

La voce senz'anima: origine e storia del Vocoder

Paolo Zavagna

(Ricevuto l'11 aprile 2013)

[J]e me contentai de supposer que Dieu formât le corps d'un homme, entièrement semblable à l'un des nôtres [...] et sans mettre en lui, au commencement, aucune âme raisonnable [...]. Car, examinant les fonctions qui pouvaient en suite de cela être en ce corps, j'y trouvais exactement toutes celles qui peuvent être en nous sans que nous y pensions, ni par conséquent que notre âme, c'est-à-dire cette partie distincte du corps [...] que la nature n'est que de penser¹.

René Descartes, *Discours de la méthode* (1637)

1 L'uomo macchina

1.1 Automi

La “favola” di Descartes, che tanto impatto avrà sull'arte medica e su tutto il pensiero occidentale – non solo seicentesco –, di una macchina che possa servire da modello per lo studioso, ci descrive un automa che è in grado di funzionare senza pensare. Datogli un “fuoco senza luce”, l'automa potrà svolgere una serie di “funzioni tutte del medesimo genere”. L'uomo macchina, che La Mettrie² sancirà in seguito quale modello da seguire nell'arte medica, è il simbolo, nella cultura occidentale, della scissione fra anima e corpo. Questa scissione determinerà lo scorporo (letteralmente parlando) fra meccanismo (corpo) e controllo (anima).

¹ “[M]i contentai di supporre che Dio formasse il corpo di un uomo interamente simile a uno dei nostri [...] e senza mettere in esso, inizialmente, alcuna anima razionale [...]. Infatti, esaminando le funzioni che in conseguenza di ciò, potevano trovarsi in quel corpo, vi trovavo esattamente tutte quelle che sono presenti in noi senza che ci pensiamo, e senza quindi che vi contribuisca la nostra anima, cioè quella parte distinta del corpo la cui natura [...] consiste unicamente nel pensare.” Descartes, *Discorso sul metodo*, pp. 58-59, parte V, traduzione leggermente modificata.

² La Mettrie, *L'homme machine*. In quest'opera viene superata la dicotomia fra anima e corpo, sostenendo l'inutilità dell'ipotesi dell'anima, e portando agli estremi la visione meccanicistica dell'uomo, macchina in tutto e per tutto.

Non è di secondaria importanza l'appartenenza di queste osservazioni alla medicina che, dall'Ottocento, da 'arte' (tecnica) diverrà 'scienza' (statistica), in quanto spesso l'aspetto di patologia e di cura accompagnerà gli studi sul comportamento della voce cosiccome dell'udito. Dal *Tentamen resolvendi problema . . .* di Kratzenstein, fino alla cibernetica, la 'sintesi' della voce troverà applicazioni in ambito medico³.

Nel 1668, il filosofo Géraud de Cordemoy ribadirà quanto espresso da Descartes riferendosi proprio alla voce e in particolare alla parola.

Car, encore que je conçoive bien qu'une pure machine pourroit proferer quelques paroles, je connois en même temps que, si les ressorts qui distriburoient le vent, ou qui feroient ouvrir les tuyaux, d'où ces voix sortiroient, avoient un certain ordre entr'eux, jamais ils ne le pourroient changer: de sorte que, dés que la premiere voix seroit entenduë, celles qui auroient accoûtumé de la suivre, le seroient necessairement aussi, pourveu que le vent ne manquât pas à la machine; au lieu que les paroles, que j'entens proferer à des corps faits comme le mien, n'ont presque jamais la même suite.

Je connois même [. . .] que l'art peut aller jusqu'à former une machine qui articuleroit des paroles semblables à celles que je prononce. Mais en même temps je conçois qu'elle ne prononceroit que celles qu'on auroit eu dessein qu'elle prononçât, & qu'elle les prononceroit toujours dans le même ordre⁴.

L'idea che un automa possa 'simulare' un corpo vivente è tutt'oggi ben presente nel pensiero scientifico e nelle arti.

In uno fra i primi testi che tratta la storia della musica elettronica, dal significativo titolo di *Musica ex machina*⁵, uscito nel 1960, il primo capitolo è dedicato a "Uomo e macchina". In esso vengono descritti una serie di casi in cui l'uomo e la macchina si confrontano, in cui le conquiste della tecnica sono protagoniste. Partendo da un ambito letterario, alle volte anche di carattere pessimistico (Orwell, Huxley), vi si descrive l'*homunculus* che "trionfa", più per la curiosità dei borghesi che per le sue reali potenzialità scientifiche. Ma se Vaucanson, Jacquet Droz, Siegmeyer vengono ricordati per i loro automi, fra i vari nomi citati spiccano, per l'argomento qui trattato, quelli dell'abate Mical e di Wolfgang von Kempelen, ai quali va sicuramente aggiunto quello di Christian Gottlieb Kratzenstein.

1.2 Christian Gottlieb Kratzenstein (1723 – 1795)

Nel *Tentamen resolvendi problema ab Academia Scientiarum Imperiali Petropolitana ad annum 1780 publice propositu*, Kratzenstein affronta il problema della genesi della voce, in particolare delle vocali. Suddiviso in due parti, la prima "de genesi vocalium" (in cui viene studiata la posizione di laringe, lingua, apertura di palato,

³Si vedano *infra* le sezioni 1.2, 2.2 e 3.8.

⁴Cordemoy, *Discours physique de la parole*, pp. 8-9.

⁵Prieberg, *Musica ex machina*.

denti e labbra) e la seconda “de construendis sistulis, vocales a, e, i, o, u, enun-
 ciantibus” (in cui vengono affrontati i “problemi” relativi alla costruzione di ogni
 singolo “organon” o “sistulam” corrispondente alle singole vocali), il libro illustra la
 costruzione di uno strumento, molto simile ad un organo, che produce i suoni delle
 vocali. L’elemento generatore di suono è un’ancia libera (come nella cornamusa,
 nella zampogna, nella fisarmonica, nell’armonium, nell’armonica a bocca e molto
 simile a quella montata dall’abate Mical nelle sue teste parlanti tre anni dopo), ricca
 di armoniche e che si comporta come la glottide, modello che resterà invariato fino
 all’introduzione dell’elettricità. In particolare, la ricchezza di armoniche dell’ancia
 libera messa in vibrazione da un flusso d’aria, assimilabile a un treno di impulsi e
 quindi molto simile a un *buzz*, la rende una sorgente ideale per un suono che deve
 attraversare una cavità risonante e venir quindi ‘equalizzato’. Anche se sembrerebbe
 evidente che la realizzazione dei tubi che fungono da risuonatori sia avvenuta per
 tentativi⁶, le conoscenze relative all’apparato vocale, all’‘acustica’ e alla fonetica,
 permetteranno a Kratzenstein di costruire una macchina che possiamo chiamare
 il primo sintetizzatore vocale. Inoltre, nell’opera citata, lo scienziato di origine
 tedesca descrive almeno due possibili applicazioni dei suoi studi: “the mechanical
 synthesis of speech as well as suggestings ways to diagnose and correct various
 speech pathologies”⁷; due tematiche, queste, che resteranno a lungo compresenti
 nel dibattito sulla produzione e riproduzione della voce.

Un confronto fra la dimensione delle canne-vocali risonanti costruite da Kratzen-
 stein e le teste parlanti dell’abate Mical⁸ lo troviamo in un giudizio espresso da un
 gruppo di scienziati dell’Accademia delle Scienze di Parigi, che erano a conoscenza
 del fatto che Kratzenstein avesse vinto il premio dell’Accademia di Pietroburgo nel
 1781 sullo stesso soggetto affrontato dall’abate Mical⁹.

1.3 L’abate Mical (1727 – 1789)

Le teste parlanti sono state a lungo oggetto di speculazioni, non sempre di
 carattere scientifico. Papa Silvestro III (ca. 1000 - ca. 1062), Roberto Grossatesta
 (1175 - 1253), Alberto Magno (ca. 1200 - 1280), Ruggero Bacone (1214 - 1294),
 sono fra i nomi che vantano fra le opere del loro ingegno teste parlanti¹⁰.

Antoine de Rivarol, in una lettera datata 20 settembre 1783, ci illustra

⁶Ohala, “Christian Gottlieb Kratzenstein: pioneer in speech synthesis”, p. 158.

⁷Ibid., p. 156.

⁸Si veda 1.3.

⁹Il giudizio lo si può trovare in Académie des sciences (France), “Procès-verbaux. T102 (1783)
 / Académie royale des sciences”. Una trascrizione del manoscritto è presente in Ranostaj, “Unter-
 suchungen zur Sprechtraktakustik”, pp. 107-110.

¹⁰Il capitolo 8 di *Instruments and the Imagination*, al quale rimando, si intitola “*Vox Mechanica: The History of Speaking Machines*”. Il testo cade nell’insidioso equivoco che esiste fra voce prodotta e voce riprodotta. Fra i progenitori dei sistemi di registrazione e riproduzione del suono gli storici spesso ricordano le teste parlanti, in questo non aiutati dal frequente uso della locuzione “talking” o “speaking machines” per indicare grammofoni e fonografi. Sulle teste parlanti si veda inoltre Lindsay, “Talking Head”; in particolare sull’abate Mical si veda Ramsay, “L’Abbé Mical et les Têtes Parlantes:

deux têtes d'airain qui parlent et qui prononcent nettement des phrases entières [...].

En suivant donc la nature pas à pas, ce grand artiste [l'Abbé Mical] s'est aperçu que l'organe vocale était dans la glotte un instrument à vent, qui avait son clavier dans la bouche; qu'en soufflant du dehors au dedans, comme dans un flûte, on n'obtenait que des sons filés; mais que pour articuler des mots, il fallait souffler du dedans au dehors. En effet l'air, en sortant de nos poumons, se change en son dans notre gosier, et ce son est morcelé en syllabes par les lèvres, et par un muscle très-mobile, qui est la langue aidée des dents et du palais. Un *son* continu n'exprimerait qu'une seule affection de l'âme, et se randerait par une seule voyelle; mais coupé à différents intervalles par la langue et les lèvres, il se charge d'une consonne à chaque coup; et se modifiant en une infinité des *tons*, il rend la variété des nos idées.

Sur ce principe, M. Mical applique deux claviers à ses *Têtes parlantes*: l'un en cylindre, par lequel on n'obtient qu'un nombre déterminé de phrases, mais sur lequel les intervalles des mots et leur prosodie sont marqués correctement. L'autre clavier contient, dans l'étendue d'un ravalement, tous les sons et tous les tons de la langue française, réduits à un petit nombre par un méthode ingénieuse et particulière à l'auteur. Avec un peu d'habitude et d'habileté, on parlera avec les doigts comme avec la langue; et on pourra donner au langage des têtes, la rapidité, le repos et toute la physionomie enfin, que peut avoir une langue qui n'est point animée par les passions¹¹.

La descrizione fornitaci da Rivarol di un vero e proprio strumento musicale coglie vari tratti essenziali del modello dell'apparato vocale e, in particolare, la differenza fra suoni vocalici e suoni consonantici, ovvero fra continuità e discontinuità, problema che dovranno affrontare tutti gli autori successivi, fino ad oggi.

