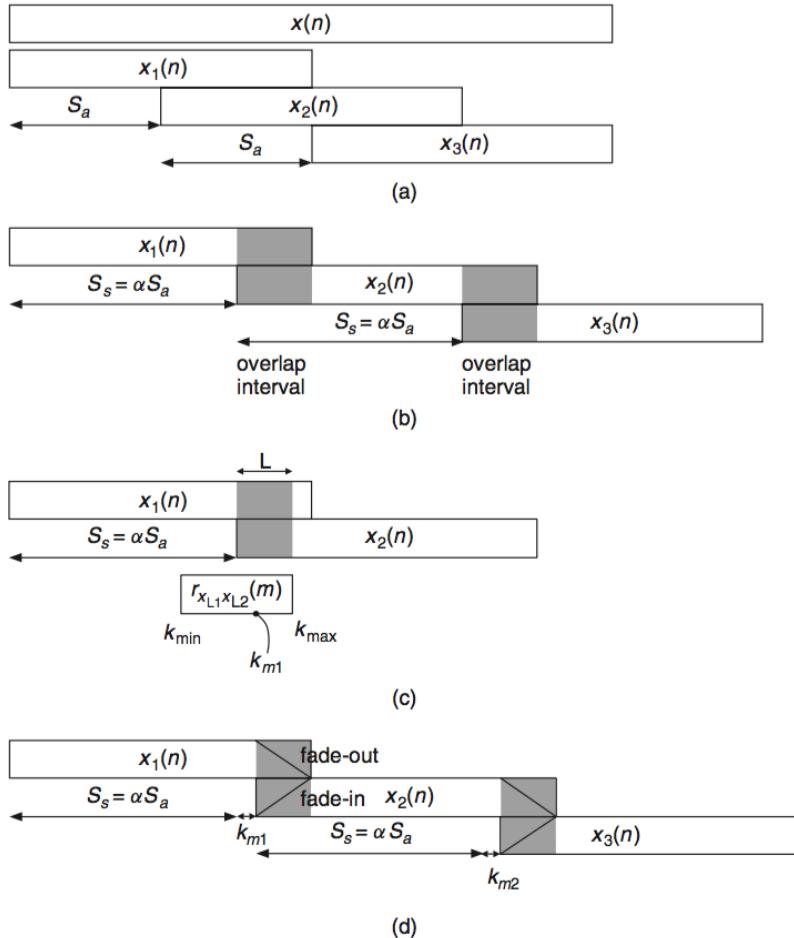


**Figura 5:** Sistema di compressione/espansione basato su magnetofono, vedi Moles (1961).

Moles riporta una tipica gamma operativa da +10% a -40% (Moles 1961, p. 76). La macchina Springer era anche conosciuta col nome di *Laufzregler* o *Zeitdehner* (Cary 1992, pp. 479-480), (Enders 1997).

### 3.2 Sovrapposizione e somma sincrona (SOLA) [Synchronous OverLap and Add]

Un semplice algoritmo per l'allungamento del tempo basato su tecniche di correlazione è proposto in Makhoul e El-Jaroudi (1986) e Roucos e Wilgus (1985). La Fig. 6 illustra l'algoritmo. Il segnale di ingresso  $x(n)$  è diviso in blocchi sovrapposti di lunghezza fissa  $N$  (vedi Fig. 6a).  $S_a$  è la dimensione del passo di analisi che determina la progressione dei blocchi successivi lungo il segnale di ingresso. In un secondo momento questi blocchi sovrapposti vengono spostati in base al fattore di scalatura temporale  $\alpha$  che porta alla dimensione del passo di sintesi  $S_s = \alpha S_a$  (vedi Fig. 6b). Poi le somiglianze in un intervallo di lunghezza  $L$  all'interno della zona di sovrapposizione vengono cercate con uno scarto temporale discreto  $k_m$  di massima somiglianza. Lo spostamento di due blocchi successivi viene corretto da questo scarto temporale per sincronizzare i blocchi e ridurre gli artefatti (vedi Fig. 6c). Infine a questo punto di massima somiglianza i blocchi sovrapposti sono pesati da una funzione fade-in e fade-out e sommati campione per campione (vedi Fig. 6d).



**Figura 6:** (a) segmentazione del segnale in ingresso, (b) riposizionamento dei blocchi, (c) calcolo della correlazione incrociata e sincronizzazione dei blocchi, (d) fading e sovrapposizione ed somma.

Descrizione dell'algoritmo:

1. Segmentazione del segnale di ingresso in blocchi di lunghezza  $N$  con uno slittamento temporale di  $S_a$  campioni.
2. Riposizionamento di blocchi con slittamento temporale  $S_s = \alpha S_a$  con fattore di scala  $\alpha$ .
3. Calcolo della correlazione incrociata

$$r_{x_{L1}x_{L2}} = \begin{cases} \frac{1}{L} \cdot \sum_{n=0}^{L-k-1} x_{L1}(n)x_{L2}(n+k), & 0 \leq k \leq L-1 \\ r_{x_{L2}x_{L1}}(-k), & -L+1 \leq k \leq 0 \end{cases} \quad (3)$$

tra  $X_{L1}(n)$  e  $X_{L2}(n)$ , che sono i segmenti di  $x_1(n)$  e  $x_2(n)$  nell'intervallo di sovrapposizione (overlap interval) di lunghezza  $L$ .

4. Estrazione dello scarto tempo discreto  $k_m$ , dove la correlazione incrociata  $r_{x_{L1}x_{L2}}(k_m) = r_{max}$  ha il suo valore massimo.
5. Fade-out  $x_1(n)$  e fade-in  $x_2(n)$ . Sovrapporre e aggiungere di  $x_1(n)$  e  $x_2(n)$  per il nuovo segnale di uscita.

L'implementazione di SOLA porta alla scalatura temporale con poca complessità, dove i parametri  $S_a$ ,  $N$ ,  $L$  sono indipendenti dal periodo del tono del segnale di ingresso.

Ci sono un certo numero di approcci analoghi all'algoritmo SOLA. Il SOLAFS, come proposto in Hejna e Musicus (1991), realizza la modifica della scalatura temporale con un passo di sintesi fisso, mentre la dimensione del passo di analisi viene variata per ottenere la massima somiglianza nella zona di sovrapposizione tra l'uscita e la nuova finestra. Un approccio simile è proposto in Verhelst e Roelands (1993) sotto forma di un algoritmo di sovrapposizione e somma per la similarità della forma d'onda (WSOLA).

Il listato **M-file 1** realizza l'implementazione dell'algoritmo di scalatura temporale SOLA.

### 3.3 Sovrapposizione e somma sincronizzata con il periodo (PSOLA) [Pitch-Synchronous OverLap and Add]

Una variante dell'algoritmo SOLA per time stretching è la sovrapposizione e somma sincronizzata con il periodo (PSOLA) proposta da Hamon, Moulines e Charpentier (1989) e Moulines e Charpentier (1990) in particolare per l'elaborazione vocale. Si basa sull'ipotesi che il suono in ingresso sia caratterizzato da un'altezza, come, per esempio, la voce umana e gli strumenti musicali monofonici.

In questo caso PSOLA può sfruttare la conoscenza dell'altezza per sincronizzare correttamente i segmenti di tempo, evitando discontinuità d'intonazione. Quando eseguiamo il time stretching di un suono in ingresso, la variazione temporale del periodo di intonazione  $P(t)$  deve essere variata di conseguenza. Se  $\tilde{t} = \alpha t$  descrive la funzione di scalatura temporale o la funzione di time-warping che associa il tempo  $t$  del segnale di ingresso nel tempo  $\tilde{t}$  del segnale di uscita, il periodo locale del segnale di uscita  $\tilde{P}(\tilde{t})$  sarà definito da  $\tilde{P}(\tilde{t}) = \tilde{P}(\alpha t) = P(t)$ .

**M-file 1** (TimeScaleSOLA.m)

---

```
% TimeScaleSOLA.m
% Authors: U. Zölzer, G. De Poli, P. Dutilleux
% Time Scaling with Synchronized Overlap and Add
%
% Parameters:
%
% analysis hop size      Sa = 256 (default parameter)
% block length            N  = 2048 (default parameter)
% time scaling factor    0.25 <= alpha <= 2
% overlap interval        L  = 256*alpha/2

clear all,close all

[DAFx_in,Fs]      = wavread('x1.wav');
DAFx_in           = signal';

% Parameters:
Sa      = 256;      % Sa must be less than N
N       = 2048;
alpha   = 1;         % 0.25 <= alpha <= 2
Ss     = round(Sa*alpha);
L       = 128;       % L must be chosen to be less than N-Ss

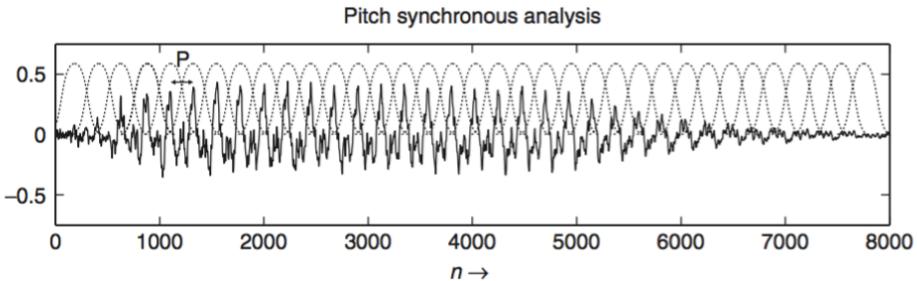
% Segmentation into blocks of length N every Sa samples
% leads to M segments
M       = ceil(length(DAFx_in)/Sa);

DAFx_in(M*Sa+N)=0;
Overlap = DAFx_in(1:N);

% ***** Main TimeScaleSOLA loop *****
for ni=1:M-1
  grain=DAFx_in(ni*Sa+1:N+ni*Sa);
  XCORRsegment=xcorr(grain(1:L),Overlap(1,ni*Ss:ni*Ss+(L-1)));
  [xmax(ni),km(ni)]=max(XCORRsegment);

  fadeout=1:(-1/(length(Overlap)-(ni*Ss-(L-1)+km(ni)-1))):0;
  fadein=0:(1/(length(Overlap)-(ni*Ss-(L-1)+km(ni)-1))):1;
  Tail=Overlap(1,(ni*Ss-(L-1))+...
               km(ni)-1:length(Overlap)).*fadeout;
  Begin=grain(1:length(fadein)).*fadein;
  Add=Tail+Begin;
  Overlap=[Overlap(1,1:ni*Ss-L+km(ni)-1) ...
            Add grain(length(fadein)+1:N)];
end;
% ***** end TimeScaleSOLA loop *****
% Output in WAV file
sound(Overlap,44100);
wavwrite(Overlap,Fs,'x1_time_stretch');
```

---



**Figura 7:** PSOLA: Analisi dell'altezza e finestratura.

Più in generale, quando il fattore di scala  $\alpha$  non è costante, una funzione di scalatura temporale non lineare può essere definita come  $\tilde{t} = \mathcal{T}(t) = \int_0^t \alpha(\tau) d\tau$  e usato al posto di  $\tilde{t} = \alpha t$ .

L'algoritmo si compone di due fasi: la prima fase analizza e segmenta il suono in ingresso (vedi Fig. 7), mentre la seconda fase sintetizza una versione time stretching mediante sovrapposizione e aggiungendo segmenti temporali estratti con l'algoritmo di analisi. Algoritmo di analisi (vedi Fig. 8):

1. Determinazione del periodo di intonazione  $P(t)$  del segnale di ingresso e di istanti (riferimenti di intonazione)  $t_i$ . Questi riferimenti di intonazione (pitch marks) sono definiti in corrispondenza della massima ampiezza o di impulsi della glottide ad un tasso sincronizzato con l'altezza durante la parte periodica del suono e definiti ad un tasso costante durante le porzioni non vocalizzate. In pratica  $P(t)$  è considerato costante  $P(t) = P(t_i) = t_{i+1} - t_i$  sull'intervallo di tempo  $(t_i, t_{i+1})$ .
2. Estrazione di un segmento centrata a ogni riferimento di intonazione  $t_i$  utilizzando una finestra Hanning con lunghezza  $L_i = 2P(t_i)$  (due periodi di riferimento di intonazione) per assicurare fade-in e fade-out.

Algoritmo di sintesi (vedi Fig. 9): per ogni riferimento di intonazione  $\tilde{t}_k$

1. Scelta del segmento di analisi corrispondente  $i$  (identificato dalla marca temporale  $t_i$ ) minimizzando la distanza temporale.
2. Sovrapposizione e somma del segmento selezionato. Si noti che alcuni segmenti di ingresso saranno ripetuti per  $\alpha > 1$  (espansione temporale) o scartati quando  $\alpha < 1$  (compressione temporale).
3. Determinazione dell'istante di tempo  $\tilde{t}_{k+1}$  in cui sarà centrato il successivo segmento di sintesi, al fine di preservare l'altezza locale tramite la relazione

$$\tilde{t}_{k+1} = \tilde{t}_k + \tilde{P}(\tilde{t}_k) = \tilde{t}_k + P(t_i)$$

L'algoritmo di sintesi PSOLA di base può essere implementato in MATLAB® come nel listato **M-file 2**.

I fattori di allungamento tipicamente variano da  $\alpha = 0.25$  a  $2$  per il parlato. Artefatti udibili appaiono in suoni non vocalizzati quando valori più grandi sono applicati, a causa della ripetizione regolare di segmenti di input identici. Per evitare che l'algoritmo introduca

**M-file 2** (psola.m)

---

```

function out=psola(in,m,alpha,beta)
% Authors: G. De Poli, U. Zölzer, P. Dutilleux
%   in      input signal
%   m      pitch marks
%   alpha time stretching factor
%   beta  pitch shifting factor

P = diff(m);    %compute pitch periods

if m(1)<=P(1), %remove first pitch mark
  m=m(2:length(m));
  P=P(2:length(P));
end

if m(length(m))+P(length(P))>length(in) %remove last pitch mark
  m=m(1:length(m)-1);
  else
    P=[P P(length(P))];
end

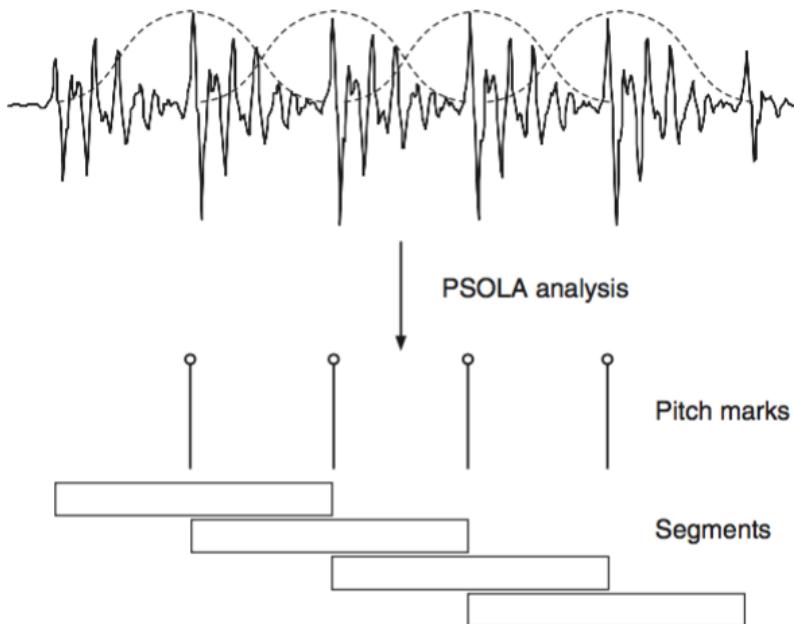
Lout=ceil(length(in)*alpha);
out=zeros(1,Lout); %output signal

tk = P(1)+1;        %output pitch mark

while round(tk)<Lout
  [minimum i] = min( abs(alpha*m - tk) ); %find analysis segment
  pit=P(i);
  gr = in(m(i)-pit:m(i)+pit) .* hanning(2*pit+1);
  iniGr=round(tk)-pit;
  endGr=round(tk)+pit;
  if endGr>Lout, break; end
  out(iniGr:endGr) = out(iniGr:endGr)+gr; %overlap new segment
  tk=tk+pit/beta;
end %while

```

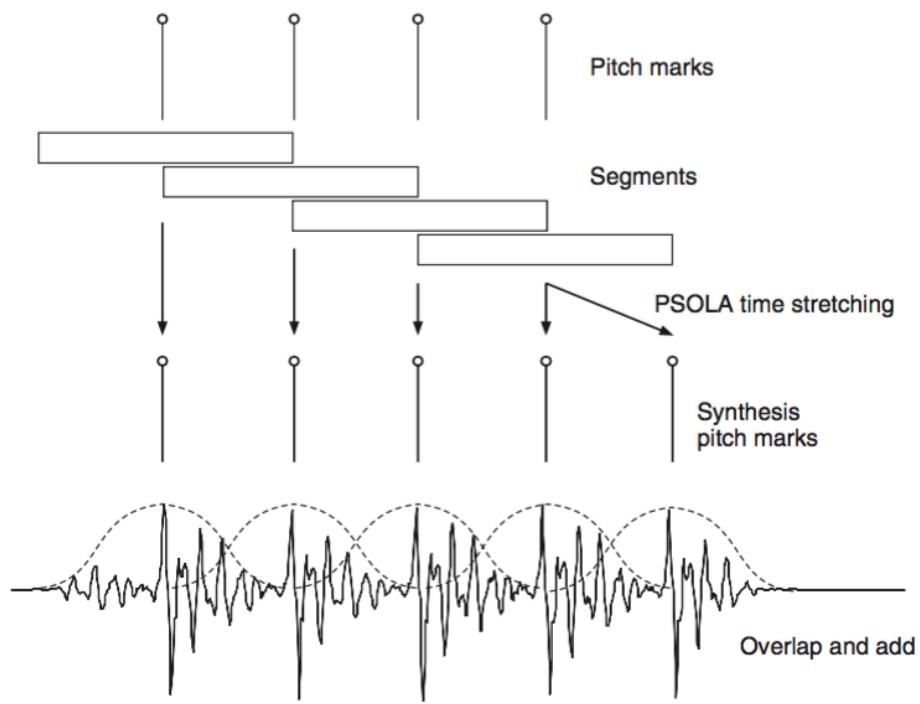
---



**Figura 8:** PSOLA Analisi dell'altezza.

tale correlazione artificiale a breve termine nel segnale di sintesi, è opportuno invertire l'asse temporale di ogni versione ripetuta di un segmento non vocalizzato. Con un tale artificio, il parlato può essere rallentato di un fattore quattro, anche se qualche effetto tonale si incontra in fricative sonore, che uniscono regioni di frequenze vocalizzate e non vocalizzate e che quindi non possono essere invertite nel tempo.

È possibile sfruttare ulteriormente la fase di analisi. Infatti, l'applicazione uniforme del time stretching può produrre qualche artefatto sulle parti non periodiche del suono. Ad esempio una consonante occlusiva può essere ripetuta se l'algoritmo di sintesi seleziona due volte il segmento temporale che contiene la consonante. L'analisi può essere quindi estesa al fine di rilevare la presenza di transizioni veloci (Knesebeck, Ziraksaz e Zölzer 2010; Peeters e Rodet 1999). Durante la sintesi, la scala temporale non verrà modificata in questi punti, quindi non si ripeteranno i segmenti. Questo approccio può essere generalizzato per i suoni non vocalizzati in cui un grande cambiamento nella scala temporale durante le transizioni (ad esempio, gli attacchi) cambierebbe radicalmente l'identità del timbro. Anche in questo caso è possibile limitare l'allungamento temporale durante le transizioni e applicarlo principalmente alla porzione stazionaria del suono in ingresso. Questa tecnica è di solito applicata agli strumenti musicali digitali basati sulla sintesi wavetable. D'altra parte, la deformazione delle parti transitorie può essere considerato una trasformazione timbrica interessante e può essere apprezzato come un effetto audio musicale creativo.



**Figura 9:** PSOLA Sintesi per allungamento temporale.

## Contrassegni di intonazione [Pitch marks]

Va notato che la determinazione del campo e della posizione dei contrassegni di intonazione non è un problema banale e potrebbe essere difficile da implementare in maniera robusta<sup>4</sup>. Una volta che i periodi di intonazione sono noti, la posizione dei contrassegni di intonazione può essere determinata. La qualità del suono risultante dalla modifica mediante l'algoritmo PSOLA dipende essenzialmente dal posizionamento dei contrassegni di intonazione, poiché i contrassegni forniscono i centri delle finestre di segmentazione del PSOLA.

Ovviamente i contrassegni possono essere posizionati solo in corrispondenza al periodo di intonazione per le parti di segnali vocalizzate. Come accennato prima, un modo per gestire le parti non vocalizzate è quello di mantenere il “periodo di intonazione” ad un valore costante fino alla successiva parte vocalizzata. Un metodo che determina i contrassegni di intonazione di una porzione vocalizzata completa è proposto in Lin e Jang (2004) e Mattheyses, Verhelst e Verhoeve (2006). Il picco massimo globale della forma d’onda viene cercato ed il primo contrassegno di intonazione  $t_i$  si trova in questa posizione. Poi i contrassegni di intonazione a destra  $t_i$  sono posizionati utilizzando il periodo di intonazione come una prima stima delle posizioni dei contrassegni di intonazione, seguito da un affinamento della posizione ricercando il massimo entro una determinata regione. Il nuovo contrassegno di intonazione viene determinato come  $t_{i+1} = \max([t_i + \delta P_0, t_i + (2 - \delta)P_0])$ , dove  $P_0$  è il periodo di intonazione e  $\delta$  è un fattore compreso tra 0.5 e 0.9. Questa procedura viene ripetuta fino a raggiungere l'estremità destra della porzione. Poi i contrassegni di intonazione a sinistra di  $t_i$  sono posizionati rispettivamente procedendo con la ricerca del massimo nella regione  $t_{i-1} = \max([t_i - \delta P_0, t_i - (2 - \delta)P_0])$ , fino al raggiungimento dell'estremità sinistra della porzione. Questo metodo è fondamentalmente un approccio di tipo *peak-search* [basato sulla ricerca dei valori di picco].

