

L'arte orale.
Poesia,
musica,
performance

a cura di
Lorenzo Cardilli
Stefano Lombardi Vallauri

aA ccademia
university
press



Trasversalmente, arte orale è ogni genere artistico che faccia uso della voce: antico o moderno; occidentale o extra-occidentale; popolare o autoriale; non scritto (nell'accezione consolidata della "tradizione orale") oppure anche scritto; non mediato oppure mediato dalla tecnologia audiovisiva; in tempo reale o differito; *in loco* o a distanza; linguistico o anche non linguistico; genere puro (poesia, vocalizzo extraverbale) o misto (teatro, melologo, canzone). E in generale è forma significativa ma insieme anche materiale presenza, sonora e corporea.

Questo volume raccoglie gli atti di un convegno svoltosi nel maggio 2019 presso l'Università IULM di Milano, che alle ordinarie sessioni accademiche affiancava anche "sessioni performative" artistiche. La prospettiva è necessariamente interdisciplinare: l'oralità è trattata come l'elemento comune che caratterizza arti diverse quali la poesia, la musica e il teatro, nonché le loro storiche commistioni (ad esempio il canto epico, la poesia lirica), fino a fenomeni contemporanei come il rap, il *poetry slam*, la *vocal performance art*. Sono convocati a dialogare studiosi di differente estrazione: estetica, teoria letteraria, poesia contemporanea, metrica, linguistica, (etno)musicologia, storia del teatro, *performance studies*. L'interazione dei vari punti di vista consente di affrontare l'oralità nella sua valenza trasversale e, al tempo stesso, negli aspetti specifici propri di ciascuna espressione artistica.

MIMESIS JOURNAL BOOKS
collana di «Mimesis Journal. Scritture della performance»

ISSN 2283-8783

comitato scientifico

Antonio Attisani Università degli Studi di Torino

Florinda Cambria Università degli Studi di Milano

Lorenzo Mango Università degli Studi L'Orientale di Napoli

Tatiana Motta Lima Universidade Federal do Estado do Rio de Janeiro

Franco Perrelli Università degli Studi di Torino

Antonio Pizzo Università degli Studi di Torino

Kris Salata Florida State University

Carlo Sini Università degli Studi di Milano

Éric Vautrin Université de Caën

**L'arte orale.
Poesia,
musica,
performance**

**a cura di
Lorenzo Cardilli
Stefano Lombardi Vallauri**

aA

**L'arte orale.
Poesia,
musica,
performance**

Volume pubblicato con il contributo
del Dipartimento di Comunicazione, arti e media
Università IULM, Milano

© 2020
Accademia University Press
via Carlo Alberto 55
I-10123 Torino



prima edizione novembre 2020
isbn 979-12-80136-18-3
edizioni digitali www.aAccademia.it/arteorale
<http://books.openedition.org/aaccademia/179>

book design boffetta.com

Introduzione	Lorenzo Cardilli, Stefano Lombardi Vallauri	VII
The Performance of Poetry: Voice and Voicing in an Age of Writing	Jonathan Culler	3
Perdere la voce. Le metamorfosi della poesia letteraria	Stefano Ghidinelli	28
La performance dalla parte dell'ascolto	Paolo Giovannetti	54
Mettere in voce il verso. La pratica del reading di poesia fra letteratura e vocalità	Rosaria Lo Russo	71
Phonodia. La voce dei poeti e l'uso delle registrazioni	Alessandro Mistrorigo	91
Sulle differenze fra poesia e canzone	Luca Zuliani	107
L'oralità di canzone. Per una filologia dei testi poetico-musicali popular	Mario Gerolamo Mossa	128
Poesia come discorso	Umberto Fiori	147
Poesia sonora, poesia d'azione: la poesia e lo scarto epigenetico	Giovanni Fontana	151
Arcipelago voce	Michela Garda	158
Il canto dei passi: voce e ritmo del corpo nella performance dei canti epici del Kosovo	Nicola Scaldaferrì	186
L'oralità simulata: produzione sonora e dimensione empatica della voce nella canzone registrata	Alessandro Bratus	202
Dalla prosodia alla musica strumentale: una sfida compositiva	Fabio Cifariello Ciardi	225
"The body of the voice / the voice of the body". Vocal Performance Art and its connection to the cognitive unconscious	Theda Weber-Lucks	245
L'oralità torna in scena	Martina Treu	263
La voce che dischiude: poetica, profetica, sonora	Enrico Pitozzi	281
Marina Abramović: silenzio, respiro e grido	Mauro Petruzzello	301

Dalla prosodia alla musica strumentale: una sfida compositiva

Fabio Cifariello Ciardi

aA

1. Introduzione

Le possibilità sonore dell'apparato vocale dell'uomo sono state ampiamente sfruttate in ogni forma di arte orale, compresa la musica, ben oltre l'ambito del canto. I compositori si sono interessati alla parola parlata e alle altre vocalizzazioni paralinguistiche sia come mezzo per trasmettere significati, anche non convenzionali, sia come fonte da esplorare per le sue proprietà acustiche¹. Diverse lingue dall'Africa e dall'Asia orientale usano l'intonazione per veicolare significati diversi mentre altre, come anche l'italiano, fanno uso di forme di accentuazione timbrica per differenziare parole ambigue (ad esempio "pèsca"/"pésca"). Tutte prevedono la possibilità di modulare l'altezza, l'intensità e il ritmo della voce per comunicare emozioni e concetti anche ben oltre le parole. In questo senso, quello della prosodia è un linguaggio che supera molti confini geografici e culturali².

225

1. Cfr. C. Lane, *Voices from the Past: compositional approaches to using recorded speech*, «Organised Sound», 11/01 (2006), pp. 3-11; A.J. Keller, *The Music of Language*, doctoral dissertation, Rice University, 2019.

2. M. Russo (ed.), *Prosodic Universals. Comparative Studies in Rhythmic Modeling and Rhythm Typology*, Aracne, Roma 2010.

La possibilità di precisare in una partitura musicale le caratteristiche prosodiche del linguaggio parlato è tuttavia limitata. Più il recitante deve seguire delle istruzioni precise, meno la sua voce risulta autentica e naturale. È forse anche a causa di questo limite che Arnold Schönberg cambiò a più riprese la notazione del suo *Sprechgesang*, una tecnica vocale a metà strada tra il canto e il parlato: dalla precisa notazione delle altezze nei *Gurre Lieder* (1901-11) e nel *Pierrot Lunaire* (1912), fino alle più generiche indicazioni in *A Survivor from Warsaw* del 1947³.

Un altro approccio, che qui approfondiremo, ha cercato di trasformare questo limite in una risorsa: invece di piegare la prosodia alle loro istanze creative, alcuni compositori hanno lasciato che fossero le caratteristiche prosodiche della voce parlata a modellare la loro musica strumentale.