Anche se il giudizio espresso dai membri dell'Accademia delle Scienze di Parigi (fra cui spiccano Lavoisier e Laplace) ci tiene a far presente che le frasi ascoltate dalle due teste “ne sont pas prononcées distinctement dans toutes leurs parties [...] et] que le sonds [sic] de la voix produite par cette machine est très different de la voix humaine”, tuttavia poco oltre riconoscono che “malgré ces defauts que M. l'abbé Mical lui même ne se dissesseule pas, le mechanisme de cette machine nous a paru interessant”¹². Le osservazioni dei sette commissari terminano con una nota d'incoraggiamento:

[n]ous en avons dit assez pour faire voier qu'il a toujours cherché à imiter la nature et c'est sans ce rapport que son travail nous a paru si interessant. La chambre à vent fait l'office des poumons; le conduit a veut fait celui de la trachée artère; le trou de la cloison repond à la glotte, la cloison et la [...]me vibrante aux lèvres de la glotte et aux ligamens du Larianx; la cavité de la boîte sonore doit être comparée aux fosses nazales, palatines et Buccales et les differentes ouvertures de la Boîte à celles de la bouche elle-même, comme

L'Histoire de Sa Vie, l'Histoire de Son Oeuvre”.

¹¹Rivarol, *Oeuvres complètes de Rivarol*, pp. 230, 231-232.

¹²Ranostaj, “Untersuchungen zur Sprechtraktakustik”, p. 107.

nous l'avons en parlant de chaque son en particulier. Nous pensons que l'academie doit applaudir aux efforts de M. l'abbé M. que sa machine est ingénieuse, que ses travaux meritent d'être encouragés et que son essai quoi qu'imparfait encor est très digne de l'approbation de l'Academie¹³.

Sebbene in questo resoconto non si menzionino tastiere, non è improbabile che parte del meccanismo che metteva in funzione gli elementi delle teste fosse comandato da tasti.

A conclusione di un libro pubblicato nel 1785, Jean-Baptiste Montmignon descrive le “teste parlanti inventate e costruite dall'abate Mical”¹⁴ e ne ribadisce il loro valore:

Le suffrage de l'Academie des Sciences de la Capitale, rendu sur le rapport de sept Commissaire, dont les noms & la réputation suffiroient seuls pour inspirer la confiance, est sans doute, pour M. l'Abbé Mical, un motif d'encouragement & d'émulation, bien capable de mettre en activité toutes les ressources de son génie, & de l'exciter à redoubler d'ardeur & de constance, pour donner à sa machine parlante, toute la perfection dont elle est susceptible¹⁵.

1.4 Wolfgang Ritter von Kempelen (1734 – 1804)

La letteratura riguardante la produzione ‘sintetica’ della voce¹⁶ riporta in maniera quasi unanime, alle origini del modello meccanicistico, un lavoro di Wolfgang von Kempelen, *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine*, pubblicato nel 1791. In esso viene descritta una macchina, analoga all'apparato vocale e in grado di emettere suoni simili a quelli della voce umana, che von Kempelen terminò di costruire nel 1778¹⁷ e i cui lavori iniziarono almeno nel 1769¹⁸. Tuttavia non è facile determinare se il perfezionamento del ‘sintetizzatore’ vocale fosse un'idea ‘originale’ dell'inventore di origine ungherese

¹³Ibid., p. 110.

¹⁴Montmignon, *Système de prononciation figurée: applicable à toutes les langues et exécuté sur les langues françoises et angloise*, pp. 129-143.

¹⁵Ibid., pp. 131-132.

¹⁶Vorrei segnalare qui soltanto alcuni contributi presenti sul web, molto diversi fra loro nell'impostazione e di carattere divulgativo, che possono introdurre all'argomento: Rubin and Vatikiotis-Bateson, *The Science of the Spoken and Written Word*, i due articoli correlati Scha, *Virtual Voices (1) e Virtual Voices (2)*, Traunmüller, *Wolfgang von Kempelen's speaking machine and its successors* e, sebbene affronti l'argomento dal 1922, anche Maxey, *Smithsonian Speech Synthesis History Project (SSSHP) 1986 - 2002*.

¹⁷Pompino-Marschall, “Von Kempelen et al.”, p. 151

¹⁸Dudley and Tarnoczy, “The Speaking Machine of Wolfgang von Kempelen”, p. 157, ricco di riferimenti anche storici. Possiamo vedere una riproduzione e una dimostrazione del funzionamento della macchina parlante di von Kempelen, costruita da Fabian Brackhane, Dominik Bauer e dall'organaro Hugo Mayer, in <<https://www.youtube.com/watch?v=zYRVqrfY3tQ>>. Per un contributo su di essa si può consultare inoltre *ad vocem* <https://en.wikipedia.org/wiki/Wolfgang_von_Kempelen's_Speaking_Machine> [ultima consultazione 11 aprile 2013].

(von Kempelen nacque a Pressburg in Ungheria ed è ricordato soprattutto per il suo giocatore automatico di scacchi) oppure se si fosse ispirato alle teste parlanti dell'abate Mical. Antoine de Rivarol infatti scrive nel 1783:

Je n'oublierai pas non plus M. Kemplein, voyageur allemand, qui nous à montré l'hiver dernier un'automate jouant aux échecs [...]. M. Kemplein avait aussi un coffret d'où il s'échappait quelques mots, à ce qu'on dit: mais cet honnête voyageur a rendu un véritable hommage à M. l'abbé Mical; dès qu'il a eu connaissance des têtes-parlantes, il a retiré son automate, son coffret et sa propre personne¹⁹.

Il modello meccanicistico è fin da subito evidente; basti osservare chi fornisce l'energia all'apparato vocale (due polmoni-mantici), la conformazione della glottide (una membrana elastica separata in due) e la lingua, una tavoletta sollevabile e abbassabile manualmente²⁰.

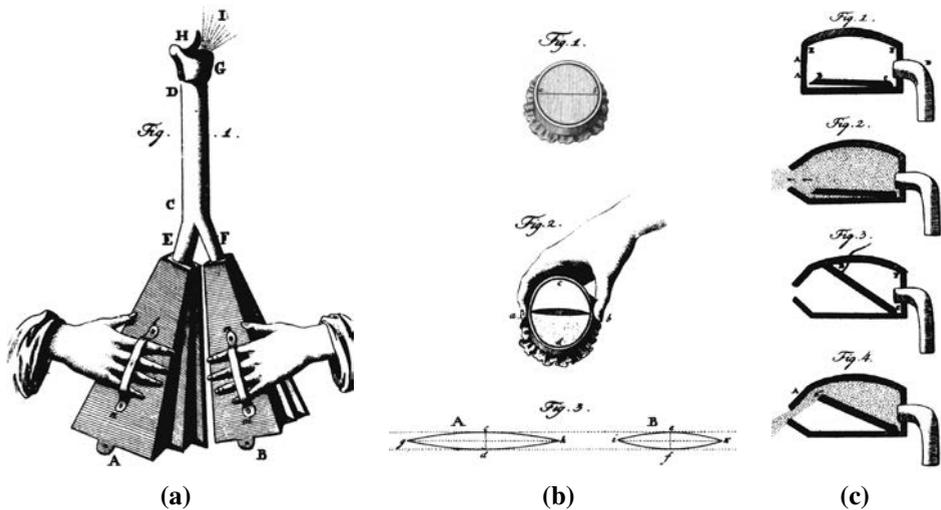


Fig. 1: Tre tavole del trattato *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine*, in cui si vedono (a) i polmoni-mantici, (b) la glottide-membrana e (c) la lingua-tavoletta.

Lo studio attento, analitico, dei meccanismi di fonazione sarà la chiave di lettura per comprendere la macchina di von Kempelen²¹. Sebbene le osservazioni sul processo naturale del parlare non siano completamente applicate alla macchina, si può

¹⁹Rivarol, *Oeuvres complètes de Rivarol*, p. 243.

²⁰Si vedano le immagini nella Fig. 1 tratte da Kempelen, *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine*, p. 76. Nella realizzazione della macchina Kempelen, tuttavia, opera alcune semplificazioni rispetto al modello da lui ipotizzato, sostituendo ad esempio con un'ancia libera – come abbiamo visto fare in tutti gli altri casi di macchine parlanti – la molto più elaborata e corrispondente all'originale membrana che vediamo in figura.

²¹A tal proposito si veda il paragrafo 3 di Pompino-Marschall, “Von Kempelen et al.”, p. 151.

affermare che proprio dalla dettagliata descrizione della posizione e articolazione di bocca, lingua, denti, labbra, abbia inizio la progettazione della macchina parlante. In maniera analoga al suo predecessore Kratzenstein, “Kempelen didn’t construct his speaking machine on the base of acoustic theories but went the engineering way of analysis-by-synthesis – or trial and error. He was mainly interested in the audible result that should be reached by a simple mechanism as close as possible to our articulatory apparatus on the one hand and playable like a musical instrument on the other”²². Dunque già le prime macchine parlanti avevano implicitamente un’interesse nei confronti dell’aspetto musicale della loro capacità produttiva.

1.5 Molti altri

“L’illustre fisico inglese” Charles Wheatstone (1802 – 1875)²³, inventore e musicista, realizzò un modello avanzato della macchina parlante di von Kempelen e inventò e costruì svariati strumenti musicali. Fra questi ricordo la *concertina*, brevettata il 19 dicembre 1829²⁴, e l’*harmonium* portatile. Entrambi questi strumenti utilizzano un’ancia libera, tipico eccitatore presente nei ‘sintetizzatori’ vocali pre-elettrici.

Molti altri sono i nomi di studiosi che realizzano sistemi per simulare la produzione della voce: Robert Willis (1800 – 1875)²⁵, Rudolph Koenig (1832 – 1901), Carl Stumpf (1848 – 1936)²⁶, William Henry Preece (1834 – 1913) e J. M. Augustus Stroh (1828 – 1914)²⁷, Dayton Clarence Miller (1866 – 1941).

Fra tanti va sicuramente ricordato Hermann von Helmholtz (1821 – 1894), il quale, durante i suoi studi sulla risonanza, aveva ideato uno strumento per produrre i suoni delle vocali. Basato su diapason intonati e controllato da una tastiera, il *Vocal-Apparates*²⁸ riproduceva le “note” corrispondenti alle vocali.

Voglio qui menzionare infine altri due casi riportati in letteratura. Del primo caso, il pluricitato *Euphonia* (1835), strumento a tastiera di Joseph Faber che imita la voce umana controllato da un ‘musicista’, non sono stati rinvenuti documenti

²²Pompino-Marschall, “Von Kempelen et al.”, p. 151.

²³Così Rudolph Koenig lo definisce nel suo libro *Quelques Expériences d’Acoustique*, p. 32. Notizie su Wheatstone si possono trovare in Bowers, *Sir Charles Wheatstone FRS. 1802-1875*. Su Koenig si può consultare Pantalony, *Altered sensations: Rudolph Koenig’s acoustical workshop in nineteenth-century Paris*.

²⁴Sebbene nel brevetto di Wheatstone, “Wind Musical Instruments”, non compaia il termine concertina, in esso si trova la prima descrizione dello strumento, simile ad una fisarmonica.

²⁵Willis, “On vowel sounds, and on reed-organ pipes”.

²⁶Stumpf, *Die Sprachlaute*.

²⁷Preece e Stroh pubblicano uno studio in cui presentano una sorta di campionatore *ante-litteram* che riproduce i suoni delle vocali registrati su un cilindro: Preece and Stroh, “Studies in Acoustics. I. On the Synthetic Examination of Vowel Sounds”.