Una funzione analoga per posizionare i contrassegni di intonazione può essere implementata in una maniera basata sull’analisi di finestre temporali [frame-based] con sovrapposizione di porzioni di finestra di lunghezza fissa. Con la conoscenza del periodo di intonazione del frame corrente e la posizione dell’ultimo contrassegno di intonazione nel frame precedente, il nuovo contrassegno di intonazione può essere prima stimato e successivamente raffinato. Questo rende possibile l’utilizzo in tempo reale dell’approccio *peak-picking* [basato sulla ricerca dei valori di picco in frequenza].

Per i segnali periodici stazionari, questo approccio produce buoni risultati. Ma nei casi in cui il segnale ha un ricco contenuto armonico, il picco massimo può saltare da un periodo all’altro. Questo porta ad un errato posizionamento del contrassegno di intonazione, cioè la distanza di due contrassegni di intonazione successivi non corrisponde al vero periodo di intonazione. Il risultato è la comparsa di artefatti nel segnale risintetizzato. Pertanto, i contrassegni di intonazione devono essere attentamente posizionati e i salti rapidi dei contrassegni di intonazione dovrebbero essere evitati. Poiché è desiderabile posizionare la finestra in modo tale che essa contenga una elevata quantità di energia del segnale, usare un approccio basato sull’individuazione del centro di energia (Knesebeck, Ziraksaz e Zölzer 2010) invece del *peak-picking*, permette un più corretto posizionamento del contrassegno di intonazione. Nei casi in cui il periodo del tono fondamentale all’interno di un frame, l’approccio al centro di energia consente di seguire la variazione dei contrassegni di intonazione in modo continuo, piuttosto che saltare al successivo picco massimo. Ci sono metodi per determinare le posizioni

---

<sup>4</sup>I periodi di intonazione possono essere determinati con uno degli algoritmi di estrazione presentati nel Capitolo 9.2 di Zölzer (2011). N. d. T.

dei contrassegni di intonazione utilizzando le informazioni spettrali. Un metodo basato sulla somma pesata di un gruppo di componenti frequenziali ritardate, che assicura anche che i marcatori siano posizionati vicino al massimo locale di energia, è descritto in Peeters e Rodet (1999).

Un algoritmo di base per il posizionamento dei contrassegni di intonazione è implementato in MATLAB® nel listato **M-file 3**. I contrassegni di intonazione sono posizionati in modo frame-based ad un tasso costante corrispondente al periodo fondamentale di intonazione del frame corrente. In caso di porzioni non sonore viene utilizzato il periodo di intonazione della porzione sonora precedente.

## 4 Pitch shifting [Trasposizione]

### Introduzione

La trasposizione è uno degli strumenti di base dei musicisti. Quando pensiamo di fornire questo effetto mediante elaborazione del segnale, abbiamo bisogno di pensare ai vari aspetti di tale elaborazione. Per un musicista, trasporre significa ripetere una melodia dopo averla traslata di un intervallo fisso. Ogni volta che l'esecutore traspone la melodia, fa uso di un altro registro del suo strumento. In questo modo, non solo l'altezza del suono è modificata, ma anche il timbro viene influenzato.

Nel regno dei DAFX, è una questione di scelta trasporre senza tener conto della modifica del timbro o se il timbro caratteristico dello strumento deve essere mantenuto in ciascuno dei suoi registri. Il primo metodo potrebbe essere chiamato “trasposizione a timbro variabile”, mentre il secondo approccio sarebbe chiamato “trasposizione a timbro costante”. Per ottenere una panoramica del problema dobbiamo prendere in considerazione le origini fisiche del segnale audio.

Il timbro di un suono dipende fortemente dall'organizzazione del suo spettro. Un modello può essere derivato dallo studio della voce cantata. L'altezza di una voce cantata è determinata dalle corde vocali e può essere correlata con la serie di frequenze disponibili nello spettro. Il timbro della voce è determinato principalmente dal tratto vocale. L'effetto di quest'ultimo è quello di evidenziare alcune parti dello spettro, che sono chiamate formanti. Un modello del segnale può essere definito osservando dove una parte di eccitazione è modificata da una parte di risonanza. Nel caso della voce, l'eccitazione è fornita dalle corde vocali, quindi è legata alle frequenze dello spettro, mentre le risonanze corrispondono ai formanti. Quando un cantante traspone una melodia ha, in una certa misura, la possibilità di modificare altezza e formanti indipendentemente. In un'implementazione di elaborazione del segnale attenta a questo effetto, entrambi questi aspetti devono essere considerati.

Se solo lo spettro dell'eccitazione è espanso o contratto, si ottiene una trasposizione d'altezza in alto o in basso con un timbro costante. Se sono espanso o contratte solo le risonanze, l'altezza rimane la stessa, ma il timbro varia. Il canto armonico si basa su questo effetto. Se sia l'eccitazione che la risonanza vengono deliberatamente e in maniera indipendente alterate, allora entriamo nel campo degli effetti che possono essere percepiti come innaturali, ma che potrebbero avere un grande potenziale musicale.

La separazione di un suono nella sua parte di eccitazione e di risonanza è un processo complesso<sup>5</sup>. Presenteremo qui metodi che alterano contemporaneamente entrambi gli aspetti,

---

<sup>5</sup>Detto processo è affrontato nel Capitolo 8 di Zölzer (*ibid.*). N. d. T.

**M-file 3 (findpitchmarks.m)**


---

```

function [m] = findPitchMarks(x,Fs,F0,hop,frameLen)
% Author: A. von dem Knezebeck
% x      input signal
% Fs     sampling frequency
% F0    fundamental frequencies
% hop   hop size of F0 detection
% frameLen length of frame

% Initialization
m       = 0;      % vector of pitch mark positions
P0      = zeros (1,length(F0));
index   = 1;
local_m = [];% local pitch marker position

% processing frames i
for i = 1:length(F0);
% set pitch periods of unvoiced frames
if (i==1 & F0(i)==0); F0(i) = 120;    % 120Hz in case no preceding pitch
elseif (F0(i)==0);      F0(i) = F0(i-1);
end
P0(i) = round(Fs/F0(i));           % fundamental period of frame i
frameRange = (1:frameLen) + (i-1)*hop; % hopping frame
last_m = index;                   % last found pitch mark

% beginning periods of 1st frame
j = 1;  %period number
if i==1
    % define limits of searchFrame
    searchUpLim = 1 * P0(i);
    searchRange = (1 : searchUpLim);
    [pk,loc] = max(x(searchRange));
    local_m(j) = round(loc);

% beginning periods of 2nd - end frame
else
    searchUpLim = searchUpLim + P0(i);
    local_m(j) = last_m + P0(i);
end % beginning periods of 1st - end frame

% remaining periods of 1st - end frame
index = local_m(1);
j = 2; % grain/period number
while( searchUpLim + P0(i) <= frameRange(end))
    % define range in which a marker is to be found
    searchUpLim = searchUpLim + P0(i);
    local_m(j) = local_m(j-1) + P0(i);
    index = local_m(j);
    j = j+1;
end %while frame end not reached
m = [m local_m];
end % processing frames i

% finishing calculated pitch marks
m = sort(m);
m = unique(m);
m = m(2:end);

```

---

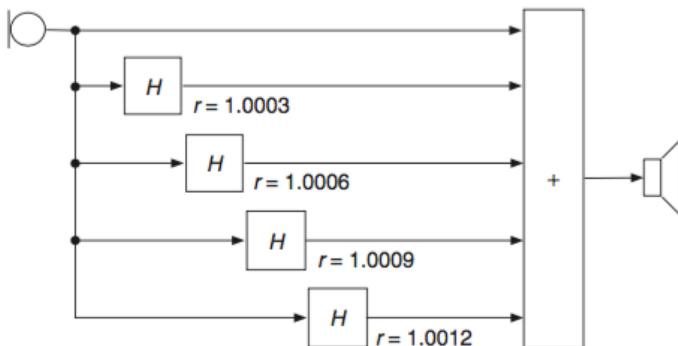
come l'harmonizer o Pitch Shifting tramite la modulazione della linea di ritardo nella Sezione 4.3. Un metodo più raffinato basato su PSOLA, che consente la correzione dell'altezza con la conservazione delle formanti, verrà discusso nella Sezione 4.<sup>6</sup>

## Applicazioni musicali

Tipiche applicazioni di pitch shifting nella musica pop sono la correzione dell'altezza degli strumenti o cantanti, così come la produzione di un effetto simile a un coro. Quando la voce di un cantante è mixata con copie di se stesso leggermente trasposte, appare un sottile effetto che dà l'impressione che si stia ascoltando un coro, invece di un solo cantante.

Anche l'armonizzatore può produrre effetti sorprendenti, come un uomo che parla con una piccola voce stridente o una donna con un grintoso tono basso. Possono essere prodotte sonorità estreme come il profondo suono del rullante in *Let's Dance* di David Bowie (White 1999). È stato anche utilizzato per la scomposizione e ricomposizione del parlato (Giordano, Rothman e Hollien 1973). In combinazione con una linea di ritardo con feedback del suono trasposto in ingresso, può essere prodotta una sorta di spirale in cui il suono viene sempre recepito più alto o più basso ad ogni iterazione.

Un effetto sottile, simile al phasing, può essere realizzato con una serie di harmonizers (Dutilleux 1988) accoppiati in parallelo e mixati al suono in ingresso, come mostrato nella Fig. 10. Il rapporto di trasposizione dell' $n^{th}$  harmonizer potrebbe essere impostato su  $1 + nr$ , dove  $r$  è dell'ordine di  $1/3000$ . Se  $f_0$  è l'altezza del suono, le uscite dell' $n^{th}$  harmonizer



**Figura 10:** Un set di armonizzatori che produce un effetto tipo phasing. È particolarmente efficace per toni bassi (tipicamente  $100\text{ Hz}$ ) di lunga durata.

forniranno un'intonazione di  $f_0 + n\Delta f$ , dove  $\Delta f = rf_0$ . Se  $f$  è abbastanza piccolo (pochi centesimi di  $\text{Hz}$ ) le interferenze tra le varie uscite degli harmonizer saranno chiaramente udibili. Quando viene applicato, ad esempio, ad un suono grave di tuba, sarà enfatizzato un armonico dopo l'altro. Effetti di flanging e di chorus possono essere ottenuti anche impostando il controllo di intonazione per una piccola quantità di trasposizione (ad esempio,  $1/10$  a  $1/5$  di un semitono) e l'aggiunta di rigenerazione (Anderton 1995, p. 53). Qui ci sembra che l'altezza di un effetto audio è molto dipendente dal suono elaborato. Succede spesso che l'altezza deve essere regolata per ogni nuovo suono o ogni nuova altezza.

<sup>6</sup>Per i metodi più avanzati di pitch-shifting si rimanda ai Capitoli 7-11 di Zölzer (2011). N. d. T.

Hans Peter Haller descrive in *Das Experimental Studio der Heinrich-Strobel-Stiftung des Südwestfunks Freiburg 1971- 1989*, pp. 51-55, alcune applicazioni dell’armonizzatore per la produzione di opere musicali di Luigi Nono e André Richard.

#### **4.1 Metodi Storici – *Harmonizer***

Le macchine basate su nastro descritte in 3.1 sono state anche in grado di modificare l’altezza dei suoni pur mantenendo la loro durata iniziale. Il *Phonogène universel* era ingombrante e non riusciva a trovare una vasta diffusione, ma nel bel mezzo degli anni settanta del Novecento, apparse un dispositivo digitale che fu chiamato *Harmonizer*. Questo implementava nel dominio digitale un processo simile a quello del *Phonogène universel*. Da qui l’effetto è diventato molto popolare. Poiché *Harmonizer* è un marchio della società Eventide, altre aziende offrono dispositivi simili con nomi come *pitch transposer* o *pitch shifter*.

La principale limitazione nell’uso dell’armonizzatore è la qualità caratteristica che dà ai suoni elaborati. Moles afferma che il raggio di azione del *Phonogène universel*, utilizzato come regolatore d’intonazione, funzionava da -4 a +3 semitonni (Moles 1961, p. 74). Geslin stima che le macchine disponibili alla fine degli anni sessanta hanno trovato applicazione nella musica concreta con rapporti di trasposizione molto più grandi (Geslin 2000).

Le implementazioni digitali in forma di harmonizer potrebbero consentire una migliore qualità, ma ci sono ancora importanti limitazioni. Per trasposizioni dell’ordine di un semitono, quasi nessuna alterazione dei suoni può essere udita. Appena il rapporto di trasposizione diventa più grande, in pratica nella gamma di più o meno due ottave, il timbro del suono in uscita ottiene un carattere che è specifico dell’armonizzatore.

Questa modifica può essere ascoltata sia nel dominio della frequenza sia nel dominio del tempo ed è dovuta alla modulazione del segnale da parte della finestra di analisi. Lo spettro del segnale di ingresso è infatti convoluto con quello della finestra. La modulazione nel dominio del tempo può essere caratterizzata dalla sua velocità e dallo spettro della finestra, che dipende dalla sua forma e dalle sue dimensioni. Più grande è la finestra, minore è il tasso e quindi la risoluzione dello spettro della finestra e meno disturba la modulazione. L’effetto di una finestra trapezoidale sarà più forte di quello di una più dolce, ad esempio la finestra coseno rialzato.

D’altra parte, una finestra più grande tende a fornire, attraverso il processo di sovrapposizione e somma, udibili copie iterate dei segnali di ingresso. Per la trasposizione di suoni percussivi è necessario ridurre la dimensione della finestra. Inoltre, per riprodurre in modo accurato transienti e non renderli troppo dolci, la finestra dovrebbe avere transizioni nette. Vediamo che un bilanciamento tra la modulazione udibile dello spettro sonoro e transitori iterati deve essere valutato per ogni tipo di suono. I musicisti che usano il computer come uno strumento musicale potrebbero sfruttare queste peculiarità dell’algoritmo per dare al loro *sound* un sapore unico.

#### **4.2 Pitch Shifting mediante dilatazione temporale e ricampionamento**

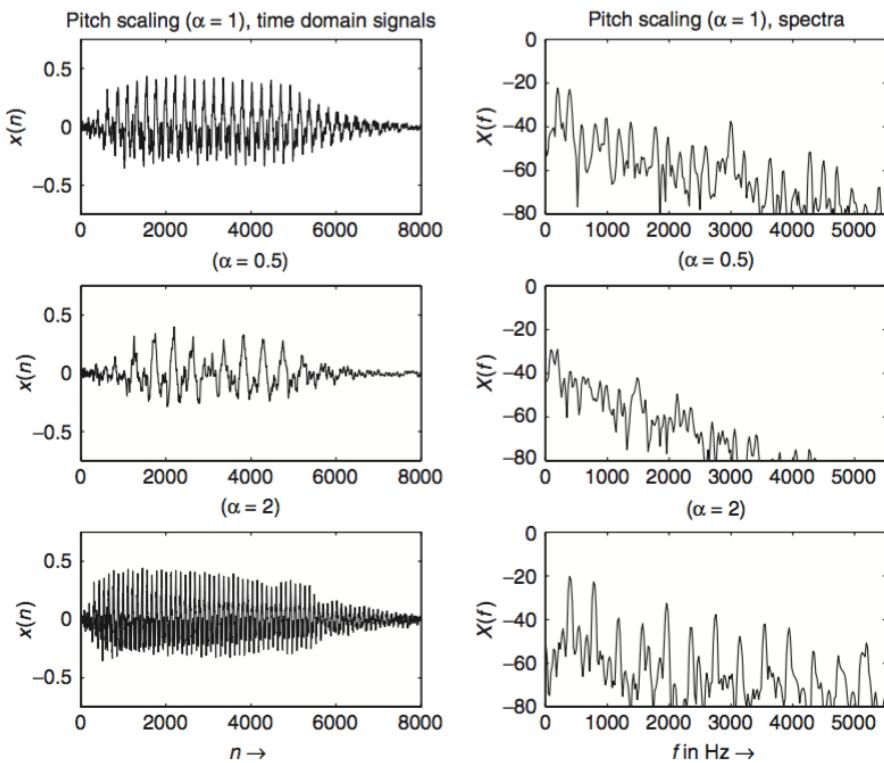
La riproduzione a velocità variabile discussa nella Sezione 2 porta ad una compressione o espansione della durata di un suono e ad un cambio di intonazione. Questo si ottiene attraverso il ricampionamento nel dominio del tempo. La Fig. 1 illustra i segnali a tempo discreto e gli spettri corrispondenti. Lo spettro del suono viene compresso o espanso nell’asse frequenziale. Le relazioni armoniche

$$f_i = i \cdot f_{\text{fondamentale}} \quad (4)$$

del suono non sono modificate, ma scalate secondo

$$f_i^{\text{new}} = \alpha \cdot f_i^{\text{old}} \quad (5)$$

Le ampiezze delle armoniche rimangono le stesse  $a_i^{\text{new}} = a_i^{\text{old}}$ . Al fine di riscalare il suono trasposto alla lunghezza originale, un ulteriore algoritmo di time-stretching può essere applicato al suono. Il risultato di trasposizione seguito da un algoritmo di time-stretching è illustrato nella Fig. 11.



**Figura 11:** Pitch shifting seguito da correzione temporale.

L'ordine di cambio di intonazione e di scalatura temporale può essere modificato, come mostrato nella Fig. 12. Innanzitutto, un algoritmo di scalatura temporale espande il segnale di ingresso da lunghezza  $N_1$  a lunghezza  $N_2$ . Poi un'operazione di ricampionamento con rapporto inverso  $N_1/N_2$  effettua il cambio di intonazione e una riduzione della lunghezza  $N_2$  alla lunghezza  $N_1$ . Il listato **M-file 4** realizza l'implementazione di scalatura temporale con algoritmo SOLA e algoritmo di cambio di intonazione.

**M-file 4 (PitchScaleSOLA.m)**


---

```
% PitchScaleSOLA.m
% Authors: G. De Poli, U. Zölzer, P. Dutilleux
% Parameters:
%   analysis hop size      Sa = 256 (default parameter)
%   block length           N  = 2048 (default parameter)
%   pitch scaling factor  0.25 <= alpha <= 2
%   overlap interval       L  = 256*alpha/2
clear all,close all
[signal,Fs]      = wavread('x1.wav');
DAFx_in          = signal';

Sa=256;N=2048;           % time scaling parameters
M=ceil(length(DAFX_in)/Sa);

n1=512;n2=256;           % pitch scaling n1/n2
Ss=round(Sa*n1/n2);
L=256*(n1/n2)/2;

DAFx_in(M*Sa+N)=0;
Overlap=DAFx_in(1:N);

% ***** Time Stretching with alpha=n2/n1*****
..... % include main loop TimeScaleSOLA.m
% ***** End Time Stretching *****

% ***** Pitch shifting with alpha=n1/n2 *****
lfen=2048;lfen2=lfen/2;
w1=hanningz(lfen);w2=w1;

% for linear interpolation of a grain of length lx to length lfen
lx=floor(lfen*n1/n2);
x=1+(0:lfen-1)''lx/lfen;
ix=floor(x);ix1=ix+1;
dx=x-ix;dx1=1-dx;
%
lmax=max(lfen,lx);
Overlap=Overlap';
DAFx_out=zeros(length(DAFX_in),1);

pin=0;pout=0;
pend=length(Overlap)-lmax;
% Pitch shifting by resampling a grain of length lx to length lfen
while pin<pend
    grain2=(Overlap(pin+ix).*dx1+Overlap(pin+ix1).*dx).* w1;
    DAFX_out(pout+1:pout+lfen)=DAFx_out(pout+1:pout+lfen)+grain2;
    pin=pin+n1;pout=pout+n2;
end;
```

---

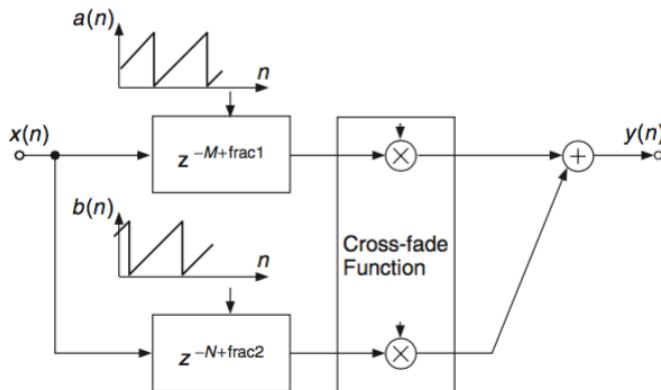


**Figura 12:** Pitch shifting mediante scalatura temporale e ricampionamento.

### 4.3 Cambio di altezza mediante modulazione della linea di ritardo

Il cambio di altezza o trasposizione basata su elaborazione di segmenti è descritto in diverse pubblicazioni. In Bogdanowicz e Blecher (1989) è proposto un cambio di altezza basato su uno schema di sovrapposizione e somma con due linee di ritardo variabili nel tempo (vedi Fig. 13). Un blocco cross-fade combina le uscite di due linee di ritardo mediante una funzione di dissolvenza incrociata.

Il segnale viene diviso in piccoli pezzi. I pezzi vengono letti più velocemente per produrre intonazioni più alte o più lentamente per produrre intonazioni più basse. Al fine di produrre un segnale continuo di uscita, due pezzi vengono letti simultaneamente con un ritardo pari alla metà della lunghezza del blocco. È quindi costruita una dissolvenza incrociata da un pezzo all'altro fino a ciascuna estremità di ciascun pezzo (Warstat e Görne 1994, pp. 257-259).