Suoni simili a risate e singhiozzi sono frequenti nelle parti strumentali della musica klezmer⁴, mentre allusioni al pianto si trovano nel “lament bass” della tradizione musicale barocca⁵. È però soprattutto la voce parlata ad essere stata imitata con ogni sorta di strumento: dalle percussioni agli strumenti elettronici, e certamente non solo per ragioni musicali⁶. Per approfondire i motivi di un interesse così ampio la fonologica e la psicologia della percezione uditiva si sono spesso occupate del rapporto fra la musica e la parola⁷, ma non hanno mai trovato delle

3. Ph. Bryn-Julson, P. Mathews, *Inside Pierrot lunaire. Performing the Sprechstimme in Schoenberg's masterpiece*, Scarecrow Press, Lanham 2008.

4. H. Netsky, *American Klezmer: A Brief History*, in M. Slobin (ed.), *American Klezmer: Its Roots and Offshoots*, University of California Press, 2002, pp. 13-23.

5. B. Brover-Lubovsky, *Tonal space in the music of Antonio Vivaldi*, Indiana University Press, 2008, pp. 151-168.

6. Oggi la voce parlata è ricreata artificialmente con una prosodia così perfetta da essere tanto persuasiva quanto una voce naturale. Ad esempio cfr. M. Dubiel et al., *Persuasive synthetic speech: voice perception and user behaviour*, in «Proceedings of the 2nd Conference on Conversational User Interfaces», 2020, <<https://dl.acm.org/doi/10.1145/3405755.3406120>>.

7. Solo per citare alcuni studi, cfr. M. Besson, D. Schön, *Comparison between language and music*, «Annals of the New York Academy of Sciences», 930/1 (2001), pp. 232-258; C. Palmer, S. Hutchins, *What is musical prosody?*, «Psychology of learning and motivation», 46 (2006), pp. 245-278; A. Patel, J.R. Iversen, J.C. Rosenberg, *Comparing the rhythm and melody of speech and music: The case of British English and French*, «The Journal of the Acoustical Society of America», 119/5 (2006), pp. 3034-3047; D. Deutsch, T. Henthorn, R. Lapidis, *Illusory transformation from speech to song*, «The Journal of the Acoustical Society of America», 129/4

ragioni d'interesse per integrare delle dimensioni musicali all'interno dei propri modelli⁸. La musicologia, da parte sua, ha approfondito il lavoro di compositori attratti dalle proprietà prosodiche del parlato (primo fra tutti Leoš Janáček), ma solo di rado ha affrontato l'argomento con un approccio sistematico⁹.

In questo quadro il presente contributo cercherà di approfondire l'uso della prosodia in un campo musicale specifico, quello della musica strumentale contemporanea. Tre saranno le domande alle quali si cercherà di rispondere. Da dove nasce l'attuale interesse per la trascrizione strumentale della voce parlata? Quali sono i principali compositori che l'hanno utilizzata? Con quali finalità e con quali metodologie hanno operato?

2. Dalla composizione con suoni referenziali alla trascrizione della parola parlata

2.1. L'imitazione dei suoni del mondo

L'interesse dei compositori per la trascrizione strumentale della voce parlata rientra in quello più ampio per la dimensione sonora del mondo che ci circonda. Non tutti i suoni sono ugualmente utili in questo senso, ma quelli che nelle varie epoche sono risultati particolarmente familiari sono stati spesso utilizzati nella musica strumentale, sia per mezzo di simbolismi sonori, sia attraverso vere e proprie imitazioni. I primi fra questi sono certamente quelli della natura: almeno dalla tradizione allusiva seicentesca¹⁰ e fino alle trascrizioni dei canti degli uccelli utilizzate da Olivier Messiaen e oltre.

In seguito alla trasformazione del paesaggio sonoro dovuto alla rivoluzione industriale anche i suoni delle

(2011), pp. 2245-2252; R.J. Zatorre, S.R. Baum, *Musical melody and speech intonation: Singing a different tune*, «PLoS Biology», 10/7 (2012); A.R. Meireles *et al.*, *Musical Speech: A New Methodology for Transcribing Speech Prosody*. «Interspeech», 2017, pp. 334-338; C. Deguchi, R. Cubelli, *Musica e linguaggio: due sistemi che interagiscono*, «Giornale italiano di psicologia», 44/1 (2017), pp. 51-94.

8. I. Chow, S. Brown, *A Musical Approach to Speech Melody*, «Frontiers in psychology», 9/247 (2018). (Vai 19810)0)an University.rd,

9. S. Glaser, *The Missing Link: Connections Between Musical and Linguistic Prosody*, «Contemporary Music Review», 19/3 (2000), pp. 129-154.

10. C. Fertonani, *Antonio Vivaldi. La simbologia musicale nei concerti a programma*, Studio Tesi, Roma 1992, p. 27.

macchine hanno progressivamente catturato l'attenzione dei compositori, prima fra tutte la locomotiva a vapore. Nel 1923 Arthur Honegger dedicò al suono di un locomotore la sua partitura orchestrale *Pacific 231* per rendere una «contemplation objective: la tranquille respiration de la machine au repos, l'effort du démarrage, puis l'accroissement progressif de la vitesse, pour aboutir à l'état lyrique, au pathétique du train de 300 tonnes lancé en pleine nuit à 120 à l'heure»¹¹. Negli stessi anni i progressi tecnologici iniziavano a rendere già disponibili alcuni dispositivi per la registrazione e riproduzione del suono offrendo ai compositori nuovi mezzi per integrare nella musica i suoni del mondo. L'anno dopo *Pacific 231*, Ottorino Respighi includeva un fonografo nell'orchestra del suo poema sinfonico *Pini di Roma* per riprodurre il suono di un vero usignolo¹². Il 15 maggio 1930, ben prima della *musique concrète* di Pierre Schaeffer, la radio di Berlino trasmetteva *Weekend*, un montaggio esclusivamente sonoro – realizzato non da un compositore, ma dal regista Walter Ruttmann – con frammenti di conversazione, canzoni e altri eventi sonori registrati e montati utilizzando una particolare pellicola cinematografica¹³. In seguito, le innovazioni e la diffusione delle tecnologie audio hanno reso possibile l'uso di una grande quantità di suoni dall'alto potere referenziale. Particolarmente rilevanti in questo ambito sono le tradizioni musicali elettroniche che dalla metà degli anni Sessanta in poi hanno utilizzato suoni ambientali riconoscibili: la fonografia di François-Bernard Mâche¹⁴, la musica aneddotica di Luc Ferrari¹⁵ o la composizione del paesaggio sonoro dei canadesi Raymond Murray Schafer, Barry Truax e Hildegard Westerkamp¹⁶.