²⁸Una dettagliata descrizione del suo funzionamento la possiamo trovare in Pisko, *Die neueren apparate der akustik: Für freunde der naturwissenschaft und der tonkunst*, pp. 17-35, testo dal quale è tratta anche la Fig. 2. Una ricostruzione moderna corredata da notizie storiche si può vedere ed ascoltare in Rees, *Helmholtz’s apparatus for the synthesis of sound: an electrical ‘talking machine’*.

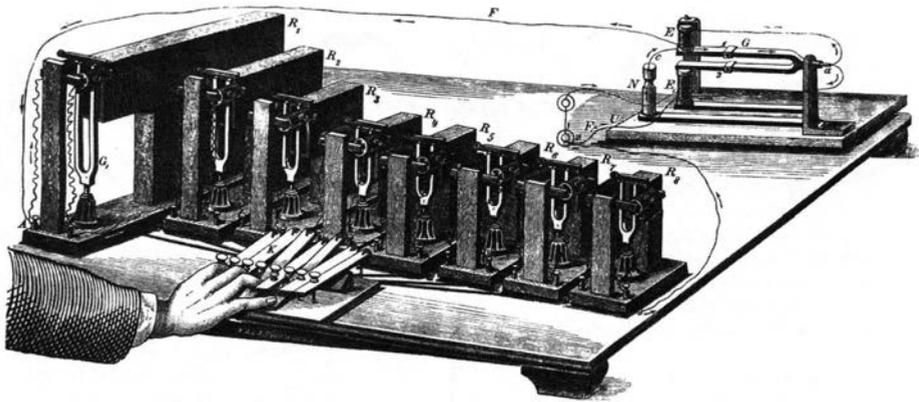


Fig. 2: Il sintetizzatore vocale di Helmholtz.

diretti ma solo testimonianze indirette; celebri quelle di Joseph Henry²⁹, che trovò lo strumento superiore a quello di Wheatstone, e di Melville Bell, che portò il figlio Alexander Graham a vederlo. Le testimonianze sono dovute in gran parte alle esibizioni che la macchina fece grazie a Phineas Taylor Barnum, che gli diede anche il nome con il quale è conosciuta. Al contrario dell'*Euphonia*, di un sintetizzatore mai costruito, il *Tecnefón* di Severino Pérez y Vázquez, vi è però un testo del 1868 che ne descrive il funzionamento³⁰.

2 L'analogia elettrica

2.1 Gli anni venti del XX secolo

Nel 1922 venne pubblicato sulla rivista *Nature* un articolo che illustra l'analogia del tratto vocale con un modello elettrico³¹. Non appena l'elettroacustica si delineò come disciplina, le prime ricerche a trattare le analogie elettro-fisiologiche riguardarono proprio la voce e l'udito. Lo stesso Graham Bell e i ricercatori che operavano nei Bell Telephone Laboratories furono al centro di questi studi³². Nel febbraio 1924 Harvey Fletcher illustrò un apparato simile a quello di Stewart³³. Finalizzata alla compressione del segnale da inviare attraverso il cavo telefonico

²⁹Si veda Millikan, *Joseph Henry and the Telephone*.

³⁰Vázquez, *El Tecnefón, y los medios representativos de la palabra*. Si veda inoltre Moro, "A 19th-Century Speaking Machine: The Tecnefón of Severino Perez y Vazquez".

³¹Stewart, "An Electrical Analogue of the Vocal Organs".

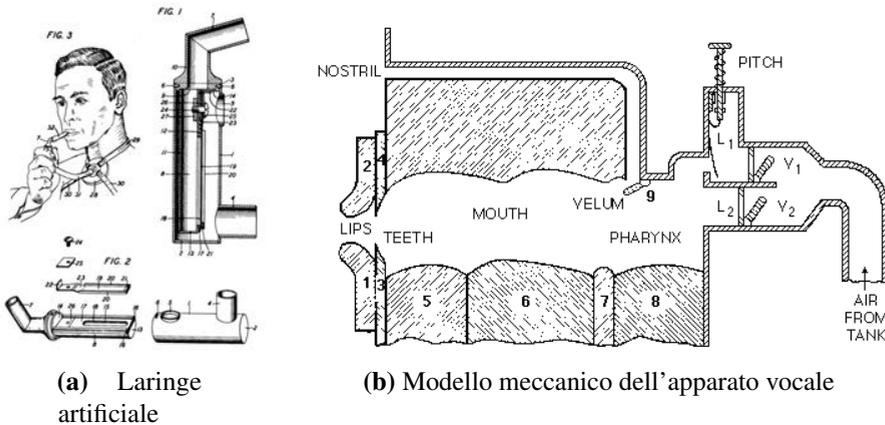
³²Per informazioni sui Bell Laboratories e sull'American Telephone & Telegraph rimando all'opera in più volumi a cura di Fagen et al., *A History of Engineering and Science in the Bell System*.

³³Dudley, Riesz, and Watkins, "A Synthetic Speaker", p. 742.

o via etere, la ricerca sulla voce comprenderà studi di psicoacustica, di teoria dei segnali e di teoria dell'informazione e della comunicazione.

2.2 Robert R. Riesz

Ma sarà il lavoro di Robert Riesz il vero *trait-d'union* fra analogia meccanica e analogia elettrica per la realizzazione del vocoder. I suoi studi sulla laringe artificiale degli anni venti dimostrano una conoscenza dettagliata del comportamento delle 'corde vocali'. Da questi studi nasceranno varie invenzioni ad uso dei laringectomizzati³⁴: possiamo vederne un esempio nella Fig. 3 (a), dove l'elemento che sostituisce la laringe è un'ancia libera, pochi anni dopo sostituita da una laringe elettrica³⁵. Nel 1937 Riesz costruisce un modello meccanico completo dell'apparato vocale³⁶) e contemporaneamente si dedica alle analogie elettriche dell'apparato fonatorio e alla produzione e trasmissione dei segnali per via elettrica.



(a) Laringe artificiale

(b) Modello meccanico dell'apparato vocale

Fig. 3: Due immagini di modelli meccanici realizzati da Robert Riesz.

In quanto coautore dell'articolo "A Synthetic Speaker", possiamo supporre che il suo apporto in merito al modello meccanico dell'apparato vocale sia stato significativo. Il suo contributo allo sviluppo del vocoder è inoltre testimoniato dalla progettazione sia del generatore di *buzz* sia del generatore di *hiss*, le due principali sorgenti per la sintesi del suono in questo strumento³⁷.

³⁴Fra cui quella descritta in Riesz, "Artificial larynx", da cui è tratta anche la Fig. 3 (a); si veda inoltre Riesz, "Description and Demonstration of an Artificial Larynx".

³⁵Riesz, "Electrical Artificial larynx".

³⁶Si veda la Fig. 3 (b), tratta da Flanagan, *Speech Analysis Synthesis and Perception*, p. 169. Non vi sono altre fonti di mia conoscenza che parlino di questo sistema meccanico. Visto il contatto diretto che Flanagan aveva con Riesz la si può considerare attendibile. Poco oltre, infatti, nello stesso testo Flanagan riporta le seguenti parole *de auditu* da Riesz: "when operated by a skilled person, the machine could be made to simulate connected speech. One of its particularly good utterances was reported to be 'cigarette'".

³⁷"Mr. R. R. Riesz is responsible in large part for the development of the gas-tube energy sources

3 Il Vocoder

3.1 Nascita

In October, 1928, Homer Dudley of Bell Telephone Laboratories sketched in his technical notebook a device which subsequently became known as “vocoder”³⁸.

L’idea di un sistema che, a seguito di un’analisi del segnale, operi una codifica e riduca la quantità di dati per trasmetterli attraverso un canale con una banda limitata per poi decodificarli (sintesi per analisi) sorge prima di quella relativa alla sua produzione *ex-novo* (sintesi). Come ci ricorda Dudley, “of these two [vocoder e voder] speech synthesizers the vocoder was constructed first. It works on the principle of automatically remaking speech under control of spoken speech instantaneously analyzed to derive the code currents for the control [...]. The voder was derived from the vocoder by substituting manipulative for automatic controls”³⁹. Per quanto qui sostenuto, è la diversa natura di questi due aspetti che differenzia sostanzialmente sistemi come il vocoder e il voder, utilizzi come quello nell’ambito della telefonia e nell’ambito musicale. Il momento cruciale del *controllo*, che in un caso (vocoder, telefonia) avviene automaticamente e nell’altro (voder, musica) necessita dell’intervento umano, permetterà a questa tecnologia di diventare paradigmatica di un’epoca e di essere oggetto di continue ricerche e approfondimenti. Nella Fig. 4 si può vedere chiaramente come i due sistemi si differenzino nella parte iniziale, in quella porzione del flusso relativa al “messaggio”, che in questo caso equivale a controllo: nel vocoder, *la macchina ricava automaticamente* i valori di controllo per il sintetizzatore, nel voder *un operatore controlla direttamente* il sintetizzatore⁴⁰.

3.2 L’acronimo

Il termine *Vocoder* ha tre caratteristiche principali: la forma di acronimo, il riferimento alla voce e il riferimento a un sistema di codifica. Il significato dell’acronimo è infatti VOICE CODER. Tuttavia lo possiamo trovare spiegato in almeno altre due maniere: ancora nell’edizione ampliata e rivista di *Electronic & Computer Music*, Manning lo definisce “Voice Operated reCORDER”⁴¹, mentre nel web, la

for the buzz and hiss to be described later” nell’articolo di Dudley, “Remaking Speech”, p. 172.

³⁸In Schroeder, “Vocoders: Analysis and Synthesis of Speech”, p. 720. La rassegna più completa della letteratura sui sistemi automatici di riconoscimento della voce fino al 1966 è forse Kyunnap, “Speech Commands in Control Systems”, che menziona 190 titoli da tutto il mondo. Val forse la pena osservare che questo testo – redatto da uno scienziato russo – si trova tradotto in inglese in un documento interno della NASA, a riprova dell’interesse che i militari avevano per le ricerche in questo campo, come vedremo in 3.7.

³⁹Dudley, “The Carrier Nature of Speech”, p. 505.

⁴⁰Le immagini sono tratte da Dudley, “The Carrier Nature of Speech”, pp. 508-509.

⁴¹Manning, *Electronic & Computer Music*, p. 39.

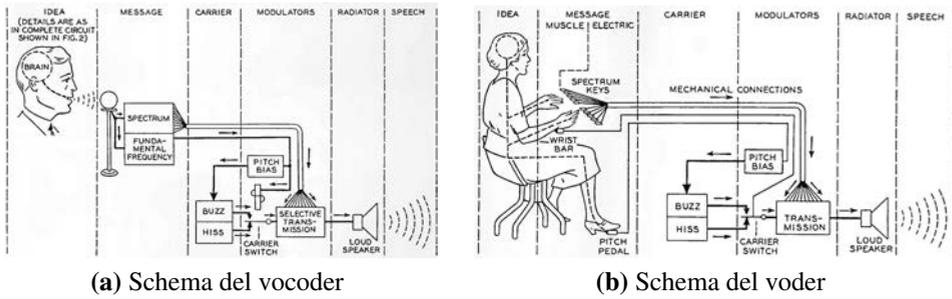


Fig. 4: Confronto fra gli schemi di vocoder e voder.

voce *Vocoder* sia in *wikipedia* sia in *wordreference* ci viene spiegata come acronimo di VOice enCODER⁴². Che il termine abbia subito varie ‘interpretazioni’ è comprensibile, vista la sua grande diffusione. Come ci ricorda Flanagan,

the original Vocoder – now referred to as a spectrum channel vocoder – has probably been described in the literature more times than any other single system⁴³.