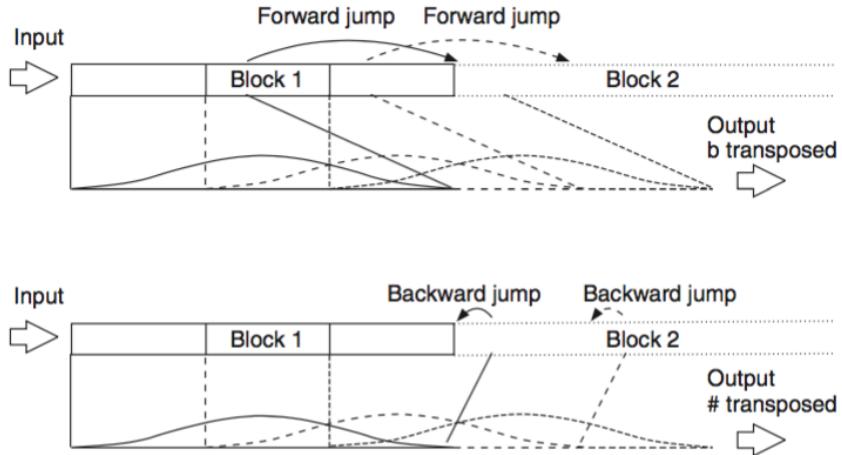


**Figura 13:** Pitch shifting.

La lunghezza delle linee di ritardo viene modulata da una funzione a dente di sega. Un approccio simile è proposto in Dattoro (1987), in cui viene utilizzata la stessa configurazione per compressione del tempo ed espansione. Un algoritmo di rilevamento della periodicità viene utilizzato per calcolare la funzione di dissolvenza incrociata per evitare cancellazioni durante le dissolvenze.

Un metodo avanzato per la trasposizione dei segnali audio è presentato in Disch e Zölzer (1999). Il metodo si basa su uno schema di sovrapposizione e somma e non necessita di alcuna stima della frequenza fondamentale. La differenza da altre applicazioni è il modo in cui i blocchi sono modulati e combinati per il segnale di uscita. Il sistema di trasposizione potenziato

si basa su uno schema di sovrapposizione e somma con tre linee di ritardo tempo-varianti parallele (vedi Fig. 15).

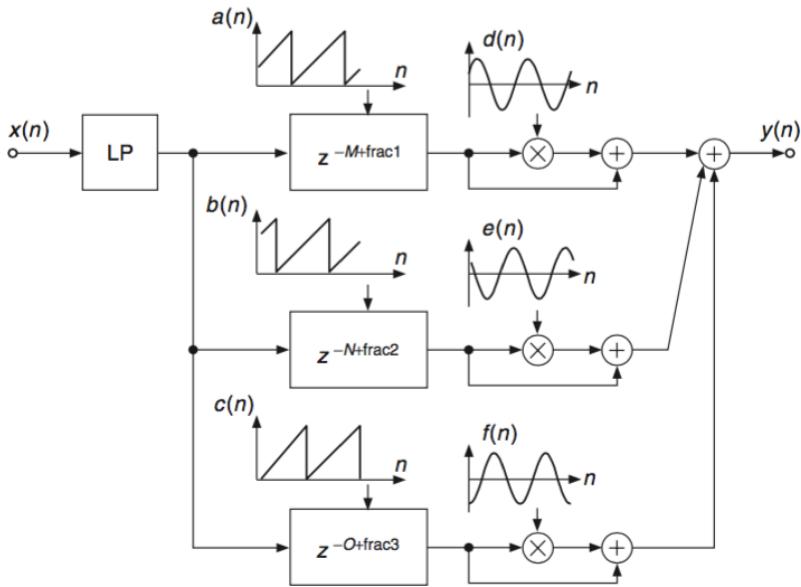


**Figura 14:** Traspositore avanzato: elaborazione dei blocchi, dilatazione temporale, sovrapposizione ed somma.

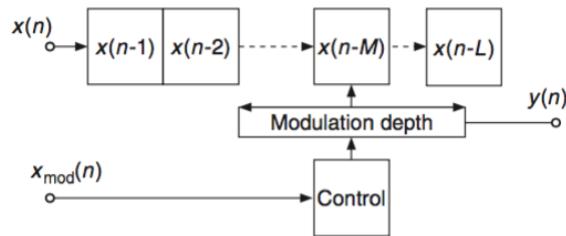
La Fig. 14 illustra come il segnale di ingresso è diviso in blocchi, che sono ricampionati (modulazione di fase con un segnale a rampa), modulati in ampiezza e sommati ottenendo un segnale di uscita della stessa lunghezza del segnale di ingresso. Blocchi adiacenti si sovrappongono con 2/3 della lunghezza del blocco.

I segnali di modulazione formano un sistema di tre funzioni coseno rialzato sfasati di 120°. La somma di queste funzioni è costante per tutti gli argomenti. La Fig. 15 mostra anche la topologia del traspositore di altezza. Dal momento che un coseno completo viene utilizzato per la modulazione, la qualità sonora percepita del segnale elaborato è molto migliore di quella prodotta utilizzando quelle applicazioni con una semplice duplice sovrapposizione e somma che utilizzano diverse finestre. La modulazione di ampiezza produce solo frequenze somma e differenza con la frequenza base del segnale di modulazione, che può essere molto bassa (6-10 Hz). Le armoniche non sono presenti nel segnale di modulazione e quindi non possono generare frequenze somma o differenza di ordine superiore. Gli artefatti percepiti sono simili all'effetto di phasinge sono meno fastidiosi rispetto alle discontinuità locali di altre applicazioni basate su metodi di sovrapposizione e somma duplice.

Se vogliamo cambiare l'altezza di un segnale controllato da un altro segnale o da un inviluppo di segnale possiamo anche fare uso della modulazione mediante linea di ritardo. L'effetto può essere ottenuto eseguendo una modulazione di fase del segnale registrato secondo  $y(n) = x(n - D(n))$ . Il fattore modulante  $D(n) = M + PROFONDIT \cdot X_{mod}(n)$  è ora dipendente da un segnale modulante  $X_{mod}(n)$ . Con questo approccio l'altezza del segnale di ingresso  $x(n)$  viene cambiata secondo l'inviluppo del segnale modulante (vedi Fig. 16).



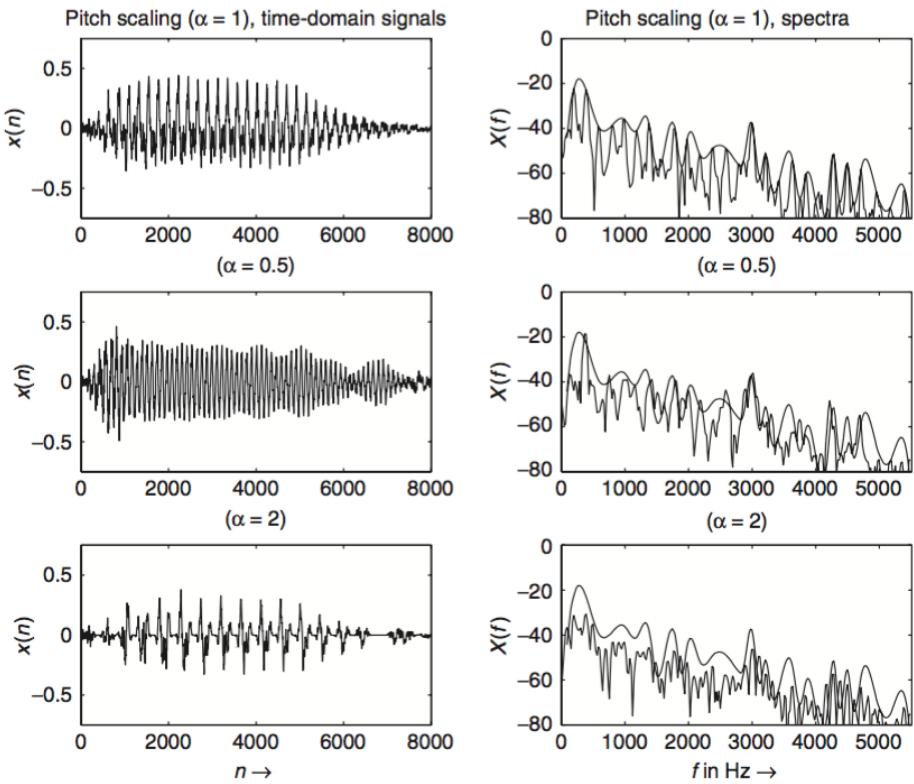
**Figura 15:** Traspositore avanzato: diagramma a blocchi.



**Figura 16:** Altezza controllata dall'inviluppo del segnale  $x_{mod}(n)$ .

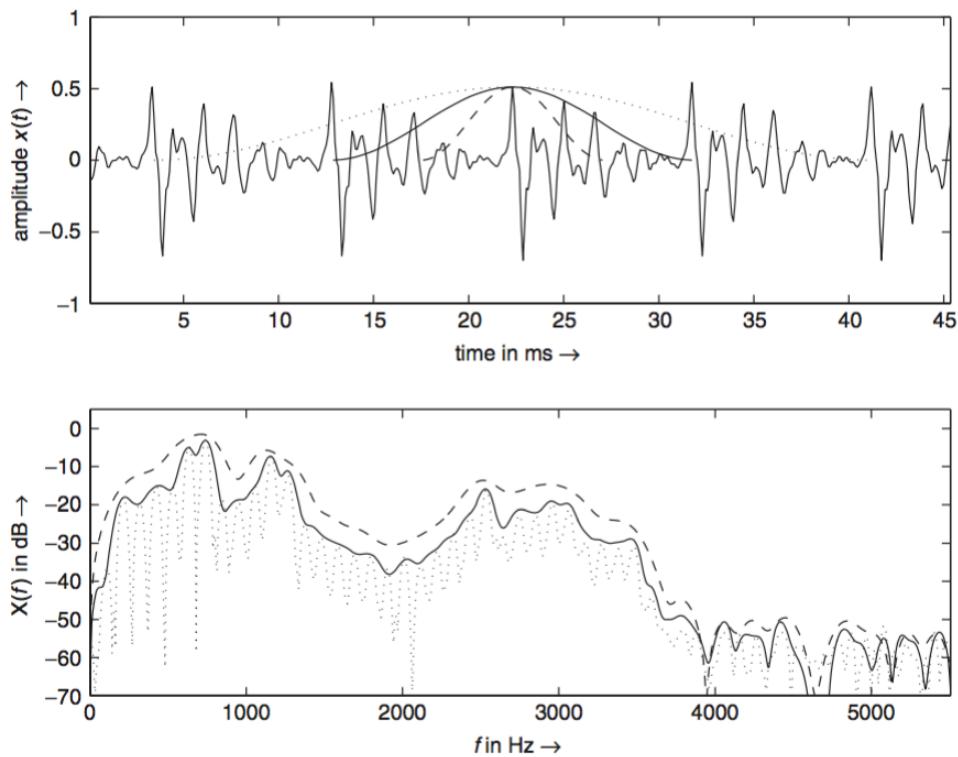
#### 4.4 Cambio di intonazione mediante PSOLA e conservazione delle formanti

Questa tecnica è l'operazione duale del ricampionamento nel dominio del tempo ma in questo caso è eseguito un ricampionamento dell'inviluppo spettrale ottenuto dall'analisi di Fourier a tempo breve. L'inviluppo spettrale a tempo breve descrive una curva in frequenza passando attraverso tutte le ampiezze delle armoniche. Questo è dimostrato nella Fig. 17, dove si vede l'inviluppo spettrale. Le armoniche sono ancora scalate in base alla nuova  $f_i^{new} = \beta \cdot f_i^{old}$ , ma le ampiezze delle armoniche  $a_i^{new} = env(f_i^{new}) \neq a_i^{old}$  sono ora determinate campionando l'inviluppo spettrale. Si può notare qualche differenza di ampiezza rispetto all'inviluppo originale. Questo dipende dalla scelta dell'algoritmo di cambio di altezza.



**Figura 17:** Pitch shifting con il metodo PSOLA: ricampionamento delle frequenze dell'inviluppo spettrale.

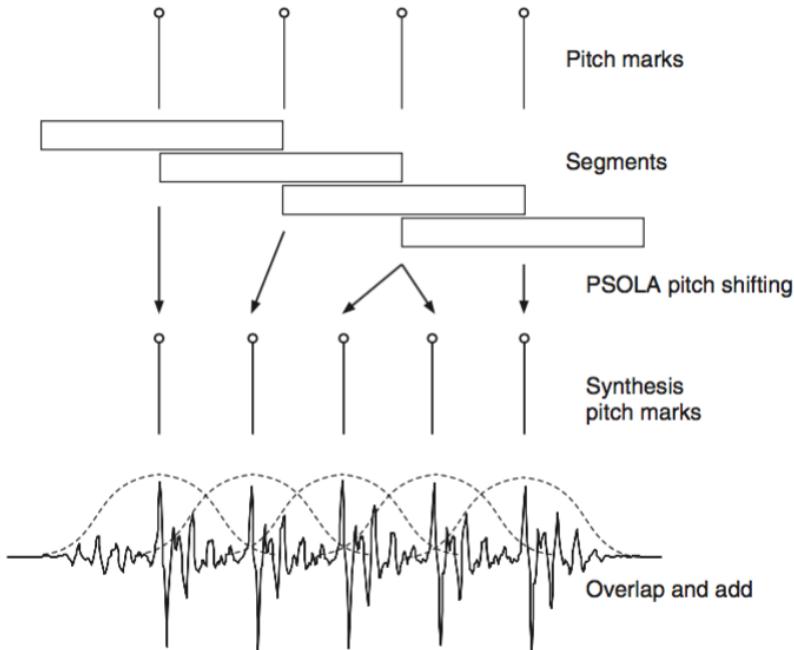
L'algoritmo PSOLA può essere convenientemente utilizzato per il cambio di intonazione di un suono vocale mantenendo la posizione delle formanti, e così l'identità vocale (Bristow-Johnson 1995; Moulines e Laroche 1995). L'idea di base consiste nello stiramento temporale della posizione dei marcatori di intonazione, mentre il segmento della forma d'onda non viene modificato. Il modello del segnale sottostante la produzione del parlato è un treno di impulsi



**Figura 18:** Spettro di segmenti estratti da una vocale /a/ usando una finestra di Hanning lunga rispettivamente 4 (linea punteggiata), 2 (linea continua) e 1 (linea tratteggiata) periodi di intonazione. Si può notare che la linea continua approssima l'inviluppo spettrale locale.

filtrato attraverso un filtro variabile nel tempo corrispondente al tratto vocale. Il segmento di ingresso corrisponde alla risposta impulsiva del filtro e determina la posizione delle formanti. Pertanto, tale risposta, non dovrebbe essere modificata. Al contrario, la distanza tra i marcatori di intonazione determina il periodo del parlato e, quindi, tale distanza deve essere modificata di conseguenza. Lo scopo dell'analisi PSOLA è quello di estrarre la risposta impulsiva locale del filtro. Come si può vedere nella Fig. 18, lo spettro di un segmento estratto utilizzando una finestra di Hanning con una lunghezza di due periodi approssima l'inviluppo spettrale locale. Finestre più lunghe tendono ad evidenziare la struttura fine delle righe dello spettro, mentre finestre più corte tendono ad offuscare la struttura delle formanti dello spettro. Così se non estendiamo il segmento, la posizione delle formanti viene mantenuta. L'operazione di sovrapposizione dei segmenti alla posizione del nuovo marcatore di intonazione ricampionerà l'inviluppo spettrale alla frequenza di intonazione desiderata. Quando desideriamo un cambio di intonazione di un fattore  $\beta$ , definito come il rapporto tra la frequenza del tono di sintesi locale rispetto a quello originale  $\beta = \tilde{f}_0(\tilde{t})/f_0(t)$ , il nuovo periodo di intonazione sarà dato da  $\tilde{P}(\tilde{t}) = P(t)/\beta$ , dove in questo caso  $\tilde{t} = t$  perché il tempo non è allungato.

L'algoritmo di analisi è lo stesso di quello visto in precedenza per allungamento temporale PSOLA in Sezione 3.3 (vedi Fig. 8). L'algoritmo di sintesi viene modificato (vedi Fig. 19) secondo le seguenti fasi:



**Figura 19:** PSOLA: algoritmo di sintesi per cambio di altezza.

- Per ogni marcatore di intonazione nella sintesi  $\tilde{t}_k$ :

1. Scelta del corrispondente segmento di analisi  $i$  (identificato dalla marca temporale  $t_i$ ) minimizzando la distanza temporale  $|t_i - \tilde{t}_k|$ .
2. Sovrapporre e aggiungere il segmento selezionato. Si noti che alcuni segmenti di ingresso saranno ripetuti per  $\beta > 1$  (altezza più alta) o scartati quando  $\beta < 1$  (altezza più bassa).
3. Determinazione dell'istante di tempo  $\tilde{t}_{k+1}$  dove il prossimo segmento di sintesi sarà centrato, in modo da preservare l'altezza locale, con la relazione

$$\tilde{t}_{k+1} = \tilde{t}_k + \tilde{P}(\tilde{t}_k) = \tilde{t}_k + P(t_i/\beta)$$

- Per grandi cambiamenti di altezza è consigliabile, per compensare la variazione di ampiezza introdotta dalla maggiore o minore sovrapposizione di segmenti, moltiplicare il segnale di uscita per  $1/\beta$ .

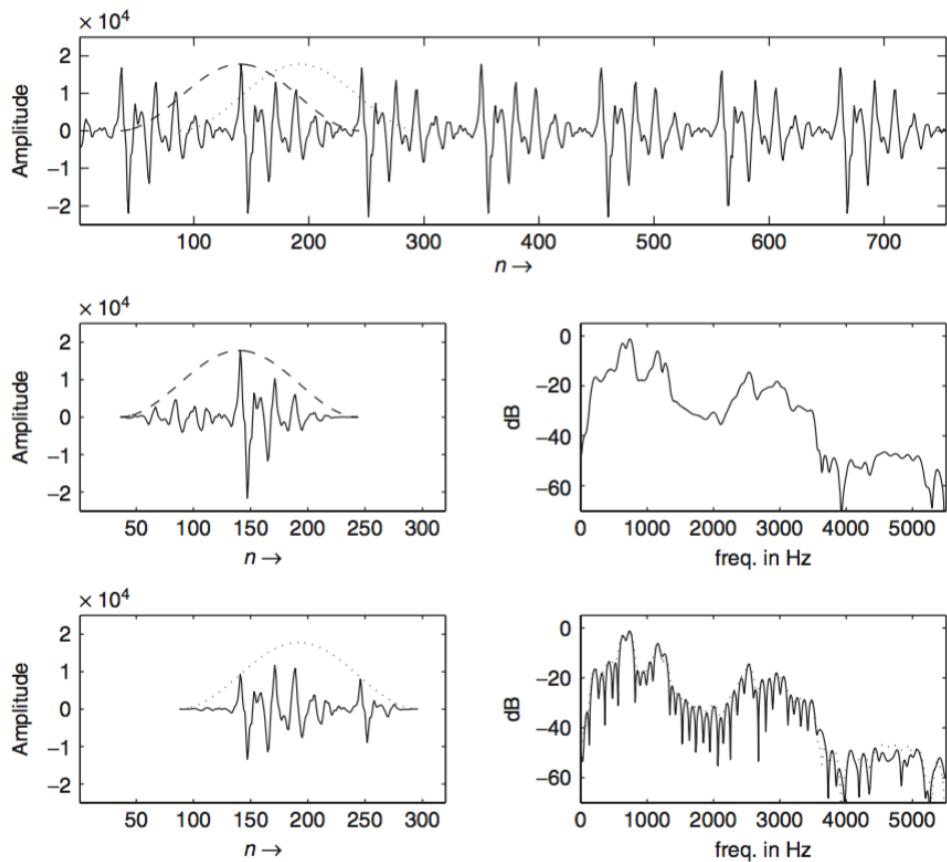
È possibile combinare l'allungamento temporale di un fattore  $\alpha$  con il cambio di altezza. In questo caso, per ogni marcatore di intonazione sintetizzato  $\tilde{t}_k$  il primo passo dell'algoritmo di sintesi presentato sopra sarà modificato dalla scelta del segmento corrispondente di analisi  $i$  (identificato dalla marcatura temporale  $t_i$ ), minimizzando la distanza temporale  $|ct_i - \tilde{t}_k|$ .

L'algoritmo PSOLA è molto efficace per la elaborazione della voce ed è computazionalmente molto efficiente una volta che il suono sia stato analizzato. Quindi è ampiamente usato per la sintesi vocale da un database di difoni, per la modifica della prosodia, per macchine automatiche telefoniche, ecc. Per un'ampia variazione dell'altezza presenta alcuni artefatti. D'altra parte, la necessità di una fase preliminare di analisi per ottenere un profilo di intonazione rende difficoltosa l'implementazione in tempo reale di una modifica del segnale di ingresso. Anche la stima di impulsi della glottide può essere difficile. Una soluzione consiste nel posizionare i marcatori di intonazione ad un tasso sincrono con il periodo, indipendentemente dalla posizione reale degli impulsi glottali. La qualità di sintesi risultante sarà solo leggermente diminuita (si veda, ad esempio, la Fig. 20).