11. A. Honegger, *Pacific (231)*, «Dissonances», 4 (1924), p. 80, cit. in P. Jost, *Composer en images? À propos des symphonies de Honegger*, in M. Duchesneau, F. de Médicis, S. Caron (éds.), *Musique et modernité en France (1900-1945)*, Presses de l'Université de Montréal, Montréal 2006, pp. 281-299.

12. Cfr. R. Gelatt, *The Fabulous Phonograph. From Edison to Stereo*, Appleton-Century, New York 1966.

13. A. Birtwistle, *Photographic Sound Art and the Silent Modernity of Walter Ruttmann's Weekend (1930)*, «The New Soundtrack», 6/2 (2016), pp. 109-127.

14. Cfr. F.B. Mâche, *Music, Myth, and Nature, or, The Dolphins of Arion*, Taylor & Francis, Abingdon 1992.

15. Cfr. L. Ferrari, presentazione di *Hétérozygote* (1964), <<http://lucferrari.com/analyses-reflexion/heterozygote/>>.

16. Cfr. R.M. Schafer, *The Soundscape. Our Sonic Environment and the Tuning of the*

Dal canto loro, i compositori di musica strumentale, soprattutto di scuola francese, hanno cercato di ricostruire con gli strumenti dell'orchestra prima la dimensione timbrica inconsueta di percussioni inarmoniche come campane, gong e tam-tam¹⁷ e quindi quella particolarmente ricca delle note più gravi di strumenti come il trombone (in *Partiels* (1975) per 18 strumenti di Gérard Grisey) o il pianoforte (in *Désintégrations* (1982) per 17 strumenti e suoni elettronici di Tristan Murail). Negli anni Settanta, i compositori di musica spettrale hanno cominciato a riferirsi a questi processi con il termine "sintesi strumentale" perché, così come accade con le sinusoidi utilizzate nella sintesi additiva, anche qui l'involuppo spettrale di un suono naturale viene ricostruito, parziale per parziale, ma con gli strumenti dell'orchestra¹⁸. Più recentemente, sia l'interesse per il paesaggio sonoro sia quello per la sintesi strumentale hanno contribuito alla formalizzazione di un ulteriore concetto, quello della "ri-sintesi strumentale"¹⁹: una pratica compositiva in cui gli strumenti acustici sono utilizzati per trascrivere vari tipi di suoni referenziali attraverso la mediazione di programmi semi-automatici di analisi (ad es. Praat, AudioSculpt, Melodyne, Spear), ri-sintesi (ad es. AudioSculpt, CataRT) e orchestrazione (ad es. Orchidée, Orchids, Orchidea, Ato-ms).

aA

229

2.2. L'imitazione strumentale del parlato

La voce parlata, il mezzo privilegiato di comunicazione di ogni nostro rapporto sociale, rientra a pieno titolo nella categoria dei suoni che ci sono familiari.

L'interesse per la trascrizione del parlato mediante strumenti acustici è già documentato da lettere, schizzi,

World, Destiny Books, Rochester 1977; B. Truax, *Acoustic Communication*, Ablex, Norwood (NJ) 1984. ilippo Del Corno d, NJ, (1984);

17. Cfr. J. Anderson, *A provisional history of spectral music*, «Contemporary Music Review», 19/2 (2000), pp. 7-22.

18. Cfr. J. Fineberg, *Guide to the basic concepts and techniques of spectral music*, «Contemporary Music Review», 19/2 (2000), pp. 81-113.

19. Cfr. G. Carpentier, *Approche computationnelle de l'orchestration musicale. Optimisation multicritère sous contraintes de combinaisons instrumentales dans de grandes banques de sons*, dissertazione dottorale, Université Paris VI, 2008; Y. Marez, *On computer-assisted orchestration*, «Contemporary Music Review», 32/1 (2013), pp. 99-109; J. O'Callaghan, *Mimetic Instrumental Resynthesis*, «Organised Sound», 20/2 (2015), pp. 231-240; N. Donin, *Sonic Imprints: Instrumental Resynthesis in Contemporary Composition*, in G. Borio (ed.), *Musical Listening in the Age of Technological Reproduction*, Ashgate, Aldershot 2015, pp. 323-341.

libri e spartiti di studiosi come l'inventore del XVII secolo Joshua Steele²⁰, di compositori dell'Ottocento come Modest Mussorgsky²¹ e soprattutto Leoš Janáček²². L'entusiasmo di quest'ultimo, testimoniato da innumerevoli trascrizioni raccolte dal 1897 al 1928 in ben 56 taccuini²³, è anche apertamente confessato in un passo del suo feuilleton del 1903 *Moje Luhačovice*, in cui si legge: «seguiamo le tracce di queste straordinarie melodie del linguaggio! Sono come fiori di ninfea che tremano sulla superficie calma di una vita felice e senza preoccupazioni»²⁴.

Negli ultimi decenni del Novecento la melodia del parlato e le altre dimensioni della prosodia hanno suscitato l'interesse di un numero crescente di compositori di diversa estrazione. Molti sono autori d'avanguardia come Peter Ablinger²⁵, Clarence Barlow²⁶, Robert Davidson, Robert Erickson²⁷, Vinko Globokar²⁸, Jonathan Harvey²⁹, Per Magnus Lindborg³⁰, François Bernard Mâche³¹, Steve Reich³²,

20. J.C. Kassler, *Representing speech through musical notation*, «Journal of Musicological Research», 24/3-4 (2005), pp. 227-239.

21. A. Orlova, *Musorgsky's Days and Works. A Biography in Documents*, UMI Research Press, Ann Arbor 1983, p. 515.

22. J. Pearl, *Eavesdropping with a Master: Leoš Janáček and the Music of Speech*, «Empirical Musicology Review», 1/3 (2006), pp.131-165.

23. Cfr. il progetto *Nápěvky mluvy*, o *Leoš Janáček. Nápěvky mluvy. Kritické vydání* (Leoš Janáček. Le melodie del parlato. Edizione critica), <www.napevkymluvy.cz>.

24. L. Janáček, *My Luhačovice* (1903), in M. Beckerman (ed.), *Janáček and His World*, Princeton University Press, Princeton-Oxford 2003, p. 227.

25. P. Ablinger, *Quadraturen*, 2006, <<https://ablinger.mur.at/docu11.html>>.

26. C. Barlow, *Music Derived from Other Sources*, «The International Journal of the Humanities», 9/7 (2011), pp. 149-160.

27. J. MacKay, *On the Music of Robert Erickson: A Survey and Some Selected Analyses*, «Perspectives of New Music», 26/2 (1988), pp. 56-85.