Una conferma di questa proliferazione ci viene ancora oggi: cercando il termine “vocoders” nella «ieeexplore digital library» si ottengono 1410 risultati, mentre il termine “vocoder” ne ottiene 625; nel sito della «Audio Engineering Society» si ottengono 149 risultati per vocoder e 54 per vocoders; nel motore di ricerca di «jstor» 358 per vocoder e 66 per vocoders⁴⁴.

3.3 Un banco di filtri

La descrizione del vocoder, come abbiamo appena visto, la si può agevolmente trovare in letteratura⁴⁵. Voglio qui rilevare solo un aspetto, che influisce molto sulla qualità del suono generato dal vocoder e sulla sua modellizzazione: il numero di canali utilizzati in fase di analisi e di risentesi.

Il sistema per la trasmissione del segnale proposto da Dudley⁴⁶ era dotato di 10 canali (10 filtri passabanda) di trasmissione fra 250 Hz e 7100 Hz con larghezza di banda variabile, che dovevano adattarsi a trasportare non solo la voce (si veda la Tabella 3.2). Il *channel vocoder* descritto in Dudley, “The Vocoder”, p. 123,

⁴²Si vedano <<https://en.wikipedia.org/wiki/Vocoder>> e <<http://www.wordreference.com/definition/vocoder>> [short for voice encoder]. Una possibile fonte di questa ‘lettura’ potrebbe essere Moore, *Elements of Computer Music*, p. 227.

⁴³Flanagan, *Speech Analysis Synthesis and Perception*.

⁴⁴Ricerche effettuate in <<https://ieeexplore.ieee.org/Xplore/guesthome.jsp>>, <aes.org> e <www.jstor.org> rispettivamente il 19 febbraio, il 15 marzo e il 6 aprile 2013.

⁴⁵Rimando alla dettagliata ed esaustiva spiegazione che ci fornisce proprio Flanagan in *Speech Analysis Synthesis and Perception* poco dopo aver scritto quanto citato sopra.

⁴⁶Si veda 3.4.

prevedeva 10 filtri passabanda che arrivavano fino a 2950 Hz con larghezza di banda fissa di 300 Hz (tranne il primo, con larghezza di banda di 250 Hz); nello stesso articolo Dudley accenna anche a un “30-channel vocoder covering the wide range of speech frequencies required for high quality”⁴⁷; Schroeder, parlando di quello che propriamente viene inserito fra gli “*spectrum-channel vocoder*”, lo descrive formato da 14 bande spettrali, da 200 Hz a 3200 Hz, con larghezze di banda variabili (200-300, 300-450 – 2800-3200 Hz), coprendo così il tipico segnale telefonico e permettendo alta intelligibilità e alta qualità⁴⁸. La diversità di proposte dipende in parte dal tipo di segnale da trasportare e in parte dall’ancora incerto stato della ricerca. Ricordo che l’attività sia teorica sia sperimentale attorno al vocoder nasce soprattutto dalla necessità di trasmettere su un’unica linea telefonica più segnali diversi, ricorrendo alla modulazione di una portante.

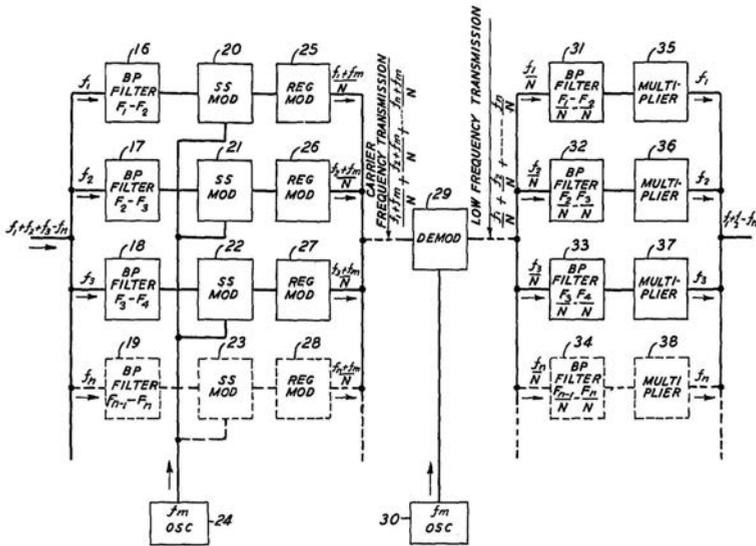


Fig. 5: Immagine tratta dal brevetto americano 2,117,739 depositato il 5 giugno 1936 da Ralph Miller per conto dei Bell Laboratories.

Un precedente spesso riferito è un “signaling system” brevettato da Ralph Miller⁴⁹ per conto dei Bell Laboratories:

[t]his invention relates to systems for the transmission of signals and more particularly to systems in which the signals are transmitted in the form of a modulated carrier wave.

An object of the invention is to effectively reduce the width of the frequency band required for the transmission of a given signal wave.

⁴⁷Dudley, “The Vocoder”, p. 123.

⁴⁸Schroeder, “Vocoders: Analysis and Synthesis of Speech”, p. 723.

⁴⁹Miller, “Signaling System”.

A further object of the invention is to produce a signal-modulated carrier wave in which the frequency of the carrier wave varies in accordance with the predominant frequency of the original signal⁵⁰.

Sebbene non sia specificato il numero di filtri passabanda impiegati (nella Fig. 5 ne vediamo un numero n generico), maggiore esso è migliore sarà la qualità del segnale modulato (a spese, ovviamente, di una maggiore banda passante necessaria)

In Dudley, “Remaking Speech”, lo spettro utile, limitato ai “circuiti del telefono commerciale”, è suddiviso come in Tabella 3.1⁵¹.

Tabella 3.1: Numero di bande e relative frequenze utilizzate da Dudley.

No. banda	Frequenze Hz	ΔF	No. banda	Frequenze Hz	ΔF
1	0 - 250	250	1	0 - 150	150
2	250 - 550	300	2	150 - 300	150
3	550 - 850	300	3	300 - 450	150
4	850 - 1150	300	4	450 - 600	150
5	1150 - 1450	300	5	600 - 750	150
6	1450 - 1750	300	6	750 - 900	150
7	1750 - 2050	300	7	900 - 1050	150
8	2050 - 2350	300	8	1050 - 1200	150
9	2350 - 2650	300	9	1200 - 1350	150
10	2650 - 2950	300	10	1350 - 1500	150
			11	1500 - 1650	150
			12	1650 - 1800	150
			13	1800 - 1950	150
			14	1950 - 2100	150
			15	2100 - 2250	150
			16	2250 - 2400	150
			17	2400 - 2550	150
			18	2550 - 2700	150
			19	2700 - 2850	150
			20	2850 - 3000	150
			21	3000 - 3160	160
			22	3160 - 3340	180
			23	3340 - 3550	210
			24	3550 - 3800	250
			25	3800 - 4110	310
			26	4110 - 4500	390
			27	4500 - 5000	500
			28	5000 - 5640	640
			29	5640 - 6460	820
			30	6460 - 7500	1040

⁵⁰Ibid., p. 1.

⁵¹I valori di sinistra della tabella si riferiscono anche a “The Vocoder”, dicembre 1939 e “Fundamentals of Speech Synthesis”, ottobre 1955. A destra vediamo invece i valori utilizzati in un primo prototipo (dalla banda 21 alla banda 30 i valori sono stati ipotizzati, poiché in Dudley, “Fundamentals of Speech Synthesis”, p. 175, viene solo affermato che l’andamento “extending logarithmically from 3000 to 7500 cps”).

3.4 Homer W. Dudley

Intorno alla seconda metà degli anni venti del secolo scorso, i problemi di degradazione del segnale nei cavi transatlantici e di compressione dei dati erano al centro delle ricerche delle compagnie telefoniche. Dalla fine degli stessi anni, Homer Dudley, dipendente dei Bell Telephone Laboratories Inc., deposita, per conto del suo datore di lavoro, una serie di brevetti relativi alla trasmissione dei segnali e in particolare della voce. Il 30 ottobre 1935 è registrato a suo nome presso l'ufficio brevetti degli Stati Uniti un sistema per la trasmissione del segnale⁵². Questo sistema, che sarà la base per la realizzazione sia del *vocoder* sia del *voder*, è citato da gran parte dei brevetti successivi che si occupano di analisi e sintesi della voce.

Nel "Signal Transmission" appena citato viene illustrato un sistema che codifica ('comprime') un segnale da una banda di 250-7100 Hz a una banda di 70-360 Hz (si veda la Fig. 6), segnale che viene successivamente decodificato ('decompresso') in maniera inversa (si veda la Fig. 7). Infatti "[t]he information transmitted by speech does not absolutely require all the frequency space allotted to it in the human voice."⁵³ Il brevetto contiene già tutti gli elementi del vocoder, dal sistema di analisi a quello di sintesi. Nel settembre 1936, durante il tricentenario della Harvard University, avviene la prima dimostazione del Vocoder⁵⁴.

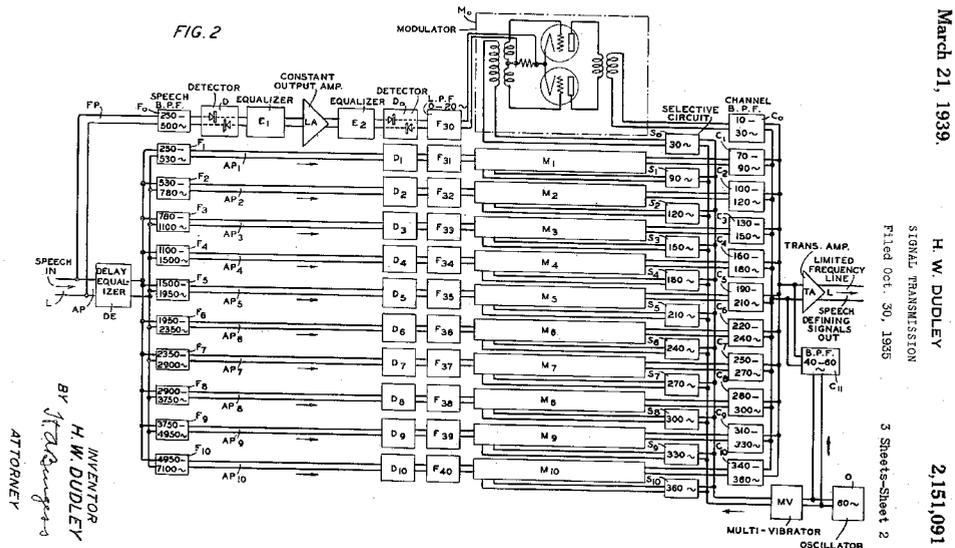


Fig. 6: Codifica del segnale nel brevetto americano 2,151,091 di Dudley.

⁵²Dudley, "Signal Transmission".

⁵³Ibid., 1:37-39.

⁵⁴Si vedano Dudley, "Synthesizing Speech" e Dudley, "Remaking Speech".