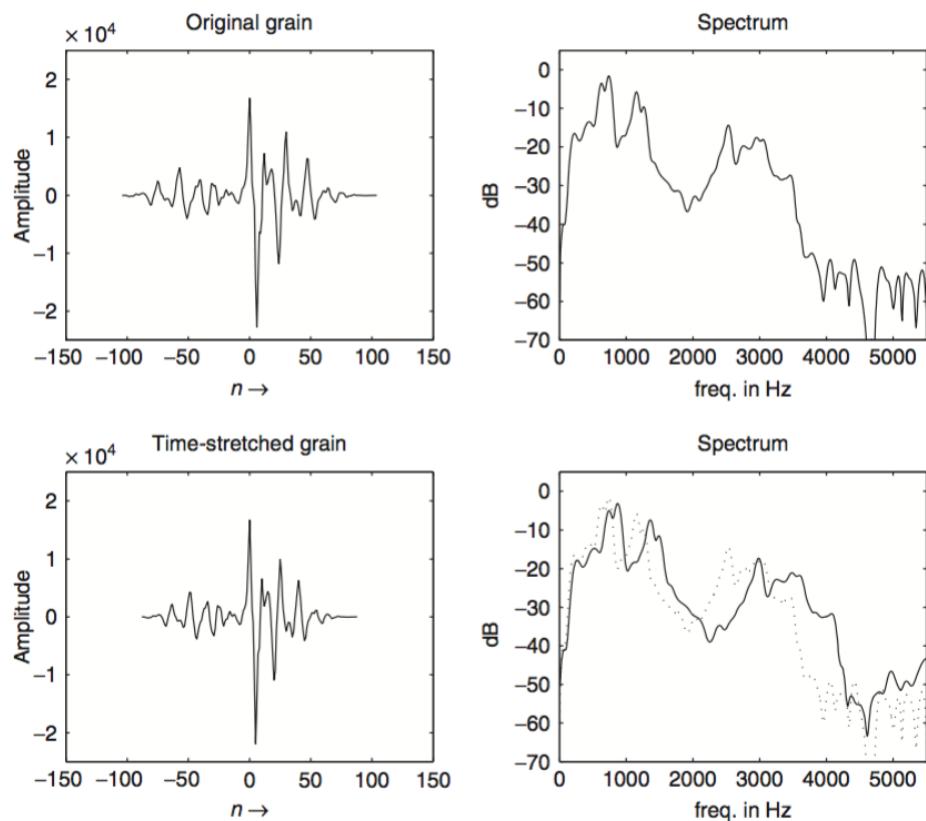
Un ulteriore effetto che può essere ottenuto da una variazione di PSOLA è una scalatura lineare delle frequenze formanti (vedi Fig. 21). Infatti, abbiamo visto che la scala temporale di un segnale corrisponde ad una scala di frequenza inversa. Così quando eseguiamo una scalatura temporale della risposta all'impulso di un filtro, si scala inversamente la frequenza delle formanti. In termini PSOLA, questo corrisponde alla scalatura temporale dei segmenti di ingresso selezionati prima del passo di sintesi di sovrapposizione e somma, senza alcuna modifica nel calcolo dei marcatori di intonazione. Per aumentare le frequenze dei formanti di un fattore  $\gamma$ , ogni segmento dovrebbe essere ridotto di un fattore  $1/\gamma$  attraverso il ricampionamento. Ad esempio, le frequenze medie delle formanti di femmine adulte sono circa il 16% superiori a quelle dei maschi adulti, e le formanti dei bambini sono circa il 20% superiori rispetto alle formanti femminili. Si deve fare attenzione quando le frequenze aumentano in modo da evitare foldover. Idealmente dovrebbe essere usato un campionamento a banda limitata.

Il cambio di altezza mediante PSOLA può essere utilizzato per sintetizzare più voci da un vero cantante per creare un effetto di *coro virtuale* (Schnell et al. 2002).

Il listato **M-file 5** implementa dell'algoritmo base di sintesi PSOLA. Si basa sull'algoritmo di stiramento temporale PSOLA mostrato nella Sezione 3.3.



**Figura 20:** Confronto di un segmento estratto in corrispondenza di un impulso glottale con uno estratto tra impulsi glottali



**Figura 21:** PSOLA: variazione di PSOLA per scalatura lineare delle formanti.

**M-file 5** (psolaf.m)

---

```

function out=psolaF(in,m,alpha,beta,gamma)
% Authors: G. De Poli, U. Zölzer, P. Dutilleux
%
%     . .
%     gamma newFormantFreq/oldFormantFreq
%     . .
%     the internal loop as
tk = P(1)+1;                      %output pitch mark
while round(tk)<Lout
    [minimum i]=min(abs(alpha*m-tk) );    % find analysis segment
    pit=P(i);pitStr=floor(pit/gamma);
    gr=in(m(i)-pit:m(i)+pit).*hanning(2*pit+1);
    gr=interp1(-pit:1:pit,gr,-pitStr*gamma:gamma:pit);% stretch segm.
    iniGr=round(tk)-pitStr;endGr=round(tk)+pitStr;
    if endGr>Lout, break; end
    out(iniGr:endGr)=out(iniGr:endGr)+gr; % overlap new segment
    tk=tk+pit/beta;
end % end of while

```

---

## 5 Ricombinazione temporale e granulazione

### 5.1 Ricombinazione temporale

#### Introduzione

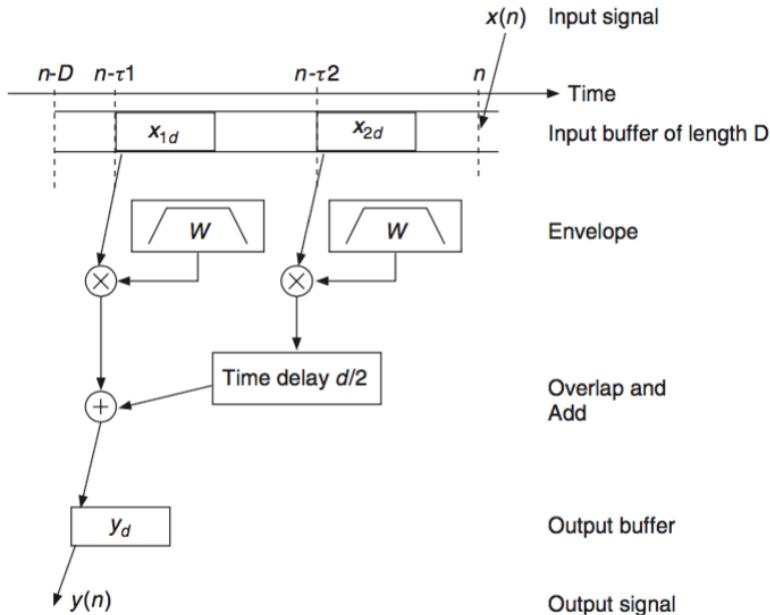
La *musique concrète* ha fatto un uso intensivo di ricombinazione di piccoli elementi di nastro magnetico. Quando ben controllato, questo assemblaggio di centinaia di frammenti di diverse decine di millisecondi permette una fusione di materiali sonori eterogenei, al limite della soglia di discriminazione temporale. Questa operazione manuale, chiamata micro-montaggio, è molto laboriosa. Bernard Parmegiani suggerì nel 1980 al Groupe de Recherches Musicales (GRM) che questa operazione può essere svolta dal computer. Una prima versione del software è stata prodotta nei primi anni ottanta. Dopo essere riscritta, migliorata e ricompilata più volte, alla fine si è reso disponibile su personal computer sotto forma di un programma chiamato *brassage* in francese che verrà tradotto qui come *ricombinazione temporale* (Geslin 1998, 2000).

#### Elaborazione del segnale

Descriviamo qui un algoritmo elementare per la ricombinazione temporale che si basa sulla sovrapposizione di due segmenti di tempo che vengono raccolti in modo casuale dal segnale di ingresso (vedi Fig. 22):

1. Siano  $x(n)$  e  $y(n)$  i segnali di ingresso e di uscita.
2. Si assegni la durata  $d$  dei frammenti e la durata  $D \geq d$  del periodo di tempo  $[n - D, n]$  da cui verranno selezionati i segmenti di tempo.
3. Memorizzare il segnale in ingresso  $x(n)$  in una linea di ritardo di lunghezza  $D$ .

4. Scegliere a caso il tempo di ritardo  $\tau_1$  con  $d \leq \tau_1 \leq D$ .
5. Selezionare il segmento del segnale  $x_{1d}$  di durata  $d$  che inizia a  $x(n - \tau_1)$ .
6. Seguire la stessa procedura (punti 4 e 5) per un secondo segmento  $x_{2d}$ .
7. Leggere  $x_{1d}$  e  $x_{2d}$  e applicare un inviluppo di ampiezza  $W$  per ciascuno di loro, al fine di appianare le discontinuità ai bordi.
8. Quando la lettura di  $x_{1d}$  e  $x_{2d}$  è terminata, iterare la procedura per ciascuno di essi.
9. Calcolare l'uscita come sovrapposizione e somma delle sequenze  $x_{1d}$  e  $x_{2d}$  con un ritardo temporale di  $d/2$ .



**Figura 22:** Ricombinazione temporale: due segmenti di ingresso, selezionati casualmente dal segnale in ingresso trascorso, sono sovrapposti e aggiunti per produrre un segmento temporale di uscita. Quando un segmento di ingresso è terminato, viene selezionato uno nuovo.

### Applicazioni musicali e controllo

La versione sopra descritta introduce disturbi locali nella evoluzione temporale del segnale, preservando la continuità globale della sua sequenza temporale. Molti ulteriori perfezionamenti di questo algoritmo sono però possibili. Un coefficiente di ampiezza casuale potrebbe essere applicato a ciascuno dei segmenti di ingresso per modificare la densità del materiale sonoro. La forma dell'inviluppo potrebbe essere modificata al fine di mantenere la struttura temporale

d'ingresso o, dall'altra parte, per smussare e fondere diversi eventi tra loro. La velocità di riproduzione dei segmenti potrebbe essere variata per produrre trasposizione o glissando.

In un'epoca in cui gli strumenti informatici non erano ancora disponibili, Bernard Parmegiani ha magnificamente illustrato la tecnica di micro-tessitura basata su nastro in opere come *Violostries* (1964) o *Dedans-Dehors* (1977) (Parmegiani 1964, 1977). L'algoritmo elementare presentato sopra può essere utilizzato in tempo reale, ma sono state implementate altre versioni che offrono numerose caratteristiche ulteriori. Esse hanno la possibilità di unire i frammenti, campionati casualmente e di qualsiasi dimensione, da alcuni campioni a diversi minuti. Così, oltre a generare fenomeni di fusione, per cui l'algoritmo è stato concepito, il software è stato in grado di produrre dissolvenza incrociata di suoni tessiturali ed anche accordi sostenuti, con infinitamente piccole variazioni nella stabilità del segnale, e nell'interpolazione di frammenti con silenzio o suoni di altri tipi (Geslin 1998). Jean-Claude Risset ha usato questo effetto per eseguire evoluzioni sonore di suoni brevi, come pietre e campanelli di metallo (Risset 1987, *Sud-I*, 3'44"-4'38"); (Geslin 2000; Risset 1998) e per la produzione di un pianoforte "balbuziente", ulteriormente elaborato da modulazione ad anello (Risset 1987, *Sud-I*, 4'30", 5'45"). A partire da "oggetti trovati" come il canto degli uccelli, Jean-Claude Risset li ha riorganizzati in modo compositivo per ottenere prima un rendering puntillistico, poi un episodio simil-stretto (ibid., 1'42"-2'49").

## 5.2 Granulazione

### Introduzione

Nelle sezioni precedenti sulla modulazione di altezza e sulla dilatazione temporale, abbiamo proposto algoritmi che presentano limitazioni per quanto riguarda il loro scopo iniziale. Al di là di una gamma limitata di modifica di altezza o di durata, appaiono evidenti artefatti. Il metodo della ricombinazione temporale considera questi artefatti da un punto di vista artistico e li dà per scontati. Tra le possibilità offerte dai metodi e dai loro limiti, questi propongono di creare nuove strutture sonore. Considerando che l'effetto di ricombinazione temporale sfrutta le possibilità di uno specifico sistema software, che potrebbe essere considerato qui come "strumento musicale", l'idea di costruire un suono complesso da un grande insieme di suoni elementari potrebbe essere un orizzonte più ampio.

Il fisico Dennis Gabor ha proposto nel 1947 l'idea del *quantum sonoro*, un'unità indivisibile di informazione dal punto di vista della psicoacustica. Secondo la sua teoria, una rappresentazione granulare potrebbe descrivere qualsiasi suono. La sintesi granulare è stata suggerita come una tecnica della computer music per la produzione di suoni complessi da Iannis Xenakis (1971) e Curtis Roads (1978). Questa tecnica accumula eventi acustici provenienti da migliaia di grani sonori. Un grano sonoro è molto breve (tipicamente da 1 a 100 ms) e si avvicina alla durata, alla frequenza e all'ampiezza dell'evento minimo percepibile (Roads 1996, 1998; Truax 2000a).

L'effetto di granulazione è un'applicazione di sintesi granulare in cui il materiale di cui sono costituiti i grani è un segnale di ingresso. Barry Truax per primo ha sviluppato questa tecnica (Truax 1988a,b) implementandola in tempo reale e utilizzandola ampiamente nelle sue composizioni.

## Elaborazione del segnale

Siano  $x(n)$  e  $y(n)$  i segnali di ingresso e di uscita. I grani  $g_k(i)$  sono estratti dal segnale di ingresso con l'aiuto di una funzione finestra  $w_k(i)$  di lunghezza  $L_k$  tramite

$$g_k(i) = x(i + i_k)w_k(i) \quad (6)$$

con  $i = 0, \dots, L_{k-1}$ . L'istante di tempo  $i_k$  indica il punto in cui il segmento viene estratto; la lunghezza  $L_k$  determina la quantità di segnale estratto; la finestra  $w_k(i)$  dovrebbe garantire fade-in e fade-out ai bordi del grano e accentuare il contenuto in frequenza del grano. Grani lunghi tendono a mantenere l'identità timbrica della porzione del segnale di ingresso, mentre quelli corti acquisiscono una qualità simile all'impulso. Quando il grano è lungo, la finestra ha una cima piatta e viene utilizzata solo per fade-in e fade-out ai bordi del segmento.

I listati **M-file 6** e **7** mostrano l'estrazione di grani corti e lunghi.

**M-file 6** (grainSh.m)

---

```
function y = grainSh(x,init,L)
% Authors: G. De Poli
% extract a short grain
% x      input signal
% init   first sample
% L      grain length (in samples)
y=x(init:init+L-1).*hanning(L)';
```

---

**M-file 7** (grainLn.m)

---

```
function y = grainLn(x,iniz,L,Lw)
% Authors: G. De Poli
% extract a long grain
% x      input signal
% init   first sample
% L      grain length (in samples)
% Lw    length fade-in and fade-out (in samples)
if length(x) <= iniz+L , error('length(x) too short.'), end
y = x(iniz:iniz+L-1);                                % extract segment
w = hanning(2*Lw+1)';
y(1:Lw)      = y(1:Lw).*w(1:Lw);                  % fade-in
y(L-Lw+1:L) = y(L-Lw+1:L).*w(Lw+2:2*Lw+1); % fade-out
```

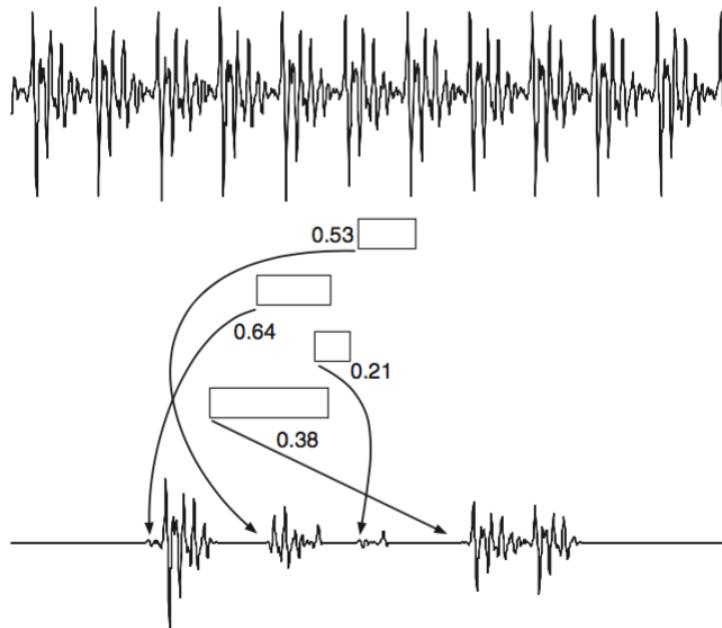
---

La formula di sintesi è data da

$$y(n) = \sum_k a_k g_k(n - n_k) \quad (7)$$

dove  $a_k$  è un eventuale coefficiente ampiezza e  $n_k$  è l'istante di tempo in cui il grano viene inserito nel segnale di uscita. Si noti che i grani possono sovrapporsi. Per sovrapporre un grano  $g_k$  (grano) all'istante  $n_k = (\text{iniOLA})$  con  $a_k$  ampiezza, possono essere utilizzate le seguenti istruzioni MATLAB<sup>®</sup>

```
endOLA = iniOLA + lungh(greno) - 1;
y(iniOLA: endOLA) = y (iniOLA:endOLA) + ak * greno;
```



**Figura 23:** Esempio di granulazione

Un esempio di granulazione con valori casuali dei parametri di punto di inizio del grano e lunghezza, punto di uscita e l'ampiezza è mostrato nella Fig. 23. Il file **M-file 8** mostra l'implementazione dell'algoritmo di granulazione.

Questa tecnica è molto generale e può essere impiegata per ottenere differenti effetti sonori. Il risultato è fortemente influenzato dal criterio utilizzato per scegliere gli istanti  $n_k$ . Se questi punti sono regolarmente distanziati nel tempo e la forma d'onda del grano non cambia troppo, la tecnica può essere interpretata come un treno di impulsi filtrati, cioè produce un suono periodico il cui inviluppo spettrale è determinato dalla forma d'onda del grano interpretato come una risposta all'impulso. Un esempio è l'algoritmo PSOLA mostrato nelle precedenti Sezioni 3.3 e 4.4. Quando la distanza tra due grani successivi è molto maggiore rispetto a  $L_k$ , il suono si tradurrà in grani separati da interruzioni o silenzi, con un carattere specifico. Quando molti grani brevi si sovrappongono (cioè la distanza è inferiore a  $L_k$ ) si ottiene un effetto di tessitura sonora.

Le strategie per la scelta degli istanti di sintesi possono essere raggruppate in due categorie semplificate: strategie sincrone, principalmente basate su funzioni deterministiche, e asincrone,

**M-file 8** (granulation.m)

---

```
% granulation.m
% Authors: G. De Poli
f=fopen('a_male.m11');
x=fread(f,'int16');
fclose(f);

Ly=length(x);  y=zeros(1,Ly);                      %output signal
% Constants
nEv=4; maxL=200;  minL=50;  Lw=20;
% Initializations
L = round((maxL-minL)*rand(1,nEv))+minL;    %grain length
initIn = ceil((Ly-maxL)*rand(1,nEv));          %init grain
initOut= ceil((Ly-maxL)*rand(1,nEv));          %init out grain
a = rand(1,nEv);                                %ampl. grain
endOut=initOut+L-1;
% Synthesis
for k=1:nEv,
    grain=grainLn(x,initIn(k),L(k),Lw);
    y(initOut(k):endOut(k))=y(initOut(k):endOut(k))+a(k)*grain;
end
```

---

basate su funzioni stocastiche. I grani possono essere organizzati in flussi. Ci sono due variabili di controllo principali: il ritardo tra i grani per un singolo flusso, e il grado di sincronicità tra i grani in flussi differenti. Dato che lo spettro locale incide sulla struttura globale del suono, è possibile utilizzare il suono in ingresso che può essere analizzato in grani senza alterare le caratteristiche complessive del suono originale, come gocce d'acqua per suoni tipo corso d'acqua.

È inoltre possibile modificare la forma d'onda del grano con una trasformazione temporale, come modulazione per spostamento di frequenza o dilatazione temporale che si estende per scalatura frequenziale (De Poli e Piccialli 1991). I principali parametri di granulazione sono: durata del grano, ordine di selezione dei grani dal suono in ingresso, ampiezza dei grani, pattern temporale per la sintesi e la densità del grano (cioè grani per secondo). La densità è un parametro primario, in quanto determina la tessitura globale, se frammentaria o continua. Si noti che è possibile estrarre grani da differenti file sonori per creare tessiture ibride, per esempio, che evolvono da una tessitura all'altra.

### Applicazioni Musicali

Esempi dell'effetto [di granulazione] possono essere ascoltati in Wishart (1994). Barry Truax ha utilizzato la tecnica della granulazione per elaborare il suono campionato come materiale compositivo. In *The Wings of Nike* (1987) egli ha elaborato solo brevi frammenti “fonemici”, ma sequenze più lunghe di suoni ambientali sono state utilizzate in brani come *Pacific* (1990). In ciascuno di questi lavori, il materiale granulato viene dilatato temporalmente da varie quantità e produce così una serie di modifiche percettive che sembrano provenire da dentro il suono (Truax 1995, 2000b).

In *Le Tombeau de Maurice*, Ludger Brümmer utilizza la tecnica della granulazione per ottenere modifiche timbriche, ritmiche e armoniche (Brümmer 1997). Una transizione dalla colorazione del suono originale di un campione orchestrale verso impulsi di rumore si ottiene riducendo progressivamente la dimensione dei grani. Con dimensione dei grani intermedia, l'altezza del suono originale è ancora riconoscibile, sebbene la struttura temporale sia scomparsa (ibid., 3'39"-4'12"). Una melodia può essere suonata selezionando grani ad altezze diverse e variando il ritmo con cui i grani vengono riprodotti (ibid., 8'38"-9'10"). Nuove melodie possono anche apparire come uscita da una granulazione a due stadi. Una prima serie di grani è definita dal campione originale, mentre la seconda è una granulazione della prima. A causa della separazione operata dall'apparato uditivo, la ritmica ed il raggruppamento armonico dei grani sono in continua evoluzione (ibid., 9'30"-10'33").

## 6 Conclusioni

Gli effetti descritti in questo capitolo sono basati sulla divisione del suono in ingresso in brevi segmenti. Questi segmenti sono trattati con metodi semplici quali la dilatazione temporale tramite ricampionamento, o moltiplicazione dell'ampiezza con un inviluppo. La forma d'onda del segmento non viene modificata, mantenendo così la caratteristica del segnale sorgente.