28. J. Warnaby, *Vinko Globokar: revaluating a phenomenon*, «Tempo», 61/240 (2007), pp. 2-18.

29. G. Nouno et al., *Making an orchestra speak*, in F. Gouyon, Á. Barbosa, X. Serra (eds.), *Proceedings of the 6th Sound and Music Computing Conference*, Porto 2009, pp. 277-282.

30. P.M. Lindborg, *About TreeTōrika: Rhetoric, C AAC and Mao*, in J. Bresson, C. Agon, G. Assayag (eds.), *The OM Composer's Book #2*, Delatour-IRCAM, Paris 2008, pp. 95-116.

31. F.B. Mâche, *Music, myth, and nature* cit.

32. M. Beirens, *Different Tracks*, in S. Gopinath, P. ap Siñon (eds.), *Rethinking Reich*, Oxford University Press, New York 2019, pp. 75-98.

Scott Johnson³³, Jacob ter Veldhuis aka Jacob TV³⁴; ma non mancano i musicisti provenienti anche da altri generi musicali come Paul Brody, Christophe Chassol, Jason Moran³⁵, Hermeto Pascoal³⁶, Steve Vai, Frank Zappa³⁷.

L'interesse compositivo riguarda i riferimenti sia lessicali che prosodici veicolati dalla voce parlata. Nel primo caso, più raro, il materiale fonetico viene decostruito e ripetuto così da mettere l'ascoltatore in grado di cogliere il significato di singole parole o perfino di brevi frasi. Esempi di questa tecnica si trovano nella parola "Why" della *Sequenza V* (1966) per trombone di Luciano Berio; nelle frasi "Why me? No money. My way" di *Orchideae Ordinariae* (1989) per grande orchestra di Clarence Barlow, oppure in "Every single year", "How did this become" e "How did we get" ricostruite sempre dall'orchestra in *Background checks* (2020) per video, voce registrata e orchestra di Fabio Cifariello Ciardi.

Nel secondo e più frequente caso, invece, ad essere trascritta non è tanto la parola come entità sonora complessa quanto i suoi riferimenti prosodici: l'articolazione nel tempo della "melodia del parlato", una successione di suoni lontana dal concetto tradizionale di linea melodica eppure, al tempo stesso, familiare.

La porzione di parlato trascritto varia da singole parole o brevi frasi, come nel caso di *Different Trains* (1988) per quartetto d'archi e nastro magnetico di Steve Reich, fino a interi discorsi come nei 59 brani per pianoforte e CD di Peter Ablinger, *Voices and Piano* (1998-2019); nei *Piccoli Studi sul Potere* (2010-2013) per video e strumenti di Fabio Cifariello Ciardi o nella prosodia immaginaria dell'ultimo discorso del controverso generale americano Douglas MacArthur

33. J. Maltman, L. Vickery, *The foundations of speech-melody composition and the dawn of the digital music era*, in «Proceedings of the Australasian Computer Music Association», Australasian Computer Music Association, Perth 2018, pp. 30-37.

34. G.H. Brown, *The Body of your Dreams: Profiling Jacob ter Veldhuis*, Sequenza 21, <<https://www.sequenza21.com/2007/05/the-body-of-your-dreams-profiling-jacob-ter-veldhuis/>>.

35. S. Simon, *Jason Moran: Finding Sound, Then Making It His Own*, NPR, 2010, <<https://www.npr.org/transcripts/128415897>>.

36. L. Costa-Lima Neto, *O cantor Hermeto Pascoal: os instrumentos da voz*, «Per Musi», 22 (2010), pp. 44-62.

37. S. Vai, *Tempo Mental*, 1983, <<https://www.vai.com/tempo-mental/>>.

realizzata per trombone solo da Robert Erickson nel suo *General Speech* (1969).

La tipologia delle informazioni che i compositori hanno cercato di veicolare è vasta. Si va dall'ampia gamma di emozioni o stati d'animo che un oratore può suscitare come la buffoneria e la tristezza dei personaggi della Commedia dell'Arte in *Les Rires du Gilles* (1981) per 5 strumenti e nastro magnetico di Michaël Lévinas; lo stile retorico di artisti (il citato *Voices and Piano* di Ablinger), poeti (ad es. *Appunti per Amanti Simultanei* per trombone, ensemble d'intonarumori ed elettronica scritto da Cifariello Ciardi nel 2009 e basato sulla voce di Filippo Tommaso Marinetti) e soprattutto di leader politici (ad es. *Man bör kalla saker och ting vid deras rätta namn* (2007-2008) per ensemble e sound files sulla voce di Olof Palme, *ReTreTorika* (2006) per 4 strumenti ed elettronica e *TreeTorika* (2006) per orchestra da camera sulla retorica di Mao Zedung, tutti scritti da PerMagnus Lindborg; il finale dell'azione di teatro musicale *Io Hitler* (2008-2009) di Filippo Del Corno; *Riforma* (2016) per ensemble su una animatissima seduta del parlamento italiano, *Cupio Dissolvi* (2017) per otetto di contrabbassi sul celebre "I have a dream" di Martin Luther King di Cifariello Ciardi); fino al contrasto fonetico tra lingue diverse come la video-opera *The Cave* (1989-1993) di Reich; *Im Januar am Nil* (1984) per orchestra da camera di Barlow; *Ab* (2007) e *Ra* (2007) per ensemble di Cifariello Ciardi. Sono stati considerati, infine, sia dialetti diversi (ad esempio la "passione" per giornalista narrante, video, ensemble ed elettronica *Voci Vicine* (2014) di Cifariello Ciardi), sia l'alternanza fra parlato e cantato come nella "operatic sound installation" *Talking Melody-Singing Story* (2016) di Paul Brody o in alcuni momenti dell'opera *Kein Licht* (2017) di Philippe Manoury.

I compositori possono cercare di imitare la natura sonora dei fonemi nel segnale vocale preso in esame, come nel caso di *Toucher* (1973) per un percussionista "parlante" di Vinko Globokar. Tuttavia, nella maggior parte dei casi, ad essere presi in considerazione sono il profilo melodico e il pattern ritmico di un'espressione verbale. Il primo è l'altezza percepita del parlato, cioè la frequenza fondamentale, detta F0, della forma d'onda glottale. Il secondo si basa sulle consonanti e le vocali che definiscono la morfologia della

sillaba. Se i fonemi vocalizzati sono utilizzati per stabilire l'altezza, quelli non vocalizzati (cioè i suoni pronunciati senza la vibrazione delle corde vocali) sono cruciali nella misurazione della durata di ciascun segmento sillabico.

Quest'ultimo passaggio è uno dei più critici. A causa del fenomeno della coarticolazione³⁸ e della rapida articolazione del parlato, la separazione fra sillabe è spesso acusticamente sfumata³⁹; mentre alla fine delle frasi, l'esaurimento dell'aria nei polmoni determina una naturale diminuzione della dinamica che rende la lunghezza delle sillabe a volte difficile da valutare.