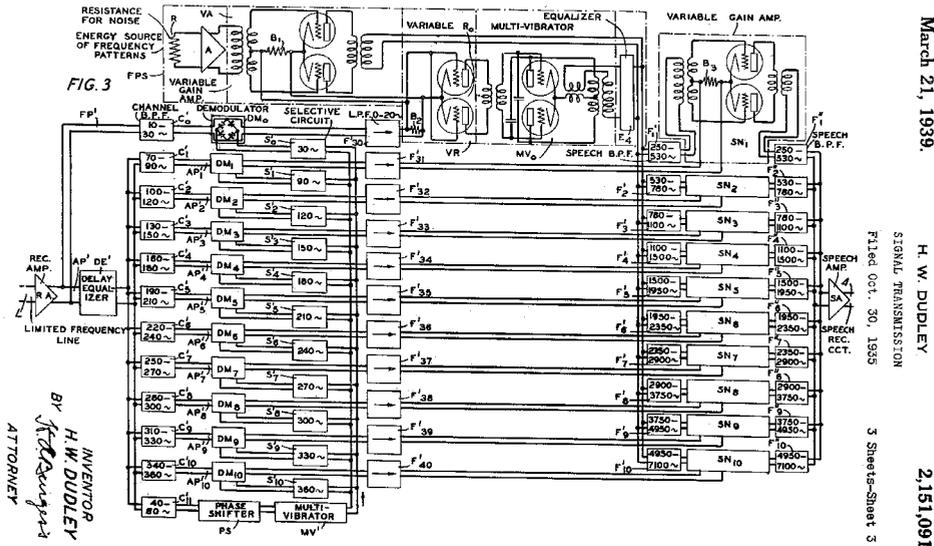


Fig. 7: Decodifica del segnale nel brevetto americano 2,151,091 di Dudley.

Tabella 3.2: Confronto fra il numero di bande e relative frequenze del “Signal Transmission” (brevetto US 2,151,091 depositato il 30 ottobre 1935), a sinistra, e quelle del “System for the artificial production of vocal or other sounds” (brevetto US 2,121,142 depositato il 7 aprile 1937), a destra.

No. banda	Frequenze Hz	ΔF	No. banda	Frequenze Hz	ΔF
1	250 - 530	280	1	0 - 225	225
2	530 - 780	250	2	225 - 450	225
3	780 - 1100	320	3	450 - 700	250
4	1100 - 1500	400	4	700 - 1000	300
5	1500 - 1950	450	5	1000 - 1400	400
6	1950 - 2350	400	6	1400 - 2000	600
7	2350 - 2900	550	7	2000 - 2700	700
8	2900 - 3750	850	8	2700 - 3800	1100
9	3750 - 4950	1200	9	3800 - 5400	1600
10	4950 - 7100	2150	10	5400 - 7500	2100

I risultati delle ricerche sulla codifica e la trasmissione dei segnali svolte da Dudley negli anni trenta verranno riassunti in “The Carrier Nature of Speech”, pubblicato nell’ottobre del 1940. In esso Dudley formalizza la teoria che fonda il sistema di trasmissione dei segnali basandosi sulle caratteristiche di “portante” della voce. Il modello della voce è un modello ‘universale’ di segnale ‘armonico’. Lo

studio del suo comportamento permette di realizzare una macchina che, in base ai dati ricevuti in ingresso e ricavati dalla voce stessa, ne replichi il risultato in uscita; permette quindi di costruire un sintetizzatore in grado di trasportare qualsiasi segnale, non solo quello vocale. Se nel sistema vocale l'onda portante è fornita dalla vibrazione della glottide e la modulazione da tutti i movimenti muscolari che ne seguono (gola, lingua, bocca, ecc.), negli strumenti musicali, scrive Dudley,

each note has a fundamental and upper harmonics of a resonant system formed for example from a stretched string, vibrating reed or air tube or chamber, and these fundamental and harmonic components of the notes form the set of high frequencies in the complex signal wave, the low frequency set corresponding to the rates of energizing such resonant systems, as for example by the hand striking the piano keys. Each key is an independently variable element in this case if one considers the keyboard as the starting point. If the hands are considered as the starting point, then the maximum number of fingers used at any one instant are the independently variable parameters. The rest of the piano mechanism is of a fixed nature⁵⁵.

Il messaggio, infatti, pur essendo un segnale complesso, “may be analyzed as the sum of component sine waves each of which is characterized by its own amplitude, frequency, and phase”⁵⁶, e ha una frequenza – compresa in un intervallo di 0-10 Hz⁵⁷ – inferiore alla frequenza della portante.

Articolazione della voce, movimento delle dita sulla tastiera o sulle chiavi di uno strumento a fiato, costituiscono il controllo del suono, originato da un'altra parte dello strumento.

3.5 Il Voder

Nell'articolo apparso nel giugno del 1939 sul «Journal of The Franklin Institute», Homer Dudley, Robert Riesz e S. S. A. Watkins dei Bell Telephone Laboratories rendono conto di una ricerca iniziata alcuni anni prima. L'articolo, intitolato “A Synthetic Speaker”, si riferisce ad un sistema per la sintesi della voce: il Voder, acronimo di Voice Operation DEMonstratoR, nato con l'intento di ridurre i dati nelle comunicazioni telefoniche e basato su una serie di analogie relative all'apparato fonatorio (bio-meccaniche, bio-elettriche, elettro-meccaniche)⁵⁸. In questo articolo viene anche chiarito un aspetto determinante degli studi sulla voce allora in corso ai Bell Laboratories; l'11 settembre 1936, uno degli autori presenta infatti “a demonstration of speech which had been analyzed electrically and reconstructed or

⁵⁵Dudley, “Signal Transmission”, 3:49-66.

⁵⁶Dudley, “The Carrier Nature of Speech”, p. 499.

⁵⁷Si veda Dudley, “Fundamentals of Speech Synthesis”.

⁵⁸Anticipato col brevetto americano n. 2,121,142, “System for the artificial production of vocal or other sounds”, depositato il 7 aprile 1937 e in seguito perfezionato col brevetto americano n. 2,194,298 con lo stesso titolo depositato il 23 dicembre 1937.

synthesized from electrical sources of energy. In that system the energy for the synthesized speech was *directly controlled* by the speech of the original talker [. . .]”⁵⁹. Poco oltre gli autori ribadiscono che “the basic distinction between the Voder and the apparatus demonstrated then is that with the Voder an operator controls the production of the synthetic speech purely by mechanical manipulation.” Il controllo è dunque la chiave interpretativa per comprendere l’uso di questi sintetizzatori, in particolare nell’uso intensivo che ne verrà fatto in seguito in musica.

Il principio su cui si basa per la risintesi semplifica il meccanismo di produzione della voce suddividendolo in due modalità⁶⁰: l’una, vocalica, periodica (‘intonata’), in cui la sorgente è un treno di impulsi (“buzzer”) emesso dalle vibrazioni della glottide, l’altra, consonantica (rumorosa, “random noise”), in cui la sorgente è un rumore.

Il Voder poteva essere controllato da un operatore, che agiva su una doppia tastiera e un pedale⁶¹, ai quali erano associati i vari caratteri del tratto vocale e del meccanismo di produzione della voce e che modificavano l’andamento della frequenza fondamentale, l’alternarsi di suoni vocalici e consonantici e le caratteristiche formantiche del suono emesso, agendo su un banco di filtri. Il modello è molto semplice (si veda la Fig. 8 (a), tratta da *ibid.*, p. 747), ma un operatore impiegava almeno un anno per far ‘parlare’ il sintetizzatore⁶². La straordinaria somiglianza con lo strumento descritto da Antoine de Rivarol (due tastiere⁶³), rendono le teste parlanti dell’abate Mical un precursore del Voder.

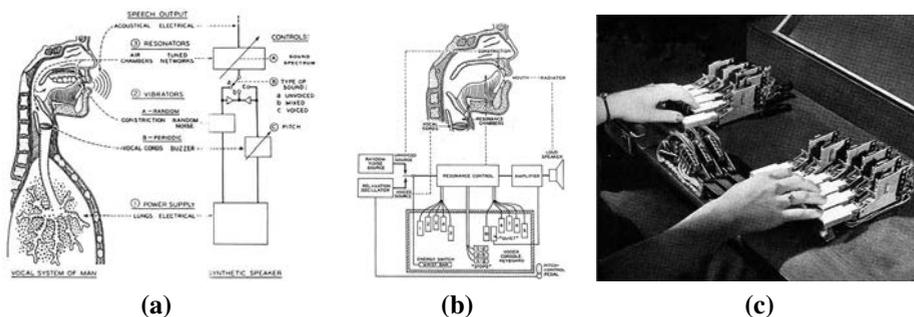


Fig. 8: Voder: (a) analogia fra sistema vocale e sintetizzatore, (b) parti essenziali del Voder e analogia tra azione svolta dall’operatore al Voder e apparato fonatorio, (c) azione svolta dalle mani dell’operatore.

⁵⁹Dudley, Riesz, and Watkins, “A Synthetic Speaker”, p. 740, corsivo mio.

⁶⁰Come si può vedere nella Fig. 8 (a), tratta da *ibid.*, p. 747.

⁶¹Si vedano le immagini (b) e (c) in Fig. 8, tratte da *ibid.*, p. 748 la (b) e da <<http://davidszondy.com/future/robot/voder.htm>> [ultima consultazione 11 aprile 2013] la (c).

⁶²“Generally speaking it seems to take about a year to develop a good technique”. In *ibid.*, p. 763.

⁶³Si veda 1.3 e le immagini (b) e (c) in Fig. 8, tratte da *ibid.*, p. 748.

3.6 Raffigurare lo spettro di un suono

Parallelamente allo studio analitico/uditivo dello spettro del suono si sviluppano anche le tecniche per la sua visualizzazione o raffigurazione⁶⁴; un primo esempio realizzato con strumenti ‘scientifici’ è stato prodotto da Koenig tramite la fiamma manometrica nella seconda metà dell’Ottocento (si veda la Fig. 9⁶⁵). Con quest’ultima, il fisico tedesco noto per la costruzione di strumenti di misura aveva realizzato un vero e proprio spettrografo, basandosi sui principi enunciati da Helmholtz.

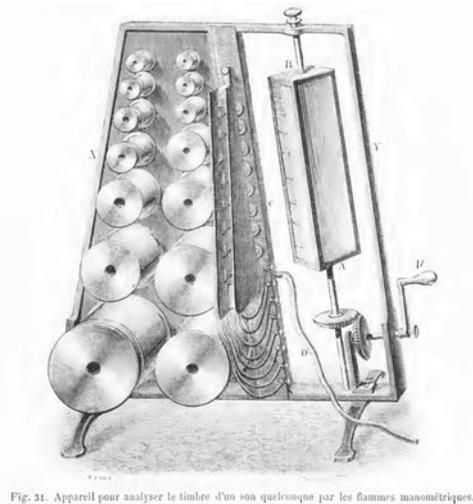


Fig. 31. Apparil pour analyser le timbre d'un son quelconque par les flammes manométriques.

(a) ‘Spettrografo’

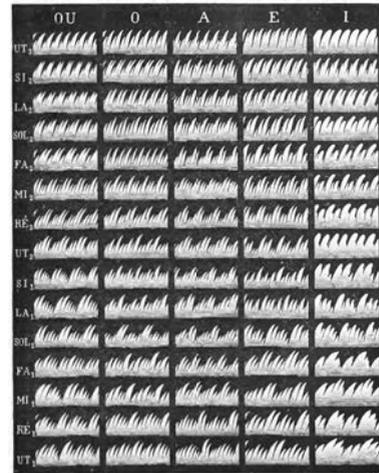


Fig. 37. Vapelles chaotiques sur différents notes, représentées par des flammes manométriques.

(b) Vocali

Fig. 9: Il visualizzatore di spettro e le vocali raffigurate osservando la fiamma manometrica di Koenig.