Si possono ottenere due categorie di effetti, a seconda della strategia utilizzata per posizionare i segmenti di tempo durante la sintesi. Se l'ordine e l'organizzazione dei segmenti estratti vengono accuratamente mantenuti, possono essere ottenuti dilatazione temporale o modifica dell'altezza. Metodi di base, SOLA e PSOLA, sono presentati e sono discusse le loro caratteristiche. Questi effetti hanno lo scopo di produrre suoni che vengono percepiti come simili all'originale, ma vengono modificati nella durata o nell'altezza. Come spesso accade con effetti audio digitali, i suoni elaborati prodotti da questi metodi possono essere utilizzati come metodo per deformazione del suono in ingresso, pur mantenendo le sue caratteristiche principali. La bassa complessità computazionale di elaborazione di un segmento temporale consente applicazioni in tempo reale efficienti. Tuttavia, questi algoritmi producono artefatti che limitano l'ambito di applicazione.

La seconda categoria di effetti cambia l'organizzazione e l'ordine dei segmenti in una larga area, e quindi porta alla ricombinazione temporale ed alla granulazione. In questo caso, il suono in ingresso può essere molto meno riconoscibile nell'output. L'elemento centrale diventa il grano con il suo inviluppo di ampiezza ed organizzazione temporale. Queste tecniche possono produrre risultati da grani sparsi una tessitura densa, con un rapporto molto libero con il suono originale. Si noti che l'ampia scelta di strategie per l'organizzazione dei grani implica un atteggiamento alla composizione sonora da parte dell'utente. Così la granulazione è diventata una sorta di metafora per la composizione musicale partendo dal micro livello.

## Suoni e musica

- Brümmer, Ludger (1993). *The Gates of H*. CD. Computer music. 1995. In: CRI, *The Listening Room*, edel 0014522TLR. Hamburg: CCRMA.  
 — (1997). *Le Tombeau de Maurice, for computer-generated tape*. CD. In: *Computer Music at CCRMA*. CCRMAV02. CCRMA.

- Eimert, Herbert (1962). *Epitaph für Aikichi Kuboyama*. Electronic music composition. Studio-Reihe neuer Musik, Wergo, LP WER 60014. Riedizione su CD, Koch/Schwann, 1996.
- Furukawa, K. (1993). *Swim, Swan, composition for clarinet and live-electronics*. CD. ZKM.
- (1997). *Den ungeborenen Göttern*. CD. Multimedia-Opera. ZKM.
- Parmegiani, Bernard (1964). *Violostries*. CD. IDEAMA 051 Target Collection, 1996. ZKM e CCRMA.
- (1977). *Dedans-Dehors*. INA-GRM.
- Risset, Jean-Claude (1987). *Risset*. C 1003. CD. Contiene: *Sud* (tracce 1-3), *Dialogues* (traccia 4), *Inharmonique* (traccia 5), *Mutations* (traccia 6). Paris: INA-GRM.
- Schaeffer, Pierre e Guy Reibel (1998). *Solfège de l'object sonore*. Con 3 CD-DA. Paris: INA-GRM.
- Truax, Barry (1995). *Granular time-shifting and transposition composition examples*. CD. In: *Computer Music Journal*, vol. 19, index 6.
- Wishart, Trevor (1994). *Audible Design*. Sound examples. CD. York.

## Bibliografia

- Anderton, C. (1995). *Multieffects for Musicians*. Amsco Publications.
- Bogdanowicz, K. e R. Blecher (1989). «Using multiple processors for real-time audio effects». In: *Proc. AES 7th Int. Conf.* Pp. 337–342.
- Bristow-Johnson, R. (1995). «A detailed analysis of a time-domain formant-corrected pitch shifting algorithm». In: *J. Audio Eng. Soc* 43.5, pp. 340–352.
- Cary, Tristram (1992). *Illustrated Compendium of Musical Technology*. Faber e Faber.
- Chion, Michel (1982). *La Musique Electroacoustique*. QSJ No 1990. PUF.
- Crochiere, R. E. e L. R. Rabiner (1983). *Multirate Digital Signal Processing*. Prentice-Hall.
- Dattoro, J. (1987). «Using digital signal processor chips in a stereo audio time compressor/expander». In: *Proc. 83rd AES Convention, Preprint 2500*.
- De Poli, Giovanni e Aldo Piccialli (1991). «Pitch-synchronous granular synthesis». In: *Representations of musical signals*. A cura di Giovanni De Poli, Aldo Piccialli e Curtis Roads. Cambridge, Mass.: MIT Press, pp. 391–412.
- Disch, S. e U. Zölzer (1999). «Modulation and delay line based digital audio effects». In: *Proc. DAFX-99 Digital Audio Effects Workshop*, pp. 5–8.
- Dutilleux, Pierre (1988). *Mise en œuvre de transformations sonores sur un système temps-reel*. Technical report, Rapport de stage de DEA. CNRS-LMA.
- Dutilleux, Pierre et al. (2011). «Time-segment processing». In: *Digital Audio Effects*. A cura di Udo Zölzer. 2<sup>a</sup> ed. John Wiley & Sons, Ltd. Cap. 6, pp. 185–217.
- Enders, B. (1997). *Lexikon Musikelektronik*. Atlantis Schott.
- Gaskell, P. S. (1987). «A hybrid approach to the variable speed replay of digital audio». In: *J. Audio Eng. Soc.* 35, pp. 230–238.
- Geslin, Y. (1998). «Sound and music transformation environments: A twenty-year experiment at the *Groupe de Recherches Musicales*». In: *Proc. DAFX-98 Digital Audio Effects Workshop*, pp. 241–248.
- (2000). «About the various types of *Phonogènes*». Personal communication.
- Giordano, T. A., H. B. Rothman e H. Hollien (1973). «Helium speech unscramblers - a critical review of the state of the art». In: *IEEE Trans. Audio and Electroacoust.* Vol. AU-21. 5.

- Haller, Hans Peter (1995). *Das Experimental Studio der Heinrich-Strobel-Stiftung des Südwestfunks Freiburg 1971- 1989. Die Erforschung der Elektronischen Klangumformung und ihre Geschichte*. Nomos.
- Hamon, C., E. Moulines e F. Charpentier (1989). «A diphone synthesis system based on time-domain prosodic modifications of speech». In: *Proc. ICASSP*, pp. 238–241.
- Hejna, Don e Bruce R. Musicus (1991). *The SOLAFS time-scale modification algorithm*. Technical report, BBN.
- Knesebeck, A. von dem, P. Ziraksaz e U. Zölzer (2010). «High quality time-domain pitch shifting using PSOLA and transient preservation». In: *Proc. of the 129th AES Convention*.
- Laroche, J. (1998). «Time and pitch scale modifications of audio signals». In: *Applications of Digital Signal Processing to Audio and Acoustics*. A cura di M. Kahrs e K.-H. Brandenburg. Kluwer, pp. 279–309.
- Lee, F. F. (1972). «Time compression and expansion of speech by the sampling method». In: *J. Audio Eng. Soc.* 20.9, pp. 738–742.
- Lin, C. Y. e J. S. R. Jang (2004). «A two-phase pitch marking method for TD-PSOLA synthesis». In: *8th Int. Conf. Spoken Lang. Process.*
- Makhoul, J. e A. El-Jaroudi (1986). «Time-scale modification in medium to low rate speech coding». In: *Proc. ICASSP*, pp. 1705–1708.
- Massie, D. C. (1998). «Wavetable sampling synthesis». In: *Applications of Digital Signal Processing to Audio and Acoustics*. A cura di M. Kahrs e K.-H. Brandenburg. Kluwer, pp. 311–341.
- Mattheyses, W., W. Verhelst e P. Verhoeve (2006). «Robust pitch marking for prosodic modification of speech using TD-PSOLA». In: *Proc. IEEE Benelux/DSP Valley Signal Process. Symp.*
- McNally, G. W. (1984). «Variable speed replay of digital audio with constant output sampling rate». In: *Proc. 76th AES Convention, Preprint 2137*.
- Moles, Abraham (1961). *Les Musiques expérimentales*. Trad. da D. Charles. Zurigo: Cercle d'art contemporain.
- Moulines, E. e F. Charpentier (1990). «Pitch synchronous waveform processing techniques for text-to speech synthesis using diphones». In: *Speech Commun.* 9.5/6, pp. 453–467.
- Moulines, E. e J. Laroche (1995). «Non-parametric technique for pitch-scale and time-scale modification of speech». In: *Speech Commun.* 16, pp. 175–205.
- Peeters, G. e X. Rodet (1999). «SINOLA: A new analysis/synthesis method using spectrum peak shape distortion, phase and reassigned spectrum». In: *Proc. ICMC*, pp. 153–156.
- Poullin, J. e D. A. Sinclair (1957). *The Application of Recording Techniques to the Production of New Musical Materials and Forms. Application to Musique Concète*. Technical Translation TT-646. National Research Council of Canada, pp. 1–29.
- Poullin, Jacques (1954). «L'apport des techniques d'enregistrement dans la fabrication de matières et formes musicales nouvelles. Applications à la musique concrète». In: *L'Onde Électrique* 324, pp. 282–291.
- Risset, Jean-Claude (1998). «Example of the musical use of digital audio effects». In: *Proc. DAFX-99 Digital Audio Effects Workshop*, pp. 254–259.
- Roads, Curtis (1996). *The Computer Music Tutorial*. Cambridge, London: The MIT Press.
- (1998). «Micro-sound, history and illusion». In: *Proc. DAFX-98 Digital Audio Effects Workshop*, pp. 260–269.
- Roucos, S. e A. M. Wilgus (1985). «High quality time-scale modification for speech». In: *Proc. ICASSP*, pp. 493–496.

- Schaeffer, Pierre (1973). «La Musique Concète». In: *Q SJ* 1287.
- Schnell, N. et al. (2002). «Synthesizing a choir in real-time using pitch synchronous overlap add (PSOLA)». In: *Proc. 1st IEEE Benelux Workshop Model based Proc. Coding Audio*.
- Springer, Anton (1955). «Ein akustischer Zeitregler». In: *Gravesaner Blätter* 1, pp. 32–37.
- Springer, Anton M. (1959). «Akustischer Tempo- und Tonlagenregler». In: *Gravesaner Blätter* 13.80, pp. 32–37.
- Truax, Barry (1988a). «Discovering inner complexity: Time-shifting and transposition with a real-time granulation technique». In: *Computer Music Journal* 18.2, pp. 38–48.
- (1988b). «Real-time granular synthesis with a digital signal processor». In: *Computer Music Journal* 12.2, pp. 14–26.
- (2000a). *Granular synthesis*. URL: <http://www.sfu.ca/truax/gran.html> (visitato il 07/01/2016).
- (2000b). *Granulation of sampled sounds*. URL: <http://www.sfu.ca/truax/gsample.html> (visitato il 07/01/2016).
- Verhelst, W. e M. Roelands (1993). «An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech». In: *Proc. ICASSP*, pp. 554–557.
- Warstat, M. e T. Görne (1994). *Studiotechnik - Hintergrund und Praxiswissen*. Elektor-Verlag.
- White, P. (1999). *Creative Recording, Effects and Processors*. Sanctuary Publishing.
- Zölzer, Udo, cur. (2011). *Digital Audio Effects*. 2<sup>a</sup> ed. John Wiley & Sons, Ltd.



## *Stato dell'arte nella sintesi di texture sonore*

Diemo Schwarz\*

(Traduzione ricevuta il 23 novembre 2015<sup>1</sup>)

### Sommario

La sintesi di *texture sonore* quali i suoni della pioggia, del vento, o di una folla, trova importanti applicazioni nel cinema e in varie forme di creatività multimediale, come giochi e installazioni. Tuttavia, nonostante ben definiti requisiti di naturalezza e di flessibilità, nessun metodo automatico si è ancora affermato. Dopo un chiarimento terminologico e relativo ai contesti d'utilizzazione della sintesi di suoni di questo tipo, daremo qui uno sguardo d'insieme sui molti metodi esistenti e sulle ancora poche implementazioni software disponibili, classificandoli in base ai differenti modelli di sintesi sottostanti, che siano di sintesi sottrattiva o additiva, granulare, concatenativa [*corpus-based concatenative synthesis*], e ancora sintesi mediante wavelet o mediante modelli fisici. Forniremo inoltre una rassegna dei metodi di analisi usati ai fini di questo tipo di sintesi, basati su criteri di segmentazione, sul rilevamento di proprietà statistiche, o sull'analisi del timbro e delle transizioni.

## 1 Introduzione

La sintesi di *texture sonora* quali i suoni della pioggia, del vento, o di una folla, trova importanti applicazioni nel cinema e in varie forme di creatività multimediale, come giochi e installazioni. Con *texture sonora* intendiamo generalmente una sonorità composta da molti micro-eventi, ma le cui caratteristiche restano stabili ad una scala di tempi più ampia, come avviene appunto nel suono della pioggia, del fuoco, del vento, nei suoni prodotti da giochi d'acqua, dal traffico automobilistico, dal movimento di masse di persone. Dobbiamo distinguere questo concetto rispetto a quello di *soundscape*, che invece indica l'insieme dei suoni che compongono una scena, alcuni dei quali a loro volta possono essere di tipo texturale.

Ci sono innumerevoli metodi per sintetizzare suoni del genere, secondo approcci molto differenti tra loro. Cercheremo di darne una classificazione e di riassumere lo stato dell'arte in questo settore. Cominceremo con una definizione dei termini e degli obiettivi d'utilizzo (sezioni 1.1, 1.2 e 1.3), per passare poi alla descrizione di metodi per la sintesi e l'analisi di texture sonore (sezioni 2 e 3) e dei primi software esistenti (sezione 4). Infine (sezioni 5 e 6), metteremo in evidenza le ricerche che appaiono più notevoli nel quadro di questa nostra ricostruzione complessiva.

---

\*Real-Time Music Interaction team, IRCAM, Centre Pompidou – Parigi.

<sup>1</sup>Il presente testo è la traduzione, a cura di Agostino Di Scipio, dell'articolo del 2011 «State of the Art in Sound Texture Synthesis», reperibile online al seguente url [http://recherche.ircam.fr/pub/dafx11/Papers/30\\_e.pdf](http://recherche.ircam.fr/pub/dafx11/Papers/30_e.pdf).

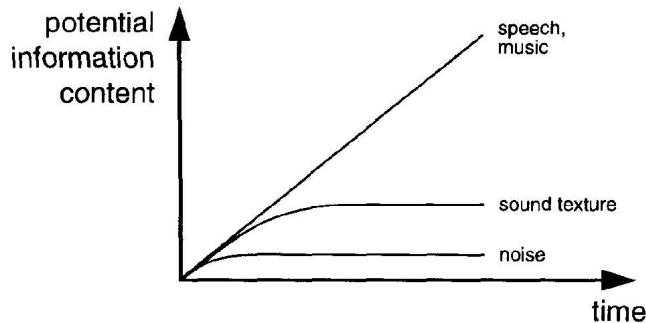
## 1.1 Definizione di texture sonora

Una prima definizione approfondita è quella avanzata da Saint-Arnaud [75] e Saint-Arnaud e Popat [76] a margine di vari esperimenti sulla percezione e generazione di texture sonore. È possibile riassumerla con la seguente analogia visiva:

Una texture sonora è come la trama di una carta da parati [wallpaper]: localmente può avere o una struttura particolare o un andamento casuale [randomness], ma nell'insieme queste caratteristiche locali rimangono costanti a scala globale.

Questa definizione è illustrata in Fig. 1. Si può arrivare dunque alla seguente definizione operativa:

1. le texture sonore sono formate da elementi sonori di base, o atomi;
2. questi elementi atomici si susseguono secondo un pattern di livello superiore che può essere periodico o casuale, o di entrambi i tipi;
3. le caratteristiche a livello generale rimangono le stesse per periodi di tempo estesi (ciò implica che non vi può essere un messaggio complesso);
4. il pattern di livello superiore dev'essere esposto nella sua completezza nel giro di pochi secondi ('tempo di attenzione');
5. una certa casualità a livello superiore è accettabile nella misura in cui gli elementi sonori all'interno del 'tempo di attenzione' sono in numero sufficiente ad evidenziare le proprietà casuali.



**Figura 1:** Contenuto di informazione potenziale di una texture sonora in funzione del tempo (da Saint-Arnaud e Popat [76]).

### 1.1.1 Ciò che una texture sonora *non è*

Per chiarire il concetto, può essere d'aiuto provare a dare una definizione negativa. Escludiamo dunque quanto segue:



**Figura 2:** Esempi di texture visive, sia naturali sia sintetiche, ‘orientate all’oscillazione’ (da Peyré [71]).

**Suoni da contatto nell’interazione tra oggetti** Impatti, frizioni, rotolamenti: tali suoni, pur trattati in certi lavori dedicati alla sintesi della texture sonora [1, 15, 49, 66, 89], contraddicono le proprietà di struttura tipo ‘carta da parati’ che abbiamo sopra indicato.

**Paesaggi sonori** [*sound scapes*] Sebbene siano talvolta trattati come texture, anche perché in effetti ne contengono sempre, i paesaggi sonori comprendono eventi sonori ricchi di informazione (come spiegheremo nella sezione 1.1.2).

**Sound design** La letteratura in questo campo spesso propone metodi utili per la sintesi di texture sonore; tuttavia il sound design in generale è un ambito più ampio di elaborazione creativa, volto a generare suoni da interazione e paesaggi sonori, oltre che texture [10, 12, 59–62].

In certi esempi di composizione e di performance musicale, la locuzione *texture sonora* [*sound texture*] viene usata per indicare materiale sonoro *non-tonico* e *non-percussivo*, o materiale musicale *non-armonico* e *non-ritmico*.

Si veda inoltre Strobl [86] per un’indagine sul termine *texture* in contesti diversi da quelli relativi al suono, come nei tessuti, in tipografia, in gastronomia.

### 1.1.2 Paesaggi sonori

Poiché le texture sonore costituiscono una parte vitale del paesaggio sonoro, è utile dare qui una breve introduzione alla classificazione e generazione automatica di paesaggi sonori. Inoltre, a sua volta chi si occupa di paesaggi sonori inevitabilmente guarda anche alla sintesi e all’organizzazione di texture sonore.

I primi tentativi in vista di una definizione e classificazione dei paesaggi sonori sono stati quelli di Raymond Murray Schafer [77], che distingue livelli diversi – come tonica [keynote], segnale [signal] e impronta sonora [sound mark] – e propone una tassonomia di referenti semantici inclusiva di attributi di tipo socio-culturale e presi dall’acustica ecologica.

In una prospettiva di ecologia acustica, invece, Gavin Gaver [37] propone una classificazione in base agli attributi fisici del suono e al tipo di interazione tra materiali.

Attualmente molteplici sono le ricerche sul paesaggio sonoro che rivestono interesse per la sintesi della texture sonora [8, 9, 34, 59–62, 90, 91].

## 1.2 Sintesi della texture: tentativi di classificazione esistenti

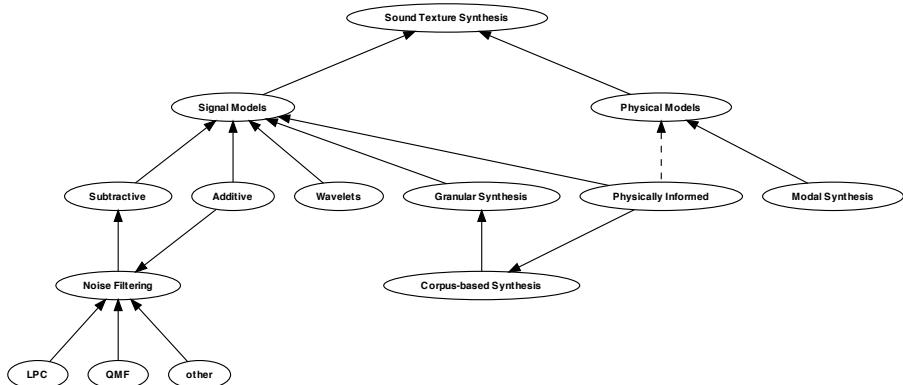
Come punto di partenza, consideriamo il tentativo di definizione di texture sonora proposto da Strobl et al. [87]. I metodi che vi sono presi in esame, appaiono divisi in due gruppi:

**Metodi provenienti dalla grafica computazionale** Trasferire nel contesto della sintesi del suono metodi usati per la generazione di texture visive [23, 65, 68]. Si veda in Fig. 2 qualche esempio di immagine texturale.

**Metodi provenienti dalla musica informatica** Applicare le tecniche di sintesi del suono usate nella *computer music* e nella sintesi del parlato [4, 6, 17, 43, 44, 95].

In una nuova rassegna degli strumenti usati nell'ambito generale del sound design e della composizione musicale [59], si propone una classificazione identica a quella che vedremo più avanti (nella sezione 2): vi si afferma che classi differenti di suoni richiedono differenti strumenti di lavoro, e viene proposta una lista di corrispondenze possibili tra differenti sonorità e metodi di sintesi che si prestano meglio ad evocarle.

In una articolo di Filatriau e Arfib [32] vengono passati in rassegna vari algoritmi per la sintesi della texture sonora dal punto di vista dei controlli e dei dispositivi di controllo. Ciò rende opportuno ricollegarsi ai diversi contesti di utilizzo pratico, come faremo ora.



**Figura 3:** Classificazione gerarchica di metodi di sintesi della texture sonora.

### 1.3 Usi e significati differenti

È importante notare che nella letteratura su questo argomento esiste un certo rischio di confusione in rapporto al preciso significato della locuzione texture sonora, e che ciò dipende dalle intenzioni d'uso. Possiamo distinguere due destinazioni d'uso più frequenti:

**Sintesi di texture expressive** Qui lo scopo è la generazione interattiva del suono nel contesto della composizione e della performance musicale, o nel contesto di altre forme di arte sonora, comunque nel senso di uno strumento musicale digitale [Digital Musical Instrument, DMI] a fini espressivi. Il termine texture sonora viene allora spesso adoperato per distinguere il materiale così generato da suoni di tipo tonico o percussivo, cioè per denotare in modo generico un materiale caratterizzato nel timbro più che nell'altezza o nel ritmo.