Così come avviene per la ri-sintesi strumentale di altri tipi di suoni, anche qui il segnale audio viene inizialmente segmentato e analizzato, quindi riportato in notazione musicale attraverso un'adeguata quantizzazione. Ciascuna caratteristica della "melodia del parlato" viene tradotta nel corrispondente parametro musicale: il contorno intonativo è trascritto con intervalli temperati o microtonali, le variazioni d'intensità sono rese con opportuni accenti e dinamiche, mentre le durate delle sillabe o dei fonemi vengono raggruppate utilizzando figure ritmiche e opportune indicazioni di tempo.

aA

233

3. *La composizione strumentale della voce parlata: i problemi*

3.1 Complessità acustica del linguaggio

Non possiamo valutare a fondo le trascrizioni di pionieri come Joshua Steele⁴⁰ o Leoš Janáček⁴¹, perché nessuna delle espressioni verbali da loro trascritte è arrivata fino a noi. Osservando i loro appunti, però, arriviamo facilmente alla conclusione che le altezze e le durate da loro ascoltate sono lontane dall'essere state trascritte con precisione. La loro conoscenza fonologica era sicuramente limitata e l'affidabilità dei loro sistemi di misurazione è quantomeno discutibi-

38. R.D. Kent, F.D. Minifie, *Coarticulation in recent speech production models*, «Journal of Phonetics», 5/2 (1977), pp. 115-133.

39. K. Munhall, A. Löfqvist, *Gestural aggregation in speech: Laryngeal gestures*. «Journal of Phonetics», 20/1 (1992), pp. 111-126.

40. J. Steele, *Prosodia Rationalis: Or, an Essay Towards Establishing the Melody and Measure of Speech, to be Expressed and Perpetuated by Peculiar Symbols*, J.Nichols, London 1779.

41. Cfr. il progetto *Nápěvky mluvy* cit.

le. Tuttavia a rendere imprecisi i loro tentativi è soprattutto la grande complessità fenomenologica del parlato.

Da un punto di vista articolatorio, il fenomeno fonatorio dipende dal modo in cui le nostre strutture fisiologiche interagiscono con il flusso d'aria proveniente dai polmoni. La forma delle parole che ascoltiamo dipende da una moltitudine di modulazioni laringee e configurazioni fisse o mobili del tratto vocale che contribuiscono alla morfologia temporale e spettrale dei suoni del linguaggio. Anche se il luogo e il modo di articolazione di questi suoni è situato su un continuum, tendiamo a dividere il flusso del parlato in unità percettivamente distinte. La fonologia divide queste unità, dette fonemi, usando solo due categorie: consonanti e vocali, a seconda della presenza o meno di ostruzioni nel passaggio dell'aria. Questa riduzione non rende giustizia alle vaste possibilità timbriche dei suoni prodotti dal sistema vocale umano, ma piuttosto, evidenzia fino a che punto le esigenze del linguaggio ci abbiano spinto a descrivere un fenomeno continuo in termini di categorie discrete. Sintassi, morfologia, fonologia usano strutture e schemi per classificare i suoni linguistici perché la categorizzazione svolge un ruolo cruciale nella comunicazione vocale: supporta la nostra capacità di riconoscere parole basate su proprietà acustiche simili, ma non identiche.

Un'altra caratteristica articolatoria che rende la trascrizione della voce parlata un compito impegnativo ha a che fare con la sua velocità di variazione. La maggior parte delle lingue utilizza tra 20 e 40 fonemi per formare strutture sillabiche altamente diversificate. I vincoli linguistici influenzano il numero di sillabe che il parlante è in grado di produrre. Anche se le procedure di misurazione non sono oggettive perché dipendono da una definizione non sempre univoca dei confini di ciascuna unità sillabica, l'evidenza sperimentale mostra una velocità media tra le 5 e le 8 sillabe al secondo nella gran parte delle lingue e in soggetti a cui non viene chiesto un discorso né rapido né accurato⁴². Invece, quando i soggetti sono incoraggiati a leggere un testo nella loro lingua madre "più velocemente

aA

42. F. Pellegrino, C. Coupé, E. Marsico, *A cross-language perspective on speech information rate*, «Language» 87/3 (2011), pp. 539-558.

del normale”, i risultati mostrano valori medi tra circa 8 sillabe per secondo per l’inglese e il tedesco e 11 sillabe per secondo per il francese⁴³. In termini musicali, ciò significa che per trascrivere il parlato usando solo un suono per sillaba, ogni durata varierà tra 200 e 91 ms (ovvero tra un sedicesimo e un trentaduesimo con un metronomo di circa 82 pulsazioni per minuto). Da un lato, questo è un valore incoraggiante per i compositori poiché simili durate possono essere eseguite da molti strumenti acustici. Dall’altro, chi lavora con la traduzione strumentale del parlato sa bene che la riduzione ad un suono per sillaba può essere sì sufficiente per sottolineare i punti in comune tra una trascrizione strumentale di un enunciato e la sua versione originale, ma non è sufficientemente verosimile se la trascrizione non è ascoltata insieme alla voce originale.

Eppure una riduzione dei dati è indispensabile e possibile: non solo dal punto di vista musicale, ma anche da quello fonologico. I risultati di Piet Martens⁴⁴ dimostrano che alcune minimali variazioni di altezza non influenzano la percezione della proprietà prosodiche percepibili e possono essere eliminate attraverso un opportuno processo di stilizzazione. Peraltro, una diminuzione della complessità è implicita quando per tradurre un fenomeno continuo come quello fonatorio viene utilizzata una rappresentazione sostanzialmente discreta come quella della notazione musicale. La quantizzazione delle altezze e soprattutto delle durate è necessaria, infine, per bilanciare il rischio di una semplificazione eccessiva con le difficoltà esecutive di una trascrizione inutilmente dettagliata. Gruppi ritmici e accentuaioni inusuali sono tipiche in trascrizioni strumentali molto precise, ma sono di difficile esecuzione, soprattutto quando sono coinvolti molti musicisti. Proprio per questo motivo, spesso, la melodia del parlato viene trascritta solo per pochi strumenti.

Per quanto complessa e dagli esiti incerti, la ricerca di un compromesso ottimale fra complessità e verosimiglianza

43. V. Dellwo, P. Wagner, *Relationships between speech rhythm and rate*, in *Proceedings of the International Congress of Phonetics Science*, Barcelona 2003, pp. 471-474.