Sistemi meccanici per la rappresentazione di segnali armonici nel dominio della frequenza sono descritti in Miller, *The Science of Musical Sounds*, e sono quelli, fra tanti, di Lord Kelvin (1876), Michelson e Stratton (1898), Henrici (1894) e in seguito quello di Backhaus (1927)⁶⁶.

The development of the sound spectrograph represented a significant contribution to speech analysis, since it displays the speech events in a way that brings into clear focus certain of the essential features of the signal such as the formant movements. The three-dimensional intensity-frequency-time representation, together with the procedure for displaying spectral sections,

⁶⁴Per un approfondimento sui sistemi di visualizzazione e raffigurazione dei suoni rimando a Zavagna, “Trascrivere documenti sonori”.

⁶⁵Koenig, *Quelques Expériences d'Acoustique*, pp. 63, 73.

⁶⁶Si vedano Thomson, “Harmonic Analyzer”; Michelson and Stratton, “A New Harmonic Analyzer”; Henrici, “On a New Harmonic Analyzer”; Backhaus, “Über Geigenklänge” e “Über die Bedeutung der Ausgleichsvorgänge in der Akustik”.

provides a means for isolating significant features for certain classes of sound, although the techniques are less successful for other classes, particularly certain types of consonants⁶⁷.

Questa dichiarazione di utilità descrive meglio di qualsiasi altra premessa quanto le forme di visualizzazione del segnale nel dominio della frequenza abbiano contribuito allo studio del suono e siano state sviluppate per finalità analitiche. Gli autori dell'articolo appena citato si riferiscono ad uno scritto del 1946, che può essere considerato fra i primi contributi moderni in materia⁶⁸.

Quest'ultimo è inserito nel volume 18, n. 1, del luglio 1946, del «Journal of Acoustical Society of America», quasi interamente dedicato alla visualizzazione⁶⁹ del suono nel dominio della frequenza. In esso sono presenti i seguenti contributi: Dudley and Gruenz, “Visible Speech Translators with External Phosphors”; Potter, “Introduction to Technical Discussions of Sound Portrayal”; Riesz and Schott, “Visible Speech Cathode-Ray Translator”; Steinberg and French, “The Portrayal of Visible Speech”. Le ricerche sulla “visible speech”⁷⁰ erano orientate alla possibilità di fornire ai sordi un sistema alternativo per comunicare. Si riteneva infatti che dalla visualizzazione spettrale del parlato – dopo un opportuno allenamento – si potesse risalire al testo⁷¹.

Dopo una breve rassegna dei metodi di rappresentazione precedenti il 1946⁷² gli autori presentano un sistema che prevede la registrazione di un segmento di suono di 2.4 secondi su un nastro magnetico montato su un disco rotante, che dia la possibilità di leggere ripetutamente, variando di volta volta manualmente la frequenza del filtro analizzatore, il contenuto del frammento sonoro; inoltre, per poter meglio valutare l'energia su tutto lo spettro dell'udibile, viene introdotta, laddove necessario, un'equalizzazione che aumenta di 6 dB le alte frequenze. Nella fase di analisi viene utilizzato un sistema ad eterodina, l'analizzatore ERPI, già presente nei laboratori di allora. La dinamica del sistema era di circa 40 dB, visibili nella variazione di densità dei grigi della carta sensibile alla luce, che tuttavia copriva una dinamica di soli 12 dB, necessitando quindi una compressione del segnale; l'estensione in frequenza arrivava fino a 3.500 Hz. Particolare attenzione viene dedicata alla variazione della larghezza di banda del filtro analizzatore (due sono quelli principalmente utilizzati, uno di 45 e l'altro di 300 Hz di larghezza di banda) e ai relativi risultati, facendo concludere agli autori che “[w]hether the wide or narrow

⁶⁷Bell et al., “Reduction of Speech Spectra by Analysis-by-Synthesis Techniques”, p. 1725.

⁶⁸Si tratta di Koenig, Dunn, and Lacy, “The Sound Spectrograph”.

⁶⁹Il termine inglese utilizzato è *portrayal*, che si può tradurre anche con rappresentazione, vivida descrizione, ‘ritratto’.

⁷⁰Nell'introduzione al numero, Potter annuncia l'imminente uscita del volume, di cui è coautore, intitolato appunto *Visible speech*.

⁷¹Ricordo che le ricerche di Graham Bell – come quelle di Edison – nella telefonia e nella registrazione dei suoni erano guidate, almeno in parte, dallo stesso scopo medico.

⁷²Fra cui ricordo Potter, “Transmission Characteristics of a Short-Wave Telephone Circuit”; Schuck and Young, “Observations on the Vibrations of Piano Strings”; Steinberg, “Application of Sound Measuring Instruments to the Study of Phonetic Problems”.

analyzing filter should be used depends on the nature of the sound and how far apart in frequency the components are”⁷³. L’articolo termina con una rassegna dei possibili suoni da analizzare oltre a quelli della voce; fra questi spiccano il canto degli uccelli e si nota l’assenza di suoni di strumenti musicali.

La rappresentazione sonografica diventerà in seguito utile strumento grafico per il compositore in fase di risintesi dei suoni. ‘Disegnare’ lo spettro di un suono è infatti una funzionalità spesso inserita nei software musicali⁷⁴.

3.7 Trasmissioni criptate

Già nel citato brevetto di Dudley, “Signal Transmission”, uno degli oggetti dell’invenzione era “di promuovere la segretezza nella trasmissione dei segnali, ad esempio nella trasmissione della parola via cavo o via radio”⁷⁵. Sottolineando le caratteristiche di codice del messaggio vocale, Dudley ci ricorda che i “codici sono utilizzati per la segretezza, come nelle comunicazioni militari”⁷⁶.

Il 15 luglio 1943, un gruppo di militari e di scienziati si riunisce fra il Pentagono (Washington) e Londra per dare formalmente il via a un servizio di telefonia segreta, il *SIGSALY*. Il coinvolgimento di un vasto gruppo di scienziati, tecnici ed inventori dei Bell Telephone Laboratories è significativo. Oliver Buckley, allora presidente dei Bell Telephone Laboratories, così ricorda l’avvenimento:

[w]e are assembled today in Washington and London to open a new service, secret telephony. It is an event of importance in the conduct of the war that others here can appraise better than I. As a technical achievement, I should like to point out that it must be counted among the major advances in the art of telephony. Not only does it represent the achievement of a goal long sought – complete secrecy in radiotelephone transmission – but it represents the first practical application of new methods of telephone transmission that promise to have far-reaching effects⁷⁷.

Il servizio si basa su una serie di tecnologie allora all’avanguardia, se non del tutto nuove. In particolare, il sistema sia di codifica sia di crittografia è illustrato in un brevetto – a lungo tenuto segreto, come tutto il progetto – di Dudley, depositato il 18 dicembre 1941 e ‘rilasciato’ il 12 ottobre 1976, 35 anni dopo⁷⁸. In esso, i segnali ‘digitali’ che rappresentano il codice criptato da inviare insieme al segnale trasmesso sono incisi su un disco che viene sincronizzato con un disco di decodifica dalla parte del ricevente. Un sistema molto simile si trova citato in un articolo di William Bennet⁷⁹ e riguarda un metodo descritto in un saggio del dicembre 1941 di Gustav

⁷³Koenig, Dunn, and Lacy, “The Sound Spectrograph”, p. 37.

⁷⁴Rimando, per questo aspetto, a Roads, *Microsound*, pp. 263-271.

⁷⁵Dudley, “Signal Transmission”, 1:28-30.

⁷⁶Dudley, “Fundamentals of Speech Synthesis”, p. 172.

⁷⁷Boone and Peterson, *Sigsaly - The Start of the Digital Revolution*.

⁷⁸Dudley, “Secret Telephony”.

⁷⁹Bennett, “Secret Telephony as a Historical Example of Spread-Spectrum Communications”, p. 102.

Guanella, “Methods for the automatic scrambling of speech”. L’autore aveva da poco – il 10 maggio – depositato in Svizzera il brevetto “Method and Apparatus for Scrambling Signals”. L’oggetto del brevetto è principalmente il sistema di crittografia, tuttavia da un altro brevetto di Guanella⁸⁰, sebbene successivo, sappiamo che egli era a conoscenza dei lavori di Dudley e la parte di codifica del segnale audio da lui descritta è chiaramente un sistema simile al vocoder. È, questa, una ulteriore testimonianza di come la tecnologia del vocoder stesse assumendo una diffusione planetaria parallelamente a quella delle comunicazioni, delle informazioni e dei sistemi che le informazioni avrebbero dovuto preservare. Il vocoder, testimone di questi sviluppi, sarà alla base di molte di queste, o simili, applicazioni.

3.8 Cibernetica

Il ne suffit pas que [l’âme raisonnable] soit logée dans le corps humain, ainsi qu’un *pilote* en son navire, sinon peut-être pour mouvoir ses membres, mais qu’il est besoin qu’elle soit jointe et unie plus étroitement avec lui, pour avoir outre cela des sentiments et des appétits semblables aux nôtres, et ainsi composer un vrai homme⁸¹.

Docet etiam natura, per istos sensus doloris, famis, sitis & c., me non tantùm adesse meo corpori ut *nauta* adest navigio, sed illi arctissime esse conjunctum & quasi permixtum, adeo ut unum quid cum illo componam⁸².

Il *nauta* (il *pilote*) controlla la sua nave, cosiccome fa il *kibernetes* (*gubernator*, timoniere): governa, pilota. In Cartesio la metafora è funzionale a mostrare che un pilota con la sua nave non bastano a formare un vero uomo, dotato di pensiero e di volontà e della capacità di reagire sia emotivamente sia volontariamente alle sensazioni. All’epoca di Cartesio, gli automi erano meccanismi ad orologeria, per quanto sofisticati, privi di sensori di qualsiasi tipo, e non erano in grado di modificare il loro comportamento in base alle informazioni ricevute dall’ambiente circostante: non avevano alcun meccanismo di *feedback*. Cartesio non poteva immaginare che il pilota della sua metafora avrebbe avuto in futuro la possibilità di reagire, se non emotivamente o razionalmente, almeno meccanicamente (automaticamente) agli stimoli esterni.

⁸⁰Si tratta di “Speech Transmission System”, il quale cita i due brevetti americani di Dudley – già visti *supra* in 3.4 – 2,098,956, “Signaling System” e 2,151,091, “Signal Transmission”, rilasciati rispettivamente il 16 novembre 1937 e il 21 marzo 1939.

⁸¹*Discours de la méthode* (1637). “Non basta che [l’anima razionale] sia posta nel corpo umano come un *pilota* nella sua nave, se non forse per muovere le sue membra, ma occorre che essa sia congiunta più strettamente con esso per avere, oltre a ciò, dei sentimenti e degli appetiti simili ai nostri, e comporre così un vero uomo”. Descartes, *Discorso sul metodo*, p. 69, corsivo mio.

⁸²Descartes, *Meditatio VI*, “De rerum materialium existentiâ, & reali mentis a corpore distinctione” (1641). “Ma, mediante queste sensazioni di fame, di sete, di dolore, eccetera, la natura mi insegna non soltanto che io sono posto nel mio corpo come un *pilota* nella sua nave, ma che gli sono anche congiunto così strettamente da comporre con lui un tutto unico”. In Descartes, *Opere filosofiche*, p. 250, corsivo mio.