Alcuni dei metodi impiegati a tale scopo possono dar luogo anche a texture evocative di sonorità naturali, come osserva Di Scipio [17]. Tuttavia, in contesti del genere non sono stati compiuti approfondimenti sullo spazio parametrico adoperabile in tal senso, e rimane quindi all'utente (al musicista) forzare i processi verso sonorità di tipo naturale. Ciò vale anche per uno dei lavori di ricerca già citati [32] e per altri che ne sono scaturiti [33].

**Sintesi di texture naturali** Qui lo scopo è di sintetizzare suoni texturali, sia prodotti dall'uomo sia prodotti dall'ambiente circostante, facenti parte del paesaggio sonoro complessivo, che tornino utili in contesti applicativi quali il cinema, i videogiochi, o altri contesti di produzione audio-visiva. Ci si sforza dunque di conseguire un certo grado di realismo (come nel rendering di immagini texturali fotorealistiche), sebbene poi in molti casi sia sufficiente conseguire una *sintesi texturale credibile* di tipo puramente simbolico o impressionistico, talché la texture ottenuta riesca a portare l'informazione ambientale desiderata anche senza effetti davvero realistici, come per esempio nella progettazione urbanistica. La maggior parte degli esempi che discuteremo qui mira a questo tipo di risultato.

## 2 Classificazione dei metodi di sintesi

Proporremo ora una classificazione dei metodi esistenti per la sintesi della texture sonora. Sembra piuttosto sensato dividere i differenti approcci in base ai processi di sintesi adoperati (ed in base ai processi di analisi, se effettivamente applicati):

- filtraggio di rumore (sezione 2.1) e sintesi additiva con sinusoidi (sezione 2.2.);
- modellazione fisica (sezione 2.3) e modelli di segnale informati in senso analogo;
- rappresentazione e sintesi mediante wavelet (sezione 2.4);
- sintesi granulare (sezione 2.5) e sue estensioni del tipo ‘sintesi concatenativa’ fondata su estese basi di dati (sezione 2.6);
- metodi non-standard, come mappe frattali o caotiche (sezione 2.7).

La Fig. 3 offre una visione d'insieme di queste diverse classi, e le loro relazioni interne. Altre possibilità classificatorie potrebbero riguardare altri aspetti: fino a che punto un metodo dipenda da un modello specifico, o fino a che punto è guidato da dati; quali sono le potenzialità

per la generazione del suono in tempo reale; se viene valutata l'attendibilità del metodo in base a test di ascolto. Alcuni di questi aspetti sono discussi nella sezione 5.

## 2.1 Sintesi sottrattiva

Il filtraggio di rumore è il metodo ‘classico’ per la sintesi di texture sonore, spesso usato a partire da una modellazione specifica della sorgente sonora.

In base alla particolare definizione operativa di cui abbiamo parlato sopra, Saint-Arnaud e Popat [76] costruiscono un primo modello di sintesi che adopera rumore filtrato mediante un sistema a sei bande con filtri in quadratura a specchio [6-band Quadrature Mirror filtered noise].

In altri contributi [4, 95] viene applicata l’analisi mediante predizione lineare a cascata sia nel dominio del tempo che della frequenza [cascaded time and frequency domain linear prediction, CTFLP], in particolare per risintetizzare il rumore sullo sfondo di eventi rilevati in primo piano (vedremo più avanti, nella sezione 3.1.3, quali tecniche di rilevamento sono utilizzate). Gli eventi sono sequenziati mediante una distribuzione di Poisson. Centrale, in questi contributi, è la volontà di ridurre la quantità di dati da trasmettere con codifica a basso bitrate.

In McDermott, Oxenham e Simoncelli [54] vengono applicati al filtraggio di rumore dati risultanti da un’analisi statistica (vi torneremo in 3.2), limitandone l’uso a texture piuttosto statiche e dall’altezza indeterminata, come nel caso della pioggia e altri suoni d’acqua, o come in suoni di fuoco.

In Peltola et al. [70] si studia il suono ottenuto con differenti modi di battere le mani, usando filtri regolati sulla base dell’analisi di suoni del genere, e combinando i vari risultati ottenuti per farne l’applauso di una folla, secondo un modello statistico con vari ‘gradi di entusiasmo’ e vari modi di occupare lo spazio da parte della folla.

La ben nota e apprezzata serie di esempi proposta da Andy Farnell [31], disponibile anche online<sup>2</sup>, fornisce parecchi suoni di carattere texturale e le rispettive *patches* di sintesi scritte in PURE DATA, a partire da oscillatori e filtri attentamente regolati secondo una comprensione intuitiva del fenomeno simulato, come si può vedere in questa citazione:

“Qual è la natura [del suono] della pioggia? Cosa lo causa?” Secondo certe filosofie solipsistiche a buon mercato, la pioggia “mi cade sempre adosso”, ma ciò non ci fornisce alcun aiuto. Consideriamo invece che si tratta di particelle quasi sferiche di acqua, con diametro tra 1 e 3 mm, che si muovono a velocità costante impattando infine su materiali diversi con un flusso tipico di circa 200 al secondo per metro quadrato. Ogni goccia perviene alla stessa velocità terminale, non vi sono gocce più veloci e gocce meno veloci. Sono tutte approssimativamente delle stesse dimensioni, essendo causate dalle stesse condizioni atmosferiche di partenza, dunque non ci sono gocce più pesanti e meno pesanti. Infine, la forma di una goccia di pioggia non è quella di una lacrima, come si ritiene normalmente, ma quella di una sfera quasi perfetta. Il fattore che impedisce il formarsi di un suono completamente uniforme, e che determina quindi una certa varietà di altezza tra goccia e goccia, sta nel materiale su cui avviene l’impatto. Certe volte si tratta di foglie d’albero, certe volte di un pavimento, o di un tetto, certe altre volte ancora di una pozzanghera d’acqua formata da gocce precedenti<sup>3</sup>.

<sup>2</sup>[http://obiwannabe.co.uk/tutorials/html/tutorials\\_main.html](http://obiwannabe.co.uk/tutorials/html/tutorials_main.html)

<sup>3</sup>[La citazione è tratta dalla pagina web [http://www.obiwannabe.co.uk/tutorials/html/tutorial\\_rain.html](http://www.obiwannabe.co.uk/tutorials/html/tutorial_rain.html), che corrisponde, con leggere varianti, a Farnell [31], p. 441.]

## 2.2 Sintesi additiva con sinusoidi e rumore

Il rumore filtrato viene spesso considerato complementare al concetto di ‘oscillatore’ quale usato nella sintesi mediante somma di parziali sinusoidali.

Nel progetto QCITY<sup>4</sup> si effettua una simulazione, non in tempo reale, di suoni del traffico, mediante una rappresentazione mista di sinusoidi e di rumore calibrati secondo misure ottenute dalle varie fasi di funzionamento di un motore, dal tubo di scarico, da vari elementi di smorzamento, ecc. Il modello permette di simulare diverse densità di traffico, differenti velocità di scorrimento, diverse tipologie di veicoli e di asfalto, l’assorbimento acustico di pareti, ecc. [39]. Il calcolo degli esempi sonori può richiedere ore.

In varie circostanze Charles Verron ha proposto metodi informati alla modellazione fisica [92–94] a partire da una classificazione tratta da Gaver [37] che distingue tre tipologie diverse di suoni ambientali: liquidi, solidi e aereodinamici. I metodi proposti per l’impatto tra solidi, per i suoni del vento, e per getti, gocce e onde d’acqua, si basano su cinque atomi sonori empiricamente definiti e parametrizzati: ‘impatto modale’, ‘impatto di rumore’, ‘impatto chirp’ [a frequenza variabile], ‘rumore a banda stretta’, ‘rumore a banda larga’. Ogni modello ha da due a quattro parametri di basso livello (con l’eccezione di ben trentadue ampiezze di banda per il rumore a banda larga).

Verron poi cerca scrupolosamente le corrispondenze tra parametri di livello superiore – come forza del vento, intensità di pioggia, dimensione dell’onda marina – con i controlli parametrici di livello inferiore e le rispettive distribuzioni probabilistiche di densità.

Il processo di sintesi, nel caso di Verron, sfrutta il metodo FFT-1 [74] esteso in modo tale da includere una codifica spaziale, o meglio binaurale, già a livello della trasformata veloce di Fourier (FFT), operando poi, in fase di risintesi, uno stadio di trasformata inversa (IFFT) per ciascuno dei canali in uscita<sup>5</sup>.

## 2.3 Modellazione fisica

La modellazione fisica può essere applicata alla sintesi della texture sonora con la limitazione che un diverso modello va specificato per ciascuna classe di suoni – per esempio suoni da frizione e da rotolamento, rumori di macchinari, suoni aerodinamici [64, 65]; quest’ultima possibilità implica l’estrazione del suono dell’impatto e la valutazione percettiva del realismo conseguito con la sintesi di suoni da rotolamento [si veda anche [48]]. Spesso si usano modelli di risonanza modale [20] nei quali i modi [di vibrazione] sintetizzabili senza sforzo [computazionale] sono precalcolati a partire da simulazioni [computazionalmente] costose del comportamento del corpo rigido.

Altri metodi di sintesi basati su segnali sono descritti come fisicamente informati, cioè ispirati per lo più intuitivamente a modelli fisici [13, 14], nel senso che essi controllano modelli di segnale mediante un modello fisico del comportamento della sorgente sonora. Si vedano anche Cook [12], Picard, Tsingos e Faure [72] e Verron [93] e il libro di testo di Menzies [56], su cui torneremo più avanti.

La sintesi di suoni liquidi descritta in Doel [21] è una combinazione tra un modello di segnale ispirato a modelli fisici, per singole bolle sonore – basato su Mallock [52] – ed un

---

<sup>4</sup><http://qcity.eu/dissemination.html>

<sup>5</sup>Al sito <http://www.charlesverron.com/thesis/> si possono trovare vari esempi sonori binaurali (a volte leggermente artificiali, all’ascolto) e un video che illustra i controlli sui parametri di livello superiore, valutando anche le differenze per sorgenti sonore puntiformi e sorgenti sonore spazialmente estese.

modello empirico, fenomenologico, del comportamento statistico delle bolle d'aria, ed offre una grande varietà di risultati, da singole gocce d'acqua alla pioggia, da bolle d'aria nell'acqua a flussi d'acqua e torrenti.

Un esempio estremo è la sintesi di suoni di liquidi mediante simulazione di fluidi [63], dove l'informazione per il controllo della sintesi viene derivata dalle relazioni armoniche della sfera nelle individuali bolle d'aria (fino a quindicimila bolle)<sup>6</sup>.

## 2.4 Wavelets

Il primo esempio di sintesi di texture sonore basato sulla scomposizione multiscala del segnale mediante coefficienti wavelet è dovuto ai contributi di El-Yaniv, Werman e Dubnov [27] e di Dubnov et al. [23], ripresi poi in Kersten e Purwins [46]. Qui, la rappresentazione multiscala del segnale viene ricampionata riorganizzando l'ordine di ramificazione dei coefficienti. Ogni percorso attraverso l'albero di ramificazioni viene usato per sintetizzare, invertendo la trasformata wavelet, solo una piccola porzione di suono.

Questi contributi traggono ispirazione dall'analisi e sintesi di immagini texturali, e cercano di modellare sia le dipendenze temporali sia le dipendenze gerarchiche relative alle varie scale di rappresentazioni in uso. Il lavoro di Kersten e Purwins si colloca ad uno stadio di approfondimento ancora precoce in cui il suono complessivo della texture è già riconoscibile come tale (stando a test quantitativi svolti in tal senso), ma la struttura risultante sembra troppo sottile e sgranata, perché effettivamente le dipendenze sequenziali della texture in esame non vengono modellate, per cui gran parte del dettaglio nella struttura fine del segnale si perde. Ciò viola la caratteristica di autocorrelazione che invece si è dimostrato essere importante sia nelle texture sonore sia in quelle visive [30, 54].

In Miner e Caudell [57] viene usata l'analisi wavelet per ricostruire un suono-modello avente carattere complessivamente stocastico. Le parametrizzazioni dell'analisi wavelet restituiscono una certa varietà di risultati sonori tutti leggermente diversi dal suono considerato, a partire da un unico insieme di modelli dinamici.

Un altro approccio basato su wavelet è quello di Kokaram e O'Regan [47] e O'Regan e Kokaram [67], a sua volta sviluppato a partire da un algoritmo proposto in Efros e Leung [24] per la sintesi di immagini texturali. Viene effettuata una sintesi multicanale che adopera dimensioni piuttosto grandi di segmenti di segnale, adeguati alle più diverse sorgenti (parole, grida di neonato, accelerazione di autoveicoli, colpi di tamburo). I risultati ottenuti in fase di risintesi sono convincenti e spesso privi di artefatti.

## 2.5 Sintesi granulare

Nella sintesi granulare si usano frammenti di un suono iniziale, che può essere anche un suono registrato, e possibilmente un modello statistico di (ri)composizione di tali 'grani sonori' nel tempo. L'esplorazione di questo metodo ai fini della sintesi di texture sonore è stata affrontata da vari ricercatori: Bascou e Pottier [6], Dubnov et al. [23], El-Yaniv, Werman e Dubnov [27], Fröjd e Horner [36], Hoskinson [43], Hoskinson e Pai [44] e Parker e Behm [68]. La dimensione ottimale del grano dipende dalla scala temporale tipica della texture considerata: se sufficientemente grande, la distribuzione di eventi temporali nel breve periodo

---

<sup>6</sup><http://gamma.cs.unc.edu/SoundingLiquids>

è preservata all'interno del grano stesso, pur permettendo di creare una struttura non ripetitiva su un periodo di tempo più esteso.

In Lu, Wenyin e Zhang [50] vengono ricombinati segmenti molto brevi di segnale, anche con trasposizioni, seguendo un modello fondato sulle probabilità di transizione (si veda sezione 3.4), cercando esplicitamente di prevenire ripetizioni sul tempo breve. La segmentazione sottostante, come vedremo, segue un criterio di ‘punteggio di novità’ [novelty score] nella forma di una matrice di similarità.

In Strobl [86] vengono studiati accuratamente metodi già introdotti in Hoskinson [43] e Hoskinson e Pai [44] e in Lu, Wenyin e Zhang [50], migliorandone i parametri e la fase di risintesi, e ottenendone texture percettivamente identiche a quelle prese a modello, provando poi anche a farne delle ibridazioni [86, capitolo 4]. Il lavoro di Gerda Strobl offre inoltre un’implementazione dei metodi di Lu, Wenyin e Zhang [50] nella forma di una patch PURE DATA in tempo reale<sup>7</sup>.

## 2.6 Sintesi concatenativa

Questa nuova forma di sintesi [78, 79] può essere vista come un’estensione della sintesi granulare, cioè come una sintesi granulare operante a partire da un database [11, 82, 84, 85]. Viene spesso qualificata come *corpus-based* nel senso che implica la navigazione e la selezione di pezzettini di segnale conservati in un database molto esteso, cioè in un *corpus* di dati. Ogni grano è posizionato secondo il proprio carattere sonoro in base a determinati descrittori audio, a loro volta estratti in rapporto all’altezza, all’intensità, alla brillantezza, e altri meta-dati di livello più generale attribuiti al suono che si intende risintetizzare. Ciò permette di esplorare interattivamente una grande quantità di dati sonori e di creare delle traiettorie al suo interno, producendo nuove evoluzioni timbriche pur preservando le sottigliezze nella microstruttura del segnale originale – il che è davvero importante per ottenere texture sonore convincenti.

Finney [34] utilizza un corpus di registrazioni non strutturate disponibili nel database collaborativo freesound<sup>8</sup>, come materiale di partenza per generare eventi e texture sonore nel contesto di una forma più generale di sintesi del paesaggio sonoro (si vedano i test qualitativi discussi più avanti, nella sezione 4). I campioni prelevati vengono segmentati (come vedremo in 3.1.2) e poi filtrati con filtri passa-basso e passa-alto. I residui reputati non conformi secondo le tecniche usate da Finney vengono rimossi dallo spettro. In fase di sintesi viene scelto un insieme di cinque segmenti reputati più vicini a quelli analizzati (come viene fatta questa scelta non viene spiegato esplicitamente).

In Schwarz e Schnell [84] si segnala che i metodi esistenti per la sintesi della texture sonora spesso prendono in considerazione l’intera estensione di una certa registrazione, di cui conservano le proprietà complessive cercando di evitare eventuali artefatti. Tuttavia, tali metodi in genere mancano di controllabilità in rapporto ai risultati da conseguire. Si propongono allora due nuovi metodi basati sulla modellazione statistica dei descrittori audio applicati a registrazioni di suoni texturali, usando istogrammi e modelli di mistura Gaussiani. I quali possono poi essere interpolati per governare la transizione della texture tra differenti registrazioni-obiettivo (per esempio, da pioggia leggera a pioggia pesante). I descrittori audio dei suoni registrati vengono derivati statisticamente a partire dal campionamento della trasformata inversa per controllare infine il processo di sintesi del suono in uscita; i controlli

---

<sup>7</sup>Si vedano gli esempi in [http://www.netsoc.tcd.ie/~dee/STS\\_EUSIPCO.html](http://www.netsoc.tcd.ie/~dee/STS_EUSIPCO.html)

<sup>8</sup><https://freesound.org/>

però possono essere ottenuti anche navigando interattivamente nello spazio dei descrittori stessi<sup>9</sup> (nella sezione 4 parleremo dell'applicazione CATART che è servita nei test interattivi per la valutazione di queste procedure). Per meglio coprire lo spazio dei descrittori relativi al suono da conseguire, in Schwarz e Schnell [84] il corpus di dati viene allargato automaticamente con varianti dei suoni originali ottenute dalle trasformazioni di volta in volta applicate, ma aggiungendo al corpus di dati [non i suoni ma] solo i relativi descrittori e parametri di trasformazione. In Masurelle [53] è stata condotta una prima validazione percettiva dei descrittori usati per suoni di vento, pioggia, e onde del mare, con test soggettivi ispirati a precedenti studi della percezione di sonorità ambientali [58, 88].

Il lavoro esposto in Picard, Tsingos e Faure [72] presenta aspetti simili alla sintesi concatenativa, ma usa una procedura di selezione di grani guidata da un modello fisico.

In Dobashi, Yamamoto e Nishita [18, 19] si impiega una sintesi concatenativa che sfrutta a sua volta un database ispirato a modelli fisici di suoni aerodinamici, come i suoni del vento o quelli del movimento di spade. Questi autori precalcolano i dati di emissione sonora aerodinamica di sorgenti puntiformi mediante una simulazione della turbolenza [d'aria] computazionalmente onerosa, per differenti angoli e differenti velocità di movimento; dopo di che, consultando i dati precalcolati e sommandone i contributi parziali, possono generare interattivamente il suono di un oggetto che si muove nell'aria in modo complesso.

## 2.7 Metodi di sintesi non-standard

I metodi di sintesi non-standard, per esempio basati su frattali e mappe caotiche prodotte mediante iterazione di funzioni nonlineari, sono usati per la generazione di texture a scopi prevalentemente espressivi da Di Scipio [17] e Filatriau, Arfib e Couturier [33], specialmente quando i processi di sintesi sono controllati con dispositivi di ingresso di tipo gestuale [3, 32].

# 3 Metodi di analisi delle texture sonore

L'analisi di proprietà texturali del suono può riguardare il grado di segmentazione del suono (sezione 3.1), l'andamento delle sue qualità statistiche (sezione 3.2), in particolare per quanto riguarda il timbro (sezione 3.3), o gli stati di transizione tipici della sorgente sonora presa in esame (sezione 3.4).

## 3.1 Segmentazione

### 3.1.1 Rilevamento dell'attacco

In O'Modhrain e Essl [66] viene descritto un metodo di analisi granulare che gli autori chiamano *granificazione*, che essi applicano a suoni nati dall'interazione di molti elementi parziali (pietre in una scatola, cereali in una busta), e con tali dati di analisi controllano in modo espressivo un metodo di sintesi granulare (siamo quindi nel caso descritto nella sezione 1.3): usando un algoritmo di rilevamento dell'attacco basato su criteri di soglia, entro un tempo di riattivazione limitato, essi derivano i tempi di attacco, il livello sonoro (il primo picco successivo all'attacco vero e proprio) e alcuni contenuti spettrali (conto degli attraversamenti dello zero da parte del segnale in una finestra di 100 campioni successiva all'attacco rilevato).

---

<sup>9</sup>Si vedano gli esempi in [http://imtr.ircam.fr/imtr/Sound\\_Texture\\_Synthesis](http://imtr.ircam.fr/imtr/Sound_Texture_Synthesis)

Tali parametri vanno a controllare il tempo di attivazione del sintetizzatore granulare, il suo livello sonoro e vari fattori di trasposizione – si veda anche Essl e O’Modhrain [29].

In Lee, Depalle e Scavone [49] si effettua una stima degli eventi di contatto in una sequenza di suoni da rotolamento, sottoponendo il segnale ad un filtro passa-alto e applicando una soglia di energia. I segmenti rilevati a loro volta vengono modellati, in fase di risintesi, mediante filtri LPC su varie bande di frequenza.

### 3.1.2 Rilevamento di modifiche spettrali

Come si è anticipato, Lu, Wenyin e Zhang [50] segmentano il segnale in base ad un ‘punteggio di novità’, assegnato seguendo i coefficienti di analisi indicati come MFCC [*Mel-Frequency Cepstral Coefficients*]. Vedremo più avanti (sezione 3.4) che ciò è utile anche a valutare le probabilità di transizione [tra configurazioni spettrali diverse]. La procedura di analisi di questi autori è stata ripresa e migliorata in Strobl [86].

Finney [34] a sua volta usa un metodo di segmentazione di registrazioni ambientali che sfrutta un *criterio di informazione bayesiano* (BIC) per valutare i coefficienti MFCC [2], ma introduce anche una relazione tra la lunghezza dei segmenti rilevati ed il tipo di suono preso in esame: la lunghezza del segmento dovrà avere una certa corrispondenza con la lunghezza tipica dell’evento sonoro complessivo.