44. P. Mertens, *The Prosogram: Semi-Automatic Transcription of Prosody Based on a Tonal Perception Model*, in *Proceedings of Speech Prosody 2004*, Nara 2004, pp. 23-26.

non ha dissuaso diversi compositori dall'orchestrare sillabe, parole o intere espressioni usando anche l'ampia palette timbrica dell'ensemble strumentale (ad es. *Le son d'une voix* di François-Bernard Mâche, *Im Januar am Nil* di Clarence Barlow, *Riforma* di Cifariello Ciardi) o anche le risorse di un'intera orchestra (ad es. *Orchideae Ordinariae* di Barlow, *Speakings* di Jonathan Harvey, o *Background checks* di Cifariello Ciardi).

3.2 La ricezione del parlato trascritto

I compositori interessati alla ri-sintesi strumentale condividono alcune preoccupazioni comuni. Come abbiamo appena visto, una prima sfida è anzitutto quella di riuscire a tradurre in modo verosimile l'evoluzione spettrale e temporale del suono originale nonostante le limitazioni implicite nelle risorse strumentali disponibili. Un'altra, però, ugualmente ardua da superare, riguarda la ricezione del risultato.

Indipendentemente dagli obiettivi dichiarati, è difficile credere che il compositore non si auguri di riuscire a mettere gli ascoltatori nella condizione di poter cogliere almeno qualcuno dei riferimenti ai quali rimanda il suono "ri-sintetizzato" dagli strumenti. Per tradurre in pratica questo intento chi scrive la partitura deve scommettere sulla propria capacità di fare delle previsioni attendibili. La trascrizione si applica a un suono specifico, scelto per il suo potere di rappresentare ed evocare una rete di informazioni almeno in parte condivise; ma come prevedere quali saranno queste informazioni e come essere sicuri che siano effettivamente condivise dalla gran parte degli ascoltatori? Il più delle volte i compositori rispondono a queste domande con delle considerazioni di buon senso.

Una prima riguarda lo stretto legame fra la capacità evocativa di un suono e la rete di informazioni che definiscono l'esperienza di cui quel suono fa parte. Ad esempio, il potere referenziale del canto di un uccello dipenderà dalle informazioni che definiscono i contesti naturali solitamente associati alla presenza di volatili, così come il rombo di un motore di un'automobile sarà associato alle caratteristiche proprie di un ambiente urbano. Tanto più univoche e ricorrenti saranno le associazioni fra un suono e l'informazione

sul suo contesto, tanto più la presenza dell'uno influenzerà l'interpretazione dell'altro.

Di conseguenza, il compositore tende ragionevolmente ad aspettarsi interpretazioni almeno in parte simili da parte dagli ascoltatori che hanno condiviso una stessa esperienza e dunque una memoria sonora collettiva basata su di essa⁴⁵.

Nel caso del parlato le informazioni basate su esperienze condivise possono essere raggruppate in tre categorie. La prima riguarda i riferimenti fonetici che ci consentono di segmentare l'evoluzione spettrale-temporale dell'enunciato in unità linguistiche e ci permettono anche di determinare il genere, l'identità, l'umore o l'emozione di chi parla. La seconda comprende i riferimenti articolatori che utilizziamo sia per produrre, sia per decodificare le parole. La terza categoria, infine, riguarda i riferimenti lessicali utilizzati per assegnare significati ai suoni che produciamo e ascoltiamo durante la fonazione.

Se da una parte la scommessa del compositore è quella di riferirsi ad una o più di queste 'qualità' del parlato, un'altra difficoltà sta nel rendere il riferimento sufficientemente evidente all'ascolto, nonostante le ambiguità e le approssimazioni causate dal processo di trascrizione.

In questo contesto delle utili indicazioni ci vengono dalla letteratura sui meccanismi cognitivi implicati nei processi di riconoscimento, dal momento che alla base dell'attribuzione di una referenza c'è inevitabilmente proprio il recupero d'informazioni dalla nostra memoria. Secondo la Dual-process theory⁴⁶, la rievocazione dei ricordi si basa su due sistemi correlati, ma diversi fra loro. Il primo è il recupero d'informazioni contestuali basate su precedenti esperienze sonore condivise. Ad esempio, gli appassionati della musica classica possono facilmente associare il motto di quattro note che apre la *Quinta Sinfonia* di Beethoven sulla base delle loro precedenti esperienze con il compositore e il pezzo. Il secondo processo, invece, si fonda su una valutazione della familiarità evocata dall'evento sonoro percepito: una sensazione di conoscenza che emerge quando le associazioni

45. Cfr. J.V. Wertsch, H.L. Roediger III, *Collective memory: conceptual foundations and theoretical approaches*, «Memory», 16/3 (2008), pp. 318-326.

46. Cfr. G. Mandler, *Recognizing: The judgment of previous occurrence*, «Psychological review», 87/3 (1980), pp. 252-271.

attivate non sono sufficienti a mettere a fuoco un'informazione precisa.

Il rapporto tra l'intensità della referenza e il riconoscimento o la familiarità è piuttosto diretto: i riferimenti condivisi derivano da interpretazioni basate su associazioni simili che, a loro volta, determinano un riconoscimento o una sensazione di una familiarità diffusa all'interno nel gruppo. Tornando all'esempio precedente, un diverso gruppo di ascoltatori potrebbe avere difficoltà ad associare il motto della *Quinta Sinfonia* al suo autore e ciononostante riconoscere il motivo come familiare.

Nel nostro caso, quando la trascrizione strumentale del parlato comporta solo un debole stato referenziale, gli ascoltatori potranno al massimo presumere che gli strumenti stiano cercando d'imitare vagamente una voce parlata; il profilo dell'intonazione potrà indicare se l'espressione trascritta è un'affermazione o una domanda, ma nulla più. Invece, quando i riferimenti prosodici diventano maggiormente evidenti l'intento mimetico della musica potrà fornire suggerimenti sull'umore, le emozioni o lo stile dell'oratore. Eventuali qualità articolatorie tradotte con successo potranno suggerire una divisione del flusso sonoro in sillabe e parole. E infine, se anche le differenze tra le diverse consonanti e vocali sono rese in modo convincente, l'ascoltatore potrà persino intravedere qualche riferimento lessicale riconoscibile nel flusso musicale.

Le variabili che determinano il grado di familiarità e dunque l'intensità del potere referenziale di un suono, sono in parte individuali e sociali. Molte riguardano il coinvolgimento emotivo e la salienza che associamo a ciò che ascoltiamo in base alle nostre conoscenze acquisite e alle molte influenze sociali a cui siamo soggetti⁴⁷, ma non solo. La capacità evocativa di un suono dipende anche da condizioni acustiche sostanzialmente indipendenti dall'ascoltatore. Pensiamo ad esempio alle diverse forme di ridondanza che riguardano il suono: quelle interne, dipendenti dalle caratteristiche del segnale (la stabilità spettrale o la regolarità nelle variazioni di ampiezza) e quelle esterne, legate alla

47. Cfr. H. Leibenstein, *Bandwagon, Snob, and Veblen effects in the theory of consumers' demand*, «The Quarterly Journal of Economics», 64/2 (1950), pp.183-207.

sua presenza ricorrente in un dato contesto⁴⁸. Da questo punto di vista, la natura sociale della comunicazione verbale e il carattere sovra-linguistico della prosodia (che comunica informazioni non sempre allineate con il significato suggerito dal segno linguistico) fanno dei suoni del parlato una categoria che si differenzia da tutti gli altri eventi sonori che ci sono familiari.