Norbert Wiener introduce il nuovo termine di cibernetica per definire l'ambito disciplinare della teoria dell'informazione e dei meccanismi che si autoregolamentano proprio perché i sensori nelle nuove macchine permettono il *feedback*; sensori sempre più sofisticati e in grado di fornire informazioni a macchine in grado di memorizzarle. Se si vuole fare un uso umano degli esseri umani non li si deve “legare a un remo e impiegarli come sorgente di energia” né “segregarli in una fabbrica e assegnarli a un compito meramente meccanico che richieda meno di un milionesimo delle loro facoltà cerebrali”⁸³. Sempre più, mansioni di tipo meccanico-automatico saranno affidate a macchine e sempre più, in futuro, bisognerà interagire con esse, permettendo così all'essere umano di “elevarsi in tutta la sua statura”⁸⁴.

Fra gli esempi di “macchine di comunicazione” che vengono proposti da Wiener in *The human use of human beings* ve ne sono anche di appartenenti ad una categoria di importanza medica diretta, adatti a sostituire un senso con un altro. Se il sensore sta alla base di ogni sistema di autoregolamentazione, uno dei sensori sviluppati e studiati all'epoca, grazie alle applicazioni nella telefonia, è senz'altro il microfono. Ma il microfono non compie alcuna operazione ‘interpretativa’ se non captare il più fedelmente possibile il fenomeno fisico del suono. Quest'informazione di base deve essere elaborata e tradotta. Il vocoder compie per Wiener queste operazioni di elaborazione e traduzione: passa da un dominio sensoriale (l'udito) ad un altro (il tatto) per fornire ai sordi le informazioni relative al parlato tramite un sistema di stimolazione tattile delle dita: l'operazione inversa del vocoder.

Le applicazioni di protesi che sostituiscano una parte del corpo o una sua funzionalità erano, allora come oggi, all'ordine del giorno della ricerca in campo medico. Si pensava così di supplire alla sordità o con sistemi tattili o con sistemi visivi. ‘Sentire’, attraverso stimolazioni sui polpastrelli, quale porzione dello spettro è attiva e quale caratteristica principale ha il suono vocale che viene emesso in quell'istante sostituiva l'ascoltare⁸⁵. Allo stesso modo, guardare le forme contenute nei sonogrammi (la “visible speech”) permette di risalire alla parola⁸⁶.

4 Il Phase Vocoder

4.1 Inizi

Il termine “Phase Vocoder”⁸⁷ compare scritto in un *abstract* pubblicato sul «Journal of Acoustical Society of America». Il 6 novembre 1965, durante il settantesimo incontro dell'Acoustical Society of America, nella sessione dedicata alla sintesi della voce, James Loton Flanagan, insieme ai colleghi D. I. S. Meinhart, Roger

⁸³ Wiener, *Introduzione alla cibernetica*, p. 30.

⁸⁴ *Ibid.*, p. 31.

⁸⁵ Si vedano Wiener, “Sound Communication with the Deaf” e Pickett, “Tactual vocoder as an aid in speech transmission to the deaf”.

⁸⁶ Si veda 3.6.

⁸⁷ D'ora in poi PV.

M. Golden e Man Mohan Sondhi dei Bell Telephone Laboratories, presentano un *paper* intitolato “Phase Vocoder”⁸⁸. Dei dodici contributi presentati nella medesima sessione, ben sei trattano del *vocoder* o di sue varianti. A giudicare dall’*abstract*, la teoria esposta è quella che verrà descritta l’anno successivo nell’articolo di Flanagan and Golden, “Phase Vocoder”⁸⁹. La novità della proposta, rispetto al tradizionale *channel vocoder*, si può individuare nella capacità, durante la fase di analisi del suono, di preservare l’informazione sulla fase. “The difference between this and a conventional channel vocoder [...] is that the phase information is preserved in each channel and that we can guarantee that the output is exactly identical to the input”⁹⁰. L’introduzione della STFT (Short Time Fourier Transform) consente di evitare l’individuazione della frequenza fondamentale e la decisione, in fase di risintesi, relativa alla scelta fra sorgente armonica (il *buzz*) o rumore (l’*hiss*). Il modello è dunque cambiato, e la possibilità di manipolare i dati acquisiti durante l’analisi verrà presto messa in luce in ambito musicale. Già dalla metà dei successivi anni settanta, gli studi di Grey e Moorer sul timbro degli strumenti musicali faranno uso di una tecnica di analisi del suono basata sul filtro ad eterodina, corrispondente al PV di Flanagan e Golden⁹¹. In essi, svolti prima all’Università di Stanford poi all’IRCAM di Parigi, si sfruttano sia la capacità analitica sia quella rappresentativa del PV, aspetti – in particolare quest’ultimo – trascurato in ambito strettamente musicale dato il principale uso in campo fonetico fino ad allora prevalente. Saranno proprio questi studi sulle caratteristiche timbriche dei suoni, svolti in istituzioni che prevedevano la compresenza di scienziati e musicisti, a dare linfa vitale alle applicazioni in musica del PV.

Due momenti salienti nella letteratura sul PV, sia per la loro portata divulgativa sia per la loro capacità di rivolgersi ai musicisti senza per questo rinunciare al rigore scientifico, sono gli articoli di Gordon and Strawn, “An Introduction to the Phase Vocoder” e di Dolson, “The Phase Vocoder: A Tutorial”. Il primo “will attempt to present the phase vocoder on an intuitive level so that a musician interested in the technique can understand the key points”⁹². Il contributo non è originale, poiché si basa esplicitamente, per gli aspetti matematici, sull’articolo di Portnoff, “Implementation of the Digital Phase Vocoder Using the Fast Fourier Transform” e, per quanto riguarda l’implementazione software, sul PV di James Moorer sviluppato a Stanford; tuttavia illustra in maniera chiara il codice e le varie possibili applicazioni in campo musicale del PV. Il secondo verrà qui preso come termine ideale di un processo che continua però a proliferare e a partire dal quale nuova letteratura si è in seguito sviluppata. Questo tutorial ‘ispirerà’ molte implementazioni del PV, fra cui

⁸⁸Flanagan et al., “Phase Vocoder”, pp. 939-940.

⁸⁹Lo si può trovare qui riprodotto alle pp. 9-25.

⁹⁰Moorer, “The Use of the Phase Vocoder in Computer Music Applications”, p. 43.

⁹¹Si vedano Moorer, *On the Segmentation and Analysis of Continuous Musical Sound by Digital Computer*; Grey, “Multidimensional perceptual scaling of musical timbres”; Grey and Moorer, “Perceptual evaluations of synthesized musical instrument tones”.

⁹²Gordon and Strawn, “An Introduction to the Phase Vocoder”, p. 222.

quella realizzata nel 1997 da Dan Ellis e Richard Karpen per l'*opcode* di *Csound* "pvoc" e di tutta la famiglia di *opcodes* da esso derivata⁹³.

4.2 Spettri che variano nel tempo

"Another way of looking at the phase vocoder is to think of the analysis side as producing an entire set of spectral data [...] for a very short segment of the input signal at a given point time"⁹⁴. Mentre nel 'banco di filtri' prevale l'aspetto temporale della rappresentazione del suono, pur essendo evidentemente ben presente anche quello della frequenza, nel modello spettrale, che permette di valutare su microscopici (dell'ordine di 20 ms) segmenti temporali, la rappresentazione nel dominio della frequenza orienta le scelte ed è oggi diventata ormai familiare al musicista, permettendogli di avere una 'visione' del suono più analitica. Ben riassume Dolson le due rappresentazioni intuitive qui accennate, parlando di "filter view" (l'interpretazione tramite banco di filtri) e di "Fourier view" (l'interpretazione tramite trasformata di Fourier), laddove la prima si sviluppa nel tempo attorno ad una singola frequenza mentre la seconda è 'istantanea' su tutto lo spettro⁹⁵.

Una delle conseguenze di questa visione è la possibilità di pilotare – poiché i dati ricavati dall'analisi sono una sequenza di frequenza+ampiezza per ogni singolo canale (*bin*) – un banco di oscillatori per risintetizzare il suono⁹⁶. Proprio a questa visione si rifà ad esempio Wishart nella composizione di alcuni passaggi di *Vox 5*⁹⁷.

4.3 Phase Vocoder e musica

L'utilizzo del PV in musica è ormai diffusissimo. Tutti i generi musicali, che in qualche modo fanno uso di tecnologie elettroniche, lo hanno prima o poi inserito nel loro campionario di 'effetti'. Che le tecniche di sintesi derivate dal vocoder fossero adatte alla musica era evidente fin da subito. Durante le dimostrazioni del vocoder effettuate nel 1939 da Dudley e Vadersen all'Acoustical Society of America e alla New York Electrical Society, grazie ad una serie di controlli dei parametri introdotti nella fase della risintesi si poteva manipolare *live* (con un ritardo di 17 ms) la voce in varie maniere:

normal speech becomes a throaty whisper when the hiss is substituted for the buzz [...]. If variation [of the original pitch] is prevented, the re-created speech is a monotone, like a chant. When the relative variations cut in half, the

⁹³Si veda il manuale di *Csound* online, <<http://www.csounds.com/manual/html/pvoc.html>> [ultima consultazione 11 aprile 2013]. Si può trovare un'implementazione aggiornata in Matlab dell'algoritmo utilizzato per il "pvoc" in Ellis, *A Phase Vocoder in Matlab*.

⁹⁴Gordon and Strawn, "An Introduction to the Phase Vocoder", p. 225.

⁹⁵Dolson, "The Phase Vocoder: A Tutorial", p. 20.

⁹⁶Opero qui una semplificazione, puramente esemplificativa, riducendo i risultati della STFT ad una serie di coppie di valori ampiezza/frequenza. Per una distinzione fra "Overlap-Add Resynthesis" e "Oscillator Bank Resynthesis" in un contesto musicale si veda Roads, *Microsound*, pp. 257-258.

⁹⁷Si veda 4.3.2.

voice seems flat and dragging; when the swings are twice normal, the voice seems more brilliant; when four times normal it sounds febrile, unnatural. The controls can be reversed so that high becomes low: the tune of a song is then unrecognizable [...]. If the basic value of the re-created pitch is “fluttered” by hand, the voice becomes that of an old person. By appropriate setting of the basic pitch, the voice may be anything from low bass to a high soprano [...]. The basic pitch set to maintain a constant ratio of 5/4 to the original is a “major third” higher and harmonizes with the original [...]. For the vocal-cord tones of the original, the Vocoder substitutes the output of a relaxation oscillator. But any sound rich in harmonics can be used: an automobile horn, an airplane roar, an organ. In some demonstrations, the sound, taken from a phonograph record, replaces the buzz input from oscillator⁹⁸.

Dudley ipotizza un utilizzo radiofonico e cinematografico della sua invenzione; ma, indipendentemente dall’uso, il carattere di controllo dei dati ricavati dall’analisi è comunque evidentemente orientato al musicale.

Una prima riflessione sull’uso del PV in musica la propone James Andy Moorer⁹⁹, che è anche autore di uno studio musicale del 1978, intitolato *Lions are Growing*, basato sull’applicazione compositiva del PV¹⁰⁰. Nel suo articolo, Moorer illustra il lavoro di Portnoff, “Implementation of the Digital Phase Vocoder Using the Fast Fourier Transform”, e rileva che

[u]sing the phase vocoder, however, we can now modify this error signal [quello ottenuto da un’analisi LPC]. We can change its pitch or its timing, then reimpose the spectral shape by applying the optimum filter again. This gives us *independent control* over timing, spectrum, and pitch of the sound, a powerful combination indeed.¹⁰¹

Saranno queste possibilità di controllare valori tipicamente ‘musicali’ ad interessare i compositori.