### 3.1.3 Segmentazione LPC

In Kauppinen e Roth [45], il segnale viene suddiviso in segmenti o frames localmente stazionari, mediante un treno di impulsi LPC [*linear predictive coding*, codifica a predizione lineare]. Gli autori ottengono la lunghezza ottimale del frame come misura statistica a seguito della predizione sul breve e sul lungo periodo. La soglia di picco è automaticamente adattata mediante un filtro mediano della derivata dello spettro.

In modo analogo, in Zhu e Wyse [95] vengono rilevati gli eventi sonori in primo piano [rispetto ad una texture continua sullo sfondo] mediante algoritmi LPC nel dominio della frequenza: rimuovendo gli eventi rilevati, è possibile quindi isolare il ‘frastuono’ del rumore di fondo. (Nella sezione 2.1 abbiamo accennato alla sintesi sottrattiva usata in questa ricerca per modellare il rumore così isolato).

### 3.1.4 Wavelets

In Hoskinson [43] e Hoskinson e Pai [44] la sorgente sonora viene segmentata in *grani naturali*, definiti in base ai *minima* nelle variazioni di energia all’interno delle prime 6 bande wavelet, dove il suono risulta di solito maggiormente stabile.

### 3.1.5 Analisi di componenti atomiche

Altri metodi effettuano l’analisi del suono di cui si intende modellare la texture in termini di componenti di evento e di componenti spettrali, per poi darne una ricombinazione statistica in sede di sintesi [50, 51, 60, 62, 68]. Si tratta di metodi ispirati alla modellazione di suoni da impatto, come in Ahmad, Hacihabiboglu e Kondoz [1].

In Bascou [5] e Bascou e Pottier [7] il suono viene scomposto in atomi tempo-frequenza e si procede alla valutazione della corrispondenza tra questi atomi ed un dizionario di grani ‘caratteristici’ costruito empiricamente dagli autori stessi a partire dal suono da scomporre.

## 3.2 Analisi di proprietà statistiche

Dubnov et al. [23] e El-Yaniv, Werman e Dubnov [27] applicano alla sintesi di texture sonore l'algoritmo basato sulle catene di Markov di El-Yaniv, Fine e Tishby [26], costruendo il modello statistico discreto di una sequenza di percorsi nello spazio di rappresentazione wavelet del segnale (ne abbiamo accennato nella sezione 2.4).

In Zhu e Wyse [95], una volta che gli eventi in primo piano sono separati dal rumore di fondo, se ne valuta la densità con una segmentazione mediante LPC (si veda 3.1.3). Masurelle [53], a partire da esperimenti precedenti di O'Modhrain e Essl [66], ha sviluppato un metodo semplice per stimare la densità di eventi di impatto, che risulta applicabile per esempio nel caso di suoni della pioggia. Anche Doel [21] cita varie ricerche di analisi statistica di questo tipo di suoni.

In McDermott, Oxenham e Simoncelli [54] si fa uso di un'analisi statistica ispirata a ricerche di neurofisiologia e riguardante la curtosi [cioè lo spessore variabile delle curve di distribuzione] di energia in determinate bande di frequenza. I dati ricavati in questo modo sono applicati alla sintesi mediante filtraggio di rumore (il metodo è stato poi applicato anche alla classificazione di suoni ambientali, si veda Ellis, Zeng e McDermott [28]).

### 3.2.1 Analisi non orientata alla sintesi

Certi autori mirano all'analisi e alla classificazione di texture sonore senza necessariamente cercarvi implicazioni rilevanti per una eventuale fase di sintesi. È il caso di Dubnov e Tishby [22], che usano spettri di ordine superiore per la sola classificazione di suoni ambientali, e di Desainte-Catherine e Hanna [16] che propongono dei descrittori statistici validi genericamente per sonorità ricche di rumore.

In Grill [38], parlando di una particolare installazione sonora, si descrive un sistema che riproduce campioni presi da un corpo di dati, valutati in tempo reale, corrispondenti ai suoni casuali prodotti dai visitatori. Se la fase di sintesi è estremamente semplice, il metodo di *matching* [ricerca delle corrispondenze] è notevole per l'uso di *patterns di fluttuazione*, cioè dello spettro di modulazione per tutte le bande di frequenza (misurate in *barks*) di un segmento di texture lungo tre secondi. Il vettore di dati risultante è costituito da ben 744 elementi, ma viene ridotto a ventiquattro [il numero di bande critiche, misurate in *barks*], prima di operare il matching finale.

## 3.3 Analisi delle qualità timbriche

In Hanna et al. [42] si fa notare che nel contesto del recupero di informazioni musicali [*Music Information Retrieval*, MIR] si è lavorato poco sulle caratteristiche specifiche di suoni ricchi di rumore, e si propone una classificazione in quattro sotto-classi: ‘rumore colorato’, ‘rumore pseudo-periodico’, ‘rumore impulsivo’ (pioggia, applausi) e ‘rumore con componenti sinusoidali’ (vento, rumori di strada, uccelli). Dato un segnale, questi autori ne rilevano le transizioni fra queste quattro categorie, con un insieme di misure probabilistiche bayesiane. L'idea può essere generalizzata in un modello di rappresentazione del suono basato su configurazioni stocastiche di sinusoidi [40, 41].

Soltanto i metodi di sintesi concatenativa basati su un corpus di dati provano a caratterizzare i suoni di una sorgente sonora in termini di descrittori audio percettivamente significativi [25, 78, 79, 82, 84, 85].

### 3.4 Clustering e modellizzazione delle transizioni

Saint-Arnaud [75] sfrutta un *modello di probabilità basato su gruppi* [Cluster based probability model], descritto anche in Popat e Picard [73], per raggruppare secondo varie medie statistiche un certo numero di atomi sonori relativi ad un segnale in ingresso. Gli atomi sono intesi come dati di ampiezza relativi al segnale in uscita da filtri, e sono misurati in rapporto al segmento o frame corrente ma anche in rapporto ad un intorno temporale [*neighbourhood mask*] dell'andamento passato del segnale, cercando così di codificare le transizioni tipiche che accadono nella texture sonora considerata. Nel suo lavoro, Saint-Arnaud propone una classificazione delle texture sonore ed effettua vari esperimenti, poi estesi in Saint-Arnaud e Popat [76], per costruire un modello di analisi basato sul filtraggio di rumore.

Come si è già segnalato, in Lu, Wenyin e Zhang [50] le probabilità di transizione vengono modellizzate con una matrice di similarità basata su coefficienti MFCC; in Hoskinson [43] e Hoskinson e Pai [44] vengono invece modellizzate in base alla continuità rilevata all'interno di un flusso di *grani naturali* – cioè, come si è già visto, in segmenti del segnale ricavati mediante analisi wavelet. Entrambi i metodi sono stati approfonditi e migliorati in Strobl [86].

## 4 Software disponibile

Rari sono i prodotti software, gratuiti o disponibili a pagamento, destinati alla sintesi di texture sonore, specialmente se consideriamo certi particolari tipi di texture. Il solo prodotto commerciale di cui personalmente sono a conoscenza è il ‘simulatore di folle’ CROWD CHAMBER<sup>10</sup>, che prende un file audio e lo elabora con varie tecniche di pitch-shifting e time-stretching. In realtà non si tratta di un sintetizzatore di texture, ma di un processore di segnale che aggiunge un ‘effetto folla’ a sonorità vocali preesistenti. Gli esempi non sono molto convincenti.

Finney [34] presenta un sistema di sintesi del paesaggio sonoro basato su tecniche concatenate di campioni preesistenti (si veda 2.6 e 3.1.2). Una parte del suo sistema è integrata al software GOOGLE STREET VIEW<sup>11</sup>. È notevole la sua capacità di simulare il rumore del traffico a partire da singoli campioni del passaggio di automobili, classificati per tipo di auto e per velocità, e ricombinati probabilisticamente in funzione dell’ora e del numero di strade laterali (vi è anche la simulazione della trama luminosa dei fari delle auto, realizzata mediante clustering). In Finney e Janer [35] troviamo una valutazione di questo sistema, particolarmente in rapporto alla qualità immersiva dei paesaggi sonori così generati, mediante test soggettivi di ascolto con otto partecipanti. È interessante che, in questa procedura di valutazione, i paesaggi sintetici risultino sistematicamente più credibili della semplice registrazione dei luoghi particolari presi in considerazione.

Il sistema PHYA [55, 56] offre vari strumenti di lavoro, concepiti come processi di filtraggio e di risonanza ispirati a modelli fisici, al fine di sintetizzare suoni di impatto, di collisione e suoni di diverse superfici con cui si simula di entrare in contatto<sup>12</sup>.

Il sistema CATART, che è un mio personale contributo [82], effettua una sintesi concatenativa basata su corpus di dati in tempo reale, con controlli interattivi, implementata in

<sup>10</sup>[http://www.quikquak.com/Prod\\_CrowdChamber.html](http://www.quikquak.com/Prod_CrowdChamber.html)

<sup>11</sup>[http://www.w3.org/2010/06/w3car/soundscape\\_generation.pdf](http://www.w3.org/2010/06/w3car/soundscape_generation.pdf)

<sup>12</sup><http://www.tech.dmu.ac.uk/~dylan/z/phya/>

MAX/MSP (e librerie aggiuntive FTM&Co)<sup>13</sup>. È un sistema disponibile gratuitamente<sup>14</sup> che permette di navigare una proiezione bidimensionale (o anche pluridimensionale) dello spazio dei descrittori di un database di segmenti sonori, col mouse o altro controller gestuale. Si tratta effettivamente di una sorta di estensione della sintesi granulare con accesso diretto, basato su contenuti, alle caratteristiche sonore desiderate; ciò rende possibile la creazione di texture sonore in evoluzione dinamica secondo controlli precisi delle variazioni timbriche, conservando intatta la coerenza della micro-struttura del suono, almeno fintanto che i grani sono di durata abbastanza estesa (come si è detto nella sezione 2.6). Una possibilità aggiuntiva è quella di aumentare la densità texturale, elevando la velocità di attivazione dei segmenti sonori indipendentemente dalla loro lunghezza, e quindi stratificando più unità di generazione del suono, il che risulta di grande efficacia nel caso di suoni texturali come quelli della pioggia, del vento, dell'acqua o di folle di persone<sup>15</sup>.

In CATART i descrittori sono calcolati da una struttura di analisi modulare [83]. Accanto a descrittori usuali, come frequenza fondamentale, periodicità, intensità, ci sono descrittori spettrali: ‘centroide spettrale’, ‘acutezza’ [*sharpness*], ‘piattezza’ [*flatness*], ‘energia a media e alta frequenza’, ‘contenuto di frequenza elevata’, ‘pendenza spettrale’ (espressa come coefficiente di autocorrelazione del primo ordine), ed ‘energia complessiva’. Dettagli in merito si possono trovare in Peeters [69] e Schwarz [80].

## 5 Discussione

Per quanto riguarda la dipendenza da uno specifico modello, possiamo vedere che i metodi presentati riflettono una dicotomia piuttosto netta, con approcci basati su regole, da una parte, e approcci basati su dati, dall'altra: i metodi che usano modelli fisici o modelli di basso livello del segnale (sezioni 2.2 e 2.3) partono quasi tutti da una modellazione molto specifica del processo di generazione della texture, con l'eccezione dei metodi basati su filtraggio di rumore e su analisi statistiche. I metodi che usano segmenti di segnale e coefficienti wavelet (sezioni 2.4, 2.5 e 2.6), essendo costitutivamente basati su dati, appaiono più generalmente applicabili a molte e differenti sonorità texturali, e sono molto più indipendenti da uno specifico modello di texture.

Inoltre, i modelli fisici non forniscono un legame diretto tra i loro parametri interni e le caratteristiche del suono risultante. Come nota Menzies [56],

in linea di principio, in ambiente virtuale un suono può essere riprodotto accuratamente attraverso una modellazione fisica accurata. Ma anche se ciò si realizzasse, non sarebbe sufficiente per chi, lavorando per esempio nel sound design e nella sonorizzazione cinematografica, ha bisogno di forgiare il suono secondo la propria immaginazione e secondo determinati riferimenti semantici: i modelli fisici esplicativi sono spesso difficili da calibrare rispetto a comportamenti sonori desiderati, anche se controllati direttamente con parametri fisici.

In tal senso, i metodi ispirati alla modellazione fisica in modo più intuitivo risultano certamente più flessibili, ma possono ancora proporre parametri di sintesi non direttamente relativi a caratteristiche sonore effettivamente percepibili. Peraltro, i parametri usati nei

<sup>13</sup><http://ftm.ircam.fr>

<sup>14</sup><http://imtr.ircam.fr/imtr/CataRT>

<sup>15</sup>Cfr. nota 9.

modelli fisici e nei modelli di segnale possono andare bene per la simulazione di certe texture sonore, ma restano limitati ad un insieme molto ristretto di sfumature e, differentemente da metodi che usano segmenti sonori registrati (sintesi mediante wavelet e tecniche concatenate), tendono in ogni caso a non includere quegli aspetti del suono dovuti al contesto della sorgente sonora.

## 5.1 Percezione e interazione

Sono piuttosto rare, in generale, le ricerche che si occupano della percezione di texture sonore ambientali. Vi sono eccezioni [54, 58, 88], ma solo di recente si inizia a considerare seriamente una valutazione sistematica della qualità texturale ottenuta mediante sintesi, ricorrendo a test di ascolto formalmente organizzati [54, 58, 88]. Solo in Kokaram e O'Regan [47] e O'Regan e Kokaram [67], partendo da esempi precedenti [23, 27], si è pensato di definire sin dall'inizio un insieme condivisibile e comparabile di suoni-test per svolgere simili procedure di valutazione.

Infine, dobbiamo notare che in queste pagine ci siamo concentrati su modelli di sintesi e di analisi applicati a sonorità texturali dell'ambiente, trascurando le possibilità di controllo e di interazione che essi rendono praticabili. I controlli di tipo gestuale sembrano essere un approccio promettente in tal senso [3, 32, 53].

## 5.2 Letture utili

Pur avendo cercato, col nostro sguardo d'insieme, di accennare al maggior numero possibile di lavori di ricerca in questo settore, va detto che alcuni di essi rivestono un interesse superiore, e possono ben rappresentare quanto di più avanzato si sia fatto fino ad oggi. In particolare pensiamo a:

- Finney [34], per l'introduzione in questo contesto del concetto di paesaggio sonoro, per i metodi di segmentazione adoperati, e per la modellazione molto efficace dei rumori del traffico automobilistico;
- Verron [93] e Farnell [31], per l'accurata descrizione di metodi di sintesi di suoni ambientali ispirati alla modellazione fisica, ricca di intuizioni importanti su questo tipo di sonorità;
- Kokaram e O'Regan [47] e O'Regan e Kokaram [67] e Schwarz e Schnell [84] per aver conseguito quelli che per ora sembrano i risultati più convincenti.

## 6 Conclusioni

Come abbiamo visto, nonostante i problemi e i contesti d'uso siano ben delineati, nel corso degli ultimi anni la ricerca sulla sintesi della texture sonora non ha ancora condotto ad un metodo prevalente capace di soddisfare tutti i requisiti di realismo e di flessibilità. Effettivamente, poiché il requisito di realismo costituisce praticamente sempre la priorità principale, la flessibilità dei modelli e dei metodi implementati appare messa in secondo piano a favore di soluzioni pragmatiche, magari anche ricorrendo all'editing manuale in fase di post-produzione digitale di registrazioni di suoni texturali, o semplicemente mettendo in loop campioni preregistrati per applicazioni interattive quali i giochi elettronici.

In ogni caso, i risultati più avanzati conseguiti con la risintesi wavelet [47, 67] e con la sintesi granulare basata su descrittori audio [84] promettono un buon livello di applicabilità grazie alla qualità dei risultati sonori conseguiti.

## 7 Ringraziamenti

L'autore ringrazia Steven Kersten e tutti coloro che, avendo visionato la versione iniziale di questo articolo, hanno apportato correzioni, ulteriori riferimenti bibliografici e osservazioni pertinenti. Il lavoro è stato sostenuto in parte dalla Agence Nationale de la Recherche, in particolare nell'ambito del progetto Topophonie (ANR-09-CORD-022, <http://topophonie.com>).

## Bibliografia

- [1] W. Ahmad, H. Hacihabiboglu e A.M. Kondoz. «Analysis-Synthesis Model for Transient Impact Sounds by Stationary Wavelet Transform and Singular Value Decomposition». In: *Proceedings of the International Computer Music Conference (ICMC)*. 2008.
- [2] X. Anguera e J. Hernando. *Xbic: Real-time cross probabilities measure for speaker segmentation*. Paper. Univ. California, Berkeley: ICSI (International Computer Science Institute), 2005. URL: [http://www.xavieranguera.com/papers/techreport\\_xbic.pdf](http://www.xavieranguera.com/papers/techreport_xbic.pdf) (visitato il 01/12/2015).
- [3] Daniel Arfib. «Gestural strategies for specific filtering processes». In: *Proceedings of the COST-G6 Conference on Digital Audio Effects (DAFx)* (2002). URL: [http://www2.hsu-hh.de/EWEB/ANT/dafx2002/papers/DAFX02\\_Arfib\\_Couturier\\_Kessous\\_gestural\\_stategies.pdf](http://www2.hsu-hh.de/EWEB/ANT/dafx2002/papers/DAFX02_Arfib_Couturier_Kessous_gestural_stategies.pdf).
- [4] M. Athineos e D. P. W. Ellis. «Sound texture modelling with linear prediction in both time and frequency domains». In: *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on* 5 (2003), V–648–51 vol.5.
- [5] C. Bascou. «Modélisation de sons bruités par la Synthèse Granulaire». Rapport de stage de DEA ATIAM. Université Aix-Marseille II, 2004.
- [6] C. Bascou e L. Pottier. «GMU, A Flexible Granular Synthesis Environment in Max/MSP». In: *Sound and Music Computing Conference*. Citeseer. 2005.
- [7] C. Bascou e L. Pottier. «New sound decomposition method applied to Granular Synthesis». In: *Proc. ICMC*. Barcelona, Spain, 2005.
- [8] D. Birchfield, N. Mattar e H. Sundaram. «Design of a generative model for soundscape creation». In: *International Computer Music Conference, Barcelona, Spain*. Citeseer. 2005.
- [9] D. Birchfield et al. «Generative Soundscapes for Experiential Communication». In: *Society for Electro Acoustic Music in the United States, Muncie, IN* (2005).
- [10] P. Cano et al. «Semi-Automatic Ambiance Generation». In: *Proceedings of the COST-G6 Conference on Digital Audio Effects (DAFx)*. Naples, Italy, 2004.

- [11] Marc Cardle. *Automated Sound Editing*. Rapp. tecn. University of Cambridge, UK: Computer Laboratory, mag. 2004. URL: <http://www.cl.cam.ac.uk/users/mpc33/Cardle-Sound-Synthesis-techreport-2004-low-quality.pdf>.
- [12] P. R. Cook. «Din of an “iquity”: Analysis and synthesis of environmental sounds». In: *International Conference on Auditory Display (ICAD2007)*. 2007, pp. 167–172.
- [13] Perry R. Cook. «Physically informed sonic modeling (PhISM): Percussive synthesis». In: *Proceedings of the International Computer Music Conference (ICMC)* (1996). URL: <http://quod.lib.umich.edu/cgi/p/pod/dod-idx/physically-informed-sonic-modeling-phism-percussive.pdf?c=icmc;idno=bbp2372.1996.071>.
- [14] Perry R. Cook. «Physically informed sonic modeling (phism): Synthesis of percussive sounds». In: *Computer Music Journal* (1997). URL: <http://www.jstor.org/stable/3681012>.
- [15] Richard Corbett et al. «TimbreFields: 3D Interactive Sound Models for Real-Time Audio». In: *Presence: Teleoperators and Virtual Environments* 16.6 (2007), pp. 643–654. eprint: <http://www.mitpressjournals.org/doi/pdf/10.1162/pres.16.6.643>. URL: <http://www.mitpressjournals.org/doi/abs/10.1162/pres.16.6.643>.
- [16] M. Desainte-Catherine e P. Hanna. «Statistical Approach FOR Sound Modeling». In: *Digital Audio Effects (DAFx)*. Citeseer, 2000.
- [17] A. Di Scipio. «Synthesis of environmental sound textures by iterated nonlinear functions». In: *Digital Audio Effects (DAFx)*. 1999.
- [18] Yoshinori Dobashi, Tsuyoshi Yamamoto e Tomoyuki Nishita. «Real-time rendering of aerodynamic sound using sound textures based on computational fluid dynamics». In: *ACM Trans. Graph.* 22 (3 2003), pp. 732–740. URL: <http://doi.acm.org/10.1145/882262.882339>.
- [19] Yoshinori Dobashi, Tsuyoshi Yamamoto e Tomoyuki Nishita. «Synthesizing sound from turbulent field using sound textures for interactive fluid simulation». In: *Proc. of Eurographics*. 2004, pp. 539–546.
- [20] Kees van den Doel, Paul G. Kry e Dinesh K. Pai. «FoleyAutomatic: physically-based sound effects for interactive simulation and animation». In: *Proceedings of the 28th annual conference on Computer graphics and interactive techniques. SIGGRAPH ’01*. New York, NY, USA: ACM, 2001, pp. 537–544. URL: <http://doi.acm.org/10.1145/383259.383322>.
- [21] Kees van den Doel. «Physically based models for liquid sounds». In: *ACM Trans. Appl. Percept.* 2 (4 2005), pp. 534–546. URL: <http://doi.acm.org/10.1145/1101530.1101554>.
- [22] S. Dubnov e N. Tishby. «Analysis of sound textures in musical and machine sounds by means of higher order statistical features». In: *IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 5. IEEE, 1997, pp. 3845–3848. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=604726](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=604726).