Scenari sonori della durata di circa 8 secondi contenenti pianto e altre vocalizzazioni prodotte da bambini sono identificati e memorizzati meglio rispetto ad altri di durata maggiore comprendenti tipi differenti di suoni familiari⁴⁹. Invece, in un esperimento di riconoscimento, Clara Suied e altri hanno dimostrato la possibilità di un riconoscimento non casuale in frammenti sonori vocali della durata di soli 2 o 4 millisecondi, finanche più brevi di un singolo ciclo nella forma d'onda del segnale⁵⁰. Anche le proprietà acustiche della voce contribuiscono a definire il suo carattere altamente referenziale. Dalla letteratura presa in esame da Assmann e Summerfield⁵¹, ad esempio, emerge come le ridondanze del parlato presenti a livello acustico, fonologico e lessicale siano tali da consentire un buon livello di intelligibilità anche quando ad essere rimossa è la gran parte dell'informazione spettrale del segnale.

aA

239

4. *La composizione strumentale della voce parlata: le soluzioni*

4.1 L'integrazione compositiva della prosodia

Come abbiamo visto, la complessità del segnale vocale e la stratificazione delle sue qualità referenziali rendono la ri-sintesi strumentale del parlato una sfida tanto intrigante quanto ricca di insidie tecniche, percettive ed estetiche. Nel valutare il proprio lavoro i compositori interessati a esplorare questo campo si trovano davanti a due domande

48. Cfr. B. Gygi, G.R. Kidd, C.S. Watson, *Spectral-temporal factors in the identification of environmental sounds*, «The Journal of the Acoustical Society of America», 115/3 (2004), pp. 1252-1265.

49. M. Marcell *et al.*, *Identifying, rating, and remembering environmental sound events*, «Behavior Research Methods», 39/3 (2007), pp. 561-569.

50. C. Suied *et al.* *Auditory gist: Recognition of very short sounds from timbre cues*. «The Journal of the Acoustical Society of America», 135/3 (2014), pp. 1380-1391.

51. P. Assmann, Q. Summerfield, *The Perception of Speech under Adverse Conditions*, in S. Greenberg *et al.* (eds.), *Speech Processing in the Auditory System*, Springer, New York (NY) 2004, pp. 231-308.

ricorrenti. Se la trascrizione delle dimensioni prosodiche di una specifica voce parlata è ritenuta efficace, come integrarla all'interno della composizione? Invece, se il risultato della semplice trascrizione non è del tutto soddisfacente, cosa fare per migliorare la sua ricezione da parte degli ascoltatori? Le soluzioni adottate si collocano in un campo circoscritto da tre possibilità.

La prima e più radicale consiste nell'utilizzare solo ed esclusivamente la trascrizione delle dimensioni prosodiche, anche a costo di eliminare la figura dell'esecutore, come accade in *A Letter From Schoenberg* (1996) per player piano di Ablinger.

La seconda consiste nell'optare per una scelta "coloniale" che punti ad incorporare le proprietà fonetiche trascritte in uno stile musicale pre-connotato. In questo caso il risultato della trascrizione viene di volta in volta 'piegato' non tanto da limitazioni strumentali, quanto dalla necessità di salvaguardare la riconoscibilità del linguaggio musicale utilizzato. Esempi di questa tendenza si trovano nel voicing jazz di *Break down* (2006) di Jason Moran e nei moltissimi "ultrascores" del pianista e compositore Christophe Chassol, raccolti negli album da *X-Pianos* (2012) a *Ludi* (2020); nel retrogusto fusion della ritmica che accompagna la chitarra di Frank Zappa in *The Jazz Party Discharge Hats* (1983); nel sapore brasiliano di *Som Da Aura Poré* (1984) di Hermeto Pascoal; nei pattern minimalisti di *Different Trains* (1988) e *The Cave* (1993) di Steve Reich o nelle letture post-minimaliste delle *speech-melodies* su cui si basa la "Boombox Music" di Jacob TV.

Una terza possibilità, infine, è quella di cercare di avvicinare la musica strumentale al materiale sonoro preso come riferimento, piuttosto che viceversa. In quest'ottica più "ecologica", applicabile a qualsiasi tipologia di ri-sintesi strumentale, le alleanze si preferiscono alle acquisizioni. Nel caso della trascrizione della voce parlata, la traduzione del materiale fonetico è centrale, ma anche gli altri eventi sonori che la circondano e la trama armonica sono in diversi modi derivati dalle dimensioni prosodiche analizzate. Nella musica di Clarence Barlow, Robert Davidson, Jonathan Harvey, PerMagnus Lindborg e molti altri il compositore non si limita all'estrazione di materiali melodici dalla voce parlata, ma cerca di esplorare le ricadute di una applicazio-

ne della prosodia sul pensiero armonico, timbrico e formale della composizione⁵².

4.2 Il supporto tecnologico

La distinzione appena considerata si riflette anche sul modo in cui vengono utilizzate le tecnologie nell'esecuzione della musica basata sul parlato. In un approccio "coloniale", i dispositivi e i media sono scelti in funzione di un pensiero compositivo che ha le sue radici lontane dal riferimento fonetico. Al contrario, in un'ottica "ecologica" l'utilizzo delle tecnologie e dei media è anzitutto funzionale a valorizzare il risultato della trascrizione. I mezzi tecnologici possono compensare la perdita di trasparenza referenziale conseguente al processo di traduzione strumentale fornendo soluzioni diverse. La prima consiste nell'integrare la parte strumentale con la riproduzione completa o parziale della voce parlata imitata dagli strumenti musicali, eventualmente completata dalle immagini dell'oratore intento a parlare. Una seconda categoria di supporti tecnologici può aiutare l'esecutore a gestire le difficoltà della complessità della notazione e la necessaria precisione nella sincronizzazione. Quando il compositore non ha previsto l'utilizzo d'immagini il discorso originale e la sua trascrizione possono essere eseguiti in serie, uno dopo l'altro, come in *Cupio Dissolvi* di Cifariello Ciardi, affidando ad uno degli esecutori l'onore di controllare la partenza di ciascun frammento registrato del discorso. Altrimenti, se come accade più spesso la voce e la sua trascrizione strumentale sono presentate in parallelo, una click-track per l'esecutore o il direttore è sufficiente a garantire una sincronizzazione potenzialmente ottimale. Un più complesso impiego della tecnologia prevede l'utilizzo contemporaneo di tecniche di analisi, elaborazione audio in tempo reale e dispositivi di sincronizzazione flessibile in grado d'integrare le informazioni prosodiche con quelle musicali in un modo decisamente meno rigido rispetto a quelli descritti in precedenza. In *Speakings* di Jonathan Harvey, ad esempio, il materiale vocale è presentato attraverso la sua traduzione strumentale; ma il rilevamento delle

aA

241

52. F. Cifariello Ciardi, *Local and Global Connotations in Sonic Composition*, «Organised Sound», 13/2 (2008), pp. 123-135; G. Nouno *et al.*, *Making an orchestra speak* cit..

caratteristiche fonetiche è ottimizzato attraverso un sistema di analisi, sintesi e sincronizzazione in tempo reale⁵³.