- [23] Shlomo Dubnov et al. «Synthesis of Audio sound Textures by Learning and Resampling of Wavelet Trees». In: *IEEE Computer Graphics and Applications* 22.4 (2002), pp. 38–48.
- [24] A. A. Efros e T. K. Leung. «Texture Synthesis by Non-parametric Sampling». In: *International Conference on Computer Vision*. Vol. 2. Citeseer. 1999, p. 1033.
- [25] Aaron Einbond, Diemo Schwarz e Jean Bresson. «Corpus-Based Transcription as an Approach to the Compositional Control of Timbre». In: *Proceedings of the International Computer Music Conference (ICMC)*. Montreal, QC, Canada, 2009.
- [26] R. El-Yaniv, S. Fine e N. Tishby. «Agnostic classification of Markovian sequences». In: *Proceedings of the 1997 conference on Advances in neural information processing systems 10*. MIT Press. 1998, pp. 465–471.
- [27] Z. B. J. R. El-Yaniv, D. L. M. Werman e S. Dubnov. «Granular Synthesis of Sound Textures using Statistical Learning». In: *Proc. ICMC*. 1999.
- [28] Daniel P. W. Ellis, Xiaohong Zeng e Josh H. McDermott. «Classifying Soundtracks with Audio Texture Features». In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (2010). URL: <http://www.ee.columbia.edu/~dpwe/pubs/EllisZM11-texture.pdf>.
- [29] G. Essl e S. O'Modhrain. «Scrubber: an interface for friction-induced sounds». In: *2005 conference on New interfaces for musical expression*. National University of Singapore. 2005, p. 75.
- [30] G. Fan e X. G. Xia. «Wavelet-based texture analysis and synthesis using hidden Markov models». In: *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 50.1 (2003).
- [31] Andy Farnell. *Designing Sound*. MIT Press, ott. 2010. URL: <http://mitpress.mit.edu/catalog/item/default.asp?ttype=2&tid=12282>.
- [32] J. J. Filatriau e D. Arfib. «Instrumental gestures and sonic textures». In: *International Conference on Sound and Music Computing (SMC)*. 2005.
- [33] J. J. Filatriau, D. Arfib e J. M. Couturier. «Using visual textures for sonic textures production and control». In: *Digital Audio Effects (DAFx)*. 2006.
- [34] N. Finney. «Autonomous generation of soundscapes using unstructured sound databases». MSc Thesis. UPF, 2009. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.168.4139&rep=rep1&type=pdf>.
- [35] N. Finney e J. Janer. «Soundscape Generation for Virtual Environments using Community-Provided Audio Databases». In: *W3C Workshop: Augmented Reality on the Web*. June 15 - 16, 2010 Barcelona, 2010. URL: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.168.4252&rep=rep1&type=pdf>>, <<http://www.w3.org/2010/06/w3car/>>.
- [36] M. Fröjd e A. Horner. «Sound texture synthesis using an overlap-add/granular synthesis approach». In: *Journal of the Audio Engineering Society* 57.1/2 (2009), pp. 29–37. URL: <http://www.aes.org/e-lib/browse.cfm?elib=14805>.

- [37] William W. Gaver. «How do we hear in the world? Explorations in ecological acoustics». In: *Ecological psychology* (1993). URL: <[http://www.tandfonline.com/doi/abs/10.1207/s15326969eco0504\\_2](http://www.tandfonline.com/doi/abs/10.1207/s15326969eco0504_2)>, <[http://www.cog.brown.edu/courses/cg195/pdf\\_files/fall07/Gaver-howdowehear.pdf](http://www.cog.brown.edu/courses/cg195/pdf_files/fall07/Gaver-howdowehear.pdf)>.
- [38] T. Grill. «Re-texturing the sonic environment». In: *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound*. ACM, 2010, pp. 1–7. URL: <<http://portal.acm.org/citation.cfm?id=1859805>>, <<http://grrrr.org/research/publications/paper-am-2010/>>.
- [39] S. Guidati e Head Acoustics GmbH. «Auralisation and psychoacoustic evaluation of traffic noise scenarios». In: *Journal of the Acoustical Society of America* 123.5 (2008), p. 3027.
- [40] P. Hanna. «Statistical modelling of noisy sounds : spectral density, analysis, musical transformations and synthesis». Tesi di dott. Laboratoire de Recherche en Informatique de Bordeaux (LaBRI), 2003. URL: <http://dept-info.labri.fr/~hanna/phd.html>.
- [41] Pierre Hanna e Myriam Desainte-Catherine. «A Statistical and Spectral Model for Representing Noisy Sounds with Short-Time Sinusoids». In: *EURASIP Journal on Advances in Signal Processing* 2005.12 (2005), pp. 1794–1806. URL: <http://www.hindawi.com/journals/asp/2005/182056.abs.html>.
- [42] Pierre Hanna et al. «Audio features for noisy sound segmentation». In: *Proceedings of the 5th International Symposium on Music Information Retrieval (ISMIR'04)*. 2004, pp. 120–123. URL: <http://www.ee.columbia.edu/~dpwe/ismir2004/CRFILES/paper184.pdf>.
- [43] R. Hoskinson. «Manipulation and Resynthesis of Environmental Sounds with Natural Wavelet Grains». Tesi di dott. The University of British Columbia, 2002. URL: [https://www.cs.ubc.ca/grads/resources/thesis/May02/Reynald\\_Hoskinson.pdf](https://www.cs.ubc.ca/grads/resources/thesis/May02/Reynald_Hoskinson.pdf).
- [44] Reynald Hoskinson e Dinesh Pai. «Manipulation and Resynthesis with Natural Grains». In: *International Computer Music Conference (ICMC)*. Havana, Cuba, set. 2001, pp. 338–341.
- [45] Ismo Kauppinen e Kari Roth. «An Adaptive Technique for Modeling Audio Signals». In: *Digital Audio Effects (DAFx)*. Limerick, Ireland, dic. 2001. URL: <http://www.csis.ul.ie/dafx01/proceedings/papers/kauppinen.pdf>.
- [46] Stefan Kersten e Hendrik Purwins. «Sound Texture Synthesis with Hidden Markov Tree Models in the Wavelet Domain». In: *International Conference on Sound and Music Computing (SMC)*. Barcelona, Spain, lug. 2010.
- [47] Anil Kokaram e Deirdre O'Regan. «Wavelet Based High Resolution Sound Texture Synthesis». In: *Audio Engineering Society Conference: 31st International Conference: New Directions in High Resolution Audio*. Giu. 2007. URL: <http://www.aes.org/e-lib/browse.cfm?elib=13952>.
- [48] M. Lagrange et al. «Objective quality measurement of the excitation of impact sounds in a source/filter model». In: *Acoustical Society of America Journal* 123 (2008), p. 3746.

- [49] Jung Suk Lee, Philippe Depalle e Gary Scavone. «Analysis / Synthesis of Rolling Sounds Using a Source Filter Approach». In: *Proceedings of the COST-G6 Conference on Digital Audio Effects (DAFx)*. Graz, Austria, set. 2010.
- [50] L. Lu, L. Wenyin e H.J. Zhang. «Audio textures: Theory and applications». In: *IEEE Transactions on Speech and Audio Processing* 12.2 (2004), pp. 156–167. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1284343](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1284343).
- [51] L. Lu et al. «Audio textures». In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Vol. 2. 2002. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.2.4586&rep=rep1&type=pdf>.
- [52] A. Mallock. «Sounds Produced by Drops Falling on Water». In: *Proceedings of the Royal Society of London* 95.667 (1918), pp. 138–143. URL: <https://ia600502.us.archive.org/15/items/philtrans00768948/00768948.pdf>.
- [53] Aymeric Masurelle. «Gestural control of environmental texture synthesis». Rapport de stage de DEA ATIAM. Ircam–Centre Pompidou, Université Paris VI, 2011.
- [54] J. H. McDermott, A. J. Oxenham e E. P. Simoncelli. «Sound Texture Synthesis via Filter Statistics». In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*. New Paltz, NY, 2009.
- [55] Dylan Menzies. «Phya and vfoley, physically motivated audio for virtual environments». In: *Proceedings of the Audio Engineering Society Conference* (2010). URL: <http://www.aes.org/e-lib/browse.cfm?elib=15171>.
- [56] Dylan Menzies. «Physically Motivated Environmental Sound Synthesis for Virtual Worlds». In: *EURASIP Journal on Audio, Speech, and Music Processing* (2011). URL: <http://www.asmp.eurasipjournals.com/content/2010/1/137878>.
- [57] Nadine E. Miner e Thomas P. Caudell. «Using wavelets to synthesize stochastic-based sounds for immersive virtual environments». In: *ACM Trans. Appl. Percept.* 2 (4 2005), pp. 521–528. URL: <http://doi.acm.org/10.1145/1101530.1101552>.
- [58] Nicolas Misdariis et al. «Environmental Sound Perception: Metadescription and Modeling Based on Independent Primary Studies». In: *EURASIP Journal on Audio, Speech, and Music Processing* (2010). URL: <http://articles.ircam.fr/textes/Misdariis10b/>.
- [59] A. Misra e P. R. Cook. «Toward synthesized environments: A survey of analysis and synthesis methods for sound designers and composers». In: *Proc. ICMC*. 2009.
- [60] A. Misra, P. R. Cook e G. Wang. «A New Paradigm for Sound Design». In: *Proceedings of the COST-G6 Conference on Digital Audio Effects (DAFx)*. 2006.
- [61] A. Misra, G. Wang e Perry Cook. «Musical Tapestry: Re-composing Natural Sounds». In: *Journal of New Music Research* 36.4 (2007), pp. 241–250.
- [62] Ananya Misra, Perry R. Cook e Ge Wang. «Tapestrea: Sound scene modeling by example». In: *ACM SIGGRAPH 2006 Sketches*. ACM. 2006, p. 177. URL: <<http://dl.acm.org/citation.cfm?id=1180070>>, <[http://soundlab.cs.princeton.edu/publications/taps\\_sketch.pdf](http://soundlab.cs.princeton.edu/publications/taps_sketch.pdf)>.

- [63] William Moss et al. «Sounding Liquids: Automatic Sound Synthesis from Fluid Simulation». In: *ACM Transactions on Graphics* 28.4 (ago. 2009), pp. 1–12. URL: <http://portal.acm.org/citation.cfm?doid=1559755.1559763>.
- [64] E. Murphy et al. «Perceptual Evaluation of a Real-time Synthesis Technique for Rolling Sounds». In: *Conference on Enactive Interfaces*. Pisa, Italy, 2008.
- [65] J. F. O'Brien, C. Shen e C. M. Gatchalian. «Synthesizing sounds from rigid-body simulations». In: *2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*. ACM New York, NY, USA. 2002, pp. 175–181.
- [66] S. O'Modhrain e G. Essl. «PebbleBox and CrumbleBag: tactile interfaces for granular synthesis». In: *2004 conference on New interfaces for musical expression*. National University of Singapore. 2004, p. 79.
- [67] D. O'Regan e A. Kokaram. «Multi-Resolution sound texture synthesis using the Dual-Tree complex wavelet transform». In: *Proc. 2007 European Signal Processing Conference (EUSIPCO)*. 2007. URL: <http://www.eurasip.org/Proceedings/Eusipco/Eusipco2007/Papers/A3L-B03.pdf>.
- [68] J. R. Parker e B. Behm. «Creating audio textures by example: tiling and stitching». In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* 4 (2004), iv–317–iv–320 vol.4.
- [69] Geoffroy Peeters. *A large set of audio features for sound description (similarity and classification) in the Cuidado project*. Rapp. tecn. version 1.0. Paris, France: Ircam – Centre Pompidou, apr. 2004. URL: [http://www.ircam.fr/anasyr/peeters/ARTICLES/Peeters\\_2003\\_cuidadoaudiofeatures.pdf](http://www.ircam.fr/anasyr/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf).
- [70] Leevi Peltola et al. «Synthesis of hand clapping sounds». In: *Audio, Speech, and Language Processing, IEEE Transactions on* 15.3 (mar. 2007), pp. 1021–1029. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=4100694](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4100694).
- [71] Gabriel Peyré. *Oriented Patterns Synthesis*. 2007.
- [72] Cecile Picard, Nicolas Tsingos e François Faure. «Retargetting Example Sounds to Interactive Physics-Driven Animations». Anglais. In: *n AES 35th International Conference, Audio in Games*. London Royaume-Uni, 2009. URL: <http://hal.inria.fr/inria-00394469/en/>.
- [73] K. Popat e R. W. Picard. «Cluster-based probability model and its application to image and texture processing.» In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 6.2 (gen. 1997), pp. 268–84. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=551697](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=551697).
- [74] Xavier Rodet e Phillippe Depalle. «A new additive synthesis method using inverse Fourier transform and spectral envelopes». In: *International Computer Music Conference (ICMC)*. Ott. 1992.
- [75] Nicolas Saint-Arnaud. «Classification of Sound Textures». Tesi di dott. Universite Laval, Quebec, MIT, 1995.
- [76] Nicolas Saint-Arnaud e Kris Popat. «Analysis and synthesis of sound textures». In: *Readings in Computational Auditory Scene Analysis*. 1995, pp. 125–131.

- [77] R. Murray Schafer. *The Soundscape*. Destiny Books, 1993, p. 320. URL: <http://www.amazon.com/Soundscape-R-Murray-Schafer/dp/0892814551>.
- [78] Diemo Schwarz. «Concatenative Sound Synthesis: The Early Years». In: *Journal of New Music Research* 35.1 (mar. 2006). A cura di Adam T. Lindsay. Special Issue on Audio Mosaicing, pp. 3–22.
- [79] Diemo Schwarz. «Corpus-Based Concatenative Synthesis. Assembling sounds by content-based selection of units from large sound databases». In: *IEEE Signal Processing Magazine* 24.2 (mar. 2007). A cura di Shih-Fu Chang. Special Section: Signal Processing for Sound Synthesis, pp. 92–104.
- [80] Diemo Schwarz. «Data-Driven Concatenative Sound Synthesis». Thèse de doctorat. Paris: Université Paris 6 – Pierre et Marie Curie, 2004. URL: <http://mediatheque.ircam.fr/articles/textes/Schwarz04a>.
- [81] Diemo Schwarz. «State of the Art in Sound Texture Synthesis». In: *Proc. of the 14th International Conference on Digital Audio Effects (DAFx-11)*. Paris, set. 2011, pp. 221–231. URL: [http://recherche.ircam.fr/pub/dafx11/Papers/30\\_e.pdf](http://recherche.ircam.fr/pub/dafx11/Papers/30_e.pdf) (visitato il 23/11/2015).
- [82] Diemo Schwarz, Roland Cahen e Sam Britton. «Principles and Applications of Interactive Corpus-Based Concatenative Synthesis». In: *Journées d'Informatique Musicale (JIM)*. GMEA, Albi, France, mar. 2008. URL: <http://mediatheque.ircam.fr/articles/textes/Schwarz08a/index.pdf>.
- [83] Diemo Schwarz e Norbert Schnell. «A Modular Sound Descriptor Analysis Framework for Relaxed-Real-Time Applications». In: *Proc. ICMC*. New York, NY, 2010.
- [84] Diemo Schwarz e Norbert Schnell. «Descriptor-based Sound Texture Sampling». In: *International Conference on Sound and Music Computing (SMC)*. Barcelona, Spain, lug. 2010, pp. 510–515. URL: <http://articles.ircam.fr/textes/Schwarz10a/index.pdf>.
- [85] Diemo Schwarz et al. «Real-Time Corpus-Based Concatenative Synthesis with CataRT». In: *Proceedings of the COST-G6 Conference on Digital Audio Effects (DAFx)*. Montreal, Canada, set. 2006, pp. 279–282.
- [86] G. Strobl. «Parametric Sound Texture Generator». MSc Thesis. Universität für Musik und darstellende Kunst, Graz; Technische Universität Graz, 2007. URL: <http://en.scientificcommons.org/43580321>.
- [87] G. Strobl et al. «Sound texture modeling: A survey». In: *Sound and Music Computing Conference*. 2006.
- [88] Patrick Susini et al. «Characterizing the sound quality of air-conditioning noise». In: *Applied Acoustics* 65-8 (2004), pp. 763–790. URL: <http://articles.ircam.fr/textes/Susini04b/>.
- [89] N. Tsingos, E. Gallo e G. Drettakis. «Perceptual audio rendering of complex virtual environments». In: *ACM SIGGRAPH 2004 Papers*. ACM. 2004, pp. 249–258.
- [90] A. Valle, V. Lombardo e M. Schirosa. «Simulating the Soundscape through an Analysis/Resynthesis Methodology». In: *Auditory Display* (2010), pp. 330–357.

- [91] Andrea Valle, Vincenzo Lombardo e Mattia Schirosa. «A Graph-based System for the Dynamic Generation of Soundscapes». In: *15th International Conference on Auditory Display*. Copenhagen, 2009, pp. 217–224.
- [92] C. Verron et al. «Spatialized Synthesis of Noisy Environmental Sounds». In: *Auditory Display*. Springer-Verlag, 2010, pp. 392–407. URL: <http://www.springerlink.com/index/J3T5177W11376R84.pdf>.
- [93] Charles Verron. «Synthèse immersive de sons d'environnement». PhD Thesis. Université Aix-Marseille I, 2010.
- [94] Charles Verron et al. «Contrôle intuitif d'un synthétiseur d'environnements sonores spatialisés». In: *10eme Congres Français d'Acoustique*. 2010. URL: <<http://cfa.sfa.asso.fr/cd1/data/articles/000404.pdf>>, <<https://hal.archives-ouvertes.fr/hal-00546833>>.
- [95] X. Zhu e L. Wyse. «Sound texture modeling and time-frequency LPC». In: *Proceedings of the COST-G6 Conference on Digital Audio Effects (DAFx)*. Vol. 4. 2004.



## *Abstracts*

### *Acoustical quanta and the theory of hearing*

Dennis Gabor

Writing in 1946 and 1947, the author considers the possibility to describe acoustical phenomena with quantum methods, and emphasizes that such an approach on sound analysis connects very neatly with issues in the psychophysiology of hearing. Central to Gabor's paper is a representation of sound signals including the explicit determination of both frequency and time variables (different from the classic Fourier analysis, where sound is analysed in terms of perfectly periodical oscillations, i.e. oscillations of indeterminate duration). The author illustrates an «information diagram» that splits the sound signal in «logons» or acoustical quanta, e.g. rectangular units of minimal time/frequency area. These latter are formally described as functions or «elementary signals» consisting of a real part (oscillation) and an imaginary part (a pseudo-Gaussian pulse). The author shows that the time/frequency unit area in the information diagram are subject to quantum uncertainty. And that the definition of such unit area of acoustical energy might definitely converge with a number of (then accepted) empirical measurements in the psychophysiology of hearing.

### *Gabor's acoustical quanta in sound and music technologies*

Agostino Di Scipio

Through a series of brilliant paper publications dating from the late 1940s, Dennis Gabor developed a new conceptual and operational framework for the analysis of sound signals, based on a quantum-oriented view of acoustical phenomena. In the present essay, I try to illustrate the shaping up of Gabor's quantum analysis of sound, especially as delineated in two papers from 1946 (*Theory of communication*) and 1947 (*Acoustical quanta and the theory of hearing*), and to overview its legacy in scientific research as well as in audio and musical applications. After some introductory remarks, I follow a hybrid path, between history of science and history of audio technology, sketching a “genealogy” of Gabor's quantum view of sound (i.e. as

connected to Heisenberg's indeterminacy principle, to Mach's analysis of sensation, etc.), and relating it to an empiricist tradition of modern science. I eventually situate his research in the context of contemporary research preoccupations shared by other, at the time, and I finally discuss some of the earliest audio devices that seem to tie back to Gabor's own practical experiments (beside his theoretical framework) and that revealed of primary interest to pioneers in analog and digital music technologies and related creative practices.

*Xenakis and the “granular connection”*

Makis Solomos

Many ways start from or go through Xenakis. One of these ways is the granular paradigm. In this paper, I will start from Xenakis, suggesting that, in his aesthetic, this approach is a “theory” in the ancient meaning of the word, and searching for the constituent elements of this “vision”. Then, I will try to incorporate the granular paradigm into a musical historicity, searching for his becoming inside other aesthetics, in particular with Horacio Vaggione's and Agostino Di Scipio's music.

*Interacting with Inner and Outer Sonic Complexity: from Microsound to Soundscape Composition*

Barry Truax

It is possible to think of the two extremes of the world of sound as the inner domain of microsound (less than 50 ms) where frequency and time are interdependent, and the external world of sonic complexity, namely the soundscape. In terms of sonic design, the computer is increasingly providing tools for dealing with each of these domains, through such practices as granular synthesis and multi-channel soundscape composition.

*Time-segment processing*

Pierre Dutilleux, Giovanni De Poli, Adrian von dem Knesebeck and Udo Zölzer

The text deals with sound processing in the discrete time domain, whereas a phase of segmentation, a subsequent analysis, and a temporal signal of the recombination strategy in order to obtain transformations on audio resynthesized. Various algorithms are presented illustrating the historical analogue techniques as well as developments and implementations using the digital audio techniques. Anthology musical compositions are also cited where composers and musicians use these calculations, contributing to the evolution of techniques of electronic music and sound.

*State of the art in sound texture synthesis*

Diemo Schwartz

The synthesis of sound textures, such as rain, wind, or crowds, is an important application for cinema, multimedia creation, games and installations. However, despite the clearly defined requirements of naturalness and flexibility, no automatic method has yet found widespread use. After clarifying the definition, terminology, and usages of sound texture synthesis, we will give an overview of the many existing methods and approaches, and the few available software implementations, and classify them by the synthesis model they are based on, such as subtractive or additive synthesis, granular synthesis, corpus-based concatenative synthesis, wavelets, or physical modeling. Additionally, an overview is given over analysis methods used for sound texture synthesis, such as segmentation, statistical modeling, timbral analysis, and modeling of transitions.