Una strada comunemente adottata per aumentare l'intensità del riferimento prosodico consiste nell'aggiungere delle informazioni visive strettamente correlate con il materiale fonetico. Ad essere proiettata può essere la rappresentazione scritta delle parole pronunciate (ad esempio, in *The art of agony* (2015) per due pianoforti, video e soundfile di Davidson o in *Mountain Top* (2008) per coro misto e gruppo di percussioni, video e audio sampler di Jacob TV), oppure il volto e i gesti di chi parla, con o senza il corrispondente contenuto audio (ad es. nei citati *Piccoli Studi sul Potere* di Cifariello Ciardi), o ancora una più articolata combinazione tra segnali testuali, visivi e uditivi (come in *The Cave* di Reich, in *Background checks* e nei *Little Studies on Background checks* per violoncello, video e voce registrata di Cifariello Ciardi).

5. Conclusioni

In questo saggio abbiamo considerato le variazioni d'intonazione, ampiezza, tempo e ritmo di una voce parlata come un campo d'indagine da esplorare compositivamente con i mezzi della musica strumentale. Negli ultimi decenni diversi compositori, anche appartenenti a tradizioni musicali molto diverse fra loro, hanno prodotto un corpus di opere che può essere valutato in più modi. Le trascrizioni vocali strumentali stupiscono gli ascoltatori perché mettono in evidenza analogie tra la voce parlata e la musica spesso inaspettate. D'altra parte, però, quando cerchiamo di valutare gli intenti estetici di queste ri-sintesi strumentali abbiamo a volte la tentazione d'interpretare l'associazione fra il parlato e la sua 'melodia' nei termini di un accostamento parodistico o tutt'al più curioso, ma non sufficiente a dar forma a un pensiero musicale compiuto. Inoltre, se gli strumenti musicali possono imitare alcune dimensioni prosodiche, invariabilmente tralasciano molte delle caratteristiche acustiche cruciali nel riconoscimento della voce parlata. L'utilizzo della sorgente audio e/o visiva originale compensa solo fino ad un certo punto i limiti di una

traduzione strumentale del discorso che rischia di rimanere monca e quindi insoddisfacente per l'ascoltatore. Infine, se la prosodia è ormai un fenomeno 'melodico' traducibile con relativa facilità, altrettanto non si può dire delle dimensioni timbriche implicite nel parlato. Le complesse transizioni tra i fonemi, anche in semplici sequenze di sillabe, producono una morfologia spettrale certamente affascinante per un compositore; ma ancora una volta, i limiti degli strumenti musicali tradizionalmente utilizzati sono tali da spegnere in parte qualsiasi facile entusiasmo.

Eppure, ciononostante, rimane il sospetto che la ri-sintesi strumentale del parlato stia ancora muovendo i suoi primi passi e che abbia tutt'altro che esaurito il suo potenziale. Ci sono vari argomenti che vale la pena di considerare a questo proposito. Il modo in cui accenti, durate e profili intonativi sono articolati nelle diverse lingue è pressoché infinito e spesso musicalmente originale. Indipendentemente dai vincoli linguistici, la possibilità di sfruttare in qualche modo l'energia emotiva che accompagna la prosodia è una tentazione quasi irresistibile per chi scrive musica. Guardando il linguaggio parlato e la musica da una prospettiva più ampia, è difficile non pensare di aver solo iniziato a scalfire la superficie del terreno comune su cui entrambe queste forme di comunicazione sono cresciute. Infine, nessun'altra categoria di suoni referenziali è stata ed è studiata in modo così ampio e multidisciplinare. Quest'ultimo punto merita di essere enfatizzato perché potrebbe contribuire ad uno storico cambio di paradigma nelle pratiche compositive.

Nel 1995 Diana Deutsch ha descritto per la prima volta la "speech-to-song illusion": «a spoken phrase is perceptually transformed to sound like song rather than speech, simply by repeating it several times over»⁵⁴. Nell'esplorare il fenomeno, la studiosa ha dato un sostegno scientifico ad un'intuizione che Steve Reich aveva già utilizzato trent'anni prima nella musica del suo *It's gonna rain* (1965): le ripetizioni ottenute con il nastro magnetico montato in loop potevano intensificare la melodia delle parole pronunciate, lasciando emergere con chiarezza la qualità musicale del

54. D. Deutsch, T. Henthorn, R. Lapidis, *Illusory transformation from speech to song* cit, p. 2245.

discorso⁵⁵. Ebbene, vogliamo immaginare che in un futuro non lontano l'ampio interesse multidisciplinare per lo studio dei riferimenti prosodici possa portare a degli esiti inversi. I prossimi risultati musicalmente interessanti nella trascrizione strumentale del discorso parlato potrebbero nascere non solo dalle intuizioni e dalla pratica dei compositori, ma anche da qualche innovativo risultato scientifico prodotto dalle ricerche sulla prosodia.

55. S. Reich, *Writings on music, 1965-2000*, Oxford University Press, New York 2002, pp. 19-21.

«**Mimesis Journal**» non è soltanto il titolo di una rivista, è anche il nome di un progetto che riunisce autori di diversi paesi e generazioni interessati all'oggi del teatro e delle arti sceniche – ovvero un plesso che comprende tutto il xx secolo, oltre alle ricerche e alle teoresi più avanzate del presente – in una prospettiva multidisciplinare, interdisciplinare e transdisciplinare, dunque non esclusivamente storiografica ma piuttosto fenomenologica.

aAaAaAaAaAa

Lorenzo Cardilli è professore a contratto di *Italian and European Culture* al Politecnico di Milano. Si interessa di poesia contemporanea italiana e di storia e teoria della critica letteraria.

Stefano Lombardi Vallauri è professore associato presso l'Università IULM di Milano, direttore della rivista «Nuove Musiche». Su temi affini all'arte orale ha pubblicato il volume *La voce mediatizzata* (2019).