Molti sono i brani che utilizzano il PV per l’elaborazione e la ri-composizione dei suoni. Con il PV sviluppato al CARL¹⁰² ricordo la parte di nastro in quadrifonia per *Transfigured Wind* (1984) di Roger Reynolds per organici vari, di Joji Yuasa *Towards the Midnight Sun* (1984) e *The Sea Darkens* (1987).

Nell’ambito della musica elettroacustica ‘colta’ ricordo due brani, che costituiscono due casi esemplari di utilizzo di PV, o comunque di tecniche ad esso

⁹⁸Dudley, Riesz, and Watkins, “A Synthetic Speaker”, pp. 125-126.

⁹⁹Moorer, “The Use of the Phase Vocoder in Computer Music Applications”.

¹⁰⁰Una breve descrizione del brano la si può trovare in <<http://www.ex-tempore.org/boulanger/boulanger.htm>> [ultima consultazione 11 aprile 2013].

¹⁰¹Si veda Moorer, “The Use of the Phase Vocoder in Computer Music Applications”, p. 44, corsivo mio.

¹⁰²“The phase vocoder program under development at the University of California, San Diego’s Computer Audio Research Laboratory (CARL) was first described in a talk presented at the International Computer Music Conference in 1983”; in Dolson, “Refinements in Phase-Vocoder-Based Modification of Music”, p. 65.

assimilabili, nella composizione. Entrambi si basano sull'idea di ibridazione, di *morphing*, di metamorfosi dei suoni.

4.3.1 *Mortuos Plango, Vivos voco* (1980) di Jonathan Harvey

Il brano di Jonathan Harvey *Mortuos Plango, Vivos Voco*¹⁰³, per nastro a otto tracce, è stato realizzato all'IRCAM nel 1980 con l'assistenza tecnica di Stanley Haynes, il quale, parlando delle tecniche di PV in un articolo del 1982, afferma che

[t]he evolution of the spectra can be modified either by changing the functions or by multiplying them by modifying functions produced by oscillators within the computer instrument. This gives the possibility of beginning with a recognizably instrumental or vocal sound and gradually shifting the amplitude and frequency of each harmonic component to produce a new spectrum. The British composer Jonathan Harvey has exploited these effects in a number of pieces, creating instrumentlike tones whose harmonic components spread out in pitch and equalize in amplitude to come to rest on chords¹⁰⁴.

Harvey utilizza il programma di analisi tramite FFT allora presente all'IRCAM, che proveniva dal pacchetto *S* di analisi del suono interattiva sviluppato all'Università di Stanford¹⁰⁵.

Harvey, cinque anni dopo la composizione di *Mortuos Plango, Vivos Voco*, parlando del timbro e dopo aver fornito alcuni esempi di pratiche di compositori quali Boulez, Stockhausen, Ligeti, Grisey, Murail, li accomuna sostenendo che “they are all playing with the identity given to objects by virtue of their having a timbre, in order to create *ambiguity*”¹⁰⁶. Il senso di ambiguità, già accennato in “Identity and Ambiguity: The Construction and Use of Trimbral Transitions and Hybrids”, viene reso grazie a ibridazione fra timbri diversi, realizzabile tramite tecniche di phase-vocoding.

Particolare importanza è stata attribuita da Harvey, per approfondire il concetto di ambiguità, allo studio dei suoni di strumenti acustici, in particolare di culture lontane da quella occidentale. In un articolo scritto a quattro mani con Jan Vanenheede

¹⁰³Un'introduzione al brano la si può leggere in Harvey, “*Mortuos Plango, Vivos Voco: A Realization at IRCAM*”, dal quale sono state tratte molte delle informazioni presenti in questo paragrafo. La prima esecuzione è avvenuta il 30 novembre 1980 al Lille Festival. Per un'analisi del brano che ripercorre, con tecnologie attuali, il processo di realizzazione dei suoni tramite l'analisi delle registrazioni della campana e della voce si veda Dirks, *An Analysis of Jonathan Harveys “Mortuos Plango, Vivos Voco”*. Un buon esercizio di ‘ricostruzione’ informatica lo propone Miller Puckette e lo si può trovare qui <<http://crca.ucsd.edu/msp/pdrp/latest/files/doc/harvey/harvey.htm>>. Ad oggi, l'unica analisi di mia conoscenza che affronta il problema della distribuzione del brano nello spazio è Zattra, *Studiare la computer music*, pp. 243-256. Un'edizione discografica stereofonica la si può trovare in Aa. Vv. *Computer Music Currents 5*.

¹⁰⁴Haynes, “The Computer as a Sound Processor: A Tutorial”, p. 14.

¹⁰⁵Harvey, “*Mortuos Plango, Vivos Voco: A Realization at IRCAM*”, p. 22.

¹⁰⁶Harvey, “The Mirror of Ambiguity”, pp. 178-179.

mentre lavorava all'IRCAM¹⁰⁷, sebbene scritto nel 1985 e legato alla realizzazione del brano *Bhakti* (1982) per 15 esecutori e nastro quadrifonico, Harvey testimonia del suo interesse per l'"ambiguità" dei timbri strumentali, per la transizione fra un timbro e l'altro, per le possibilità di ibridazione timbrica, aspetti che lo avevano già interessato nella composizione di *Mortuos plango, vivos voco*. Nella risintesi dei suoni vocali viene tuttavia utilizzata una tecnica 'granulare' (la FOF, motore di sintesi del programma *CHANT* sviluppato all'IRCAM¹⁰⁸). La compresenza di tecniche sia di sintesi (granulare tramite FOF) sia di sintesi per analisi (tramite filtro ad eterodina) evidenzia la diversità degli strumenti utilizzati dai compositori, e soprattutto la diversità dei parametri di controllo a loro disposizione per generare nuovi mondi sonori: da un punto di vista compositivo ragionare sui parametri della sintesi granulare o sui parametri di un banco di oscillatori significa creare suoni, strutture e processi musicali anche molto distanti fra loro.

I materiali concreti utilizzati da Harvey in *Mortuos Plango, Vivos voco* sono due: la voce del figlio e la grande campana della cattedrale di Winchester, per la quale Harvey ha scritto molta musica corale. Il testo letto e cantato dalla voce è l'iscrizione presente sulla campana: "Horas Avolantes Numero, Mortuos Plango: Vivos ad Preces Voco"¹⁰⁹.

La fase di analisi del suono ha occupato una buona parte del lavoro di Harvey, in quanto da essa si sono ricavate molte informazioni per la costruzione del brano, utili sia alla sintesi dei suoni sia alle relazioni formali e alle scelte strutturali. Il suono della campana, analizzato mezzo secondo dopo l'inizio del rintocco tramite FFT, riporta i valori di altezza 'trascritti' in notazione tradizionale riprodotti in Fig. 10.



Fig. 10: Le parziali della campana della cattedrale di Winchester ottenute da Harvey tramite FFT seguite dalla "strike note".

"Le otto sezioni del lavoro, con le loro altezze centrali, sono strutturate attorno alle parziali mostrate nella Fig. 11"¹¹⁰.

La sintesi e il mixaggio dei suoni sono stati realizzati con la versione dell'IRCAM del *Music V*. La possibilità di controllare l'ampiezza di ogni singola parziale in fase di risintesi ha permesso ad Harvey di variarne il decadimento singolarmente, ad esempio aumentando la durata di quelle acute, normalmente rapide, e diminuendo la durata di quelle gravi, normalmente lente. Harvey, alcuni anni dopo, sottolinea che "instead of making all the models with FOFs as we originally intended, we

¹⁰⁷Vandenheede and Harvey, "Identity and Ambiguity: The Construction and Use of Trimbral Transitions and Hybrids".

¹⁰⁸Si veda Rodet, Potard, and Barrière, *Chant*.

¹⁰⁹"Conto le ore che volano, piango i morti: i vivi chiamo alla preghiera".

¹¹⁰Harvey, "*Mortuos Plango, Vivos Voco: A Realization at IRCAM*", p. 22.



Fig. 11: Le note/frequenze, tratte dallo spettro della campana, attorno alle quali ruotano le otto sezioni del brano.

finally had to use different types of synthesis for the models. So the bell, although being possible with pure FOFs is a pure additive synthesis instrument, to get a more precise control over its time evolution”¹¹¹, proprio ciò che aveva messo in pratica in *Mortuos plango, vivos voco*.

Un’interessante applicazione della risintesi è quella che prevede la distribuzione nello spazio ottonico delle singole parziali ricavate dall’analisi del suono della campana, che “danno all’ascoltatore la curiosa sensazione di essere dentro la campana”¹¹². La suddivisione spaziale dello spettro di un suono è a mio avviso uno degli aspetti centrali nella composizione elettroacustica; di questo utilizzo ci fornisce una descrizione anche Robert Normandeau, quando afferma che “what is unique in electroacoustic music is the possibility to fragment sound spectra amongst a network of speakers. [...] With multichannel electroacoustic music, timbre can be distributed over all virtual points available in the defined space”¹¹³.

4.3.2 *Vox 5* (1986) di Trevor Wishart

Quinto di sei brani che compongono il *Vox Cycle* (1977-1986), *Vox 5* è l’unico del ciclo interamente elettroacustico. Wishart descrive il brano, nelle sue linee generali, in un articolo del 1988¹¹⁴. “The sound transformations in *Vox 5* were largely achieved using the Phase Vocoder”¹¹⁵. La genesi di questo brano ci fornisce molte indicazioni su come e dove, nei primi anni ottanta del secolo scorso, le tecnologie a disposizione dei compositori di musica elettroacustica stessero evolvendosi. Nel 1979 Wishart, dopo aver esplorato a fondo le apparecchiature audio analogiche e cercato un modo per realizzare con esse metamorfosi sonore, si rivolge all’IRCAM di Parigi – “at that time the only facility for advanced computer-music in Europe”¹¹⁶ – per realizzare un brano. Durante il soggiorno di sei settimane scopre alcuni

¹¹¹Vandenhede and Harvey, “Identity and Ambiguity: The Construction and Use of Trimbral Transitions and Hybrids”, p. 100.

¹¹²Harvey, “*Mortuos Plango, Vivos Voco*: A Realization at IRCAM”, p. 24.

¹¹³Normandeau, “The visitors and the residents”, p. 62.

¹¹⁴Wishart, “The Composition of *Vox 5*”. Una segmentazione analitica del brano è proposta da Williams, “*Vox 5* by Trevor Wishart”, in cui è riprodotta la ‘partitura’ per la diffusione redatta da Wishart e una partitura d’ascolto grafica redatta dall’autore dell’articolo. Un’edizione discografica è Wishart, *The Vox Cycle*. La prima esecuzione è avvenuta nel 1987 a Parigi all’interno dell’INA/GRM Cycle Acousmatique di Radio France. Ne esiste anche una versione quadrifonica.

¹¹⁵Wishart, *Sound Composition*, p. 63.

¹¹⁶Ibid., p. 61.

‘strumenti’, fra i quali l’LPC (Linear Predictive Coding). A seguito di questa breve permanenza a Parigi viene invitato a realizzare le sue idee musicali ma, a causa di un rinnovamento delle apparecchiature, quando decide di riprendere il suo progetto compositivo dovrà aspettare il 1986 per poterlo mettere in opera.

It was suggested to me that the CARL Phase Vocoder (Moore, Dolson) might be a better tool to use, but no-one at IRCAM at that time had inside knowledge of the workings of this program, so I took apart the data files it produced to work out for myself what was going on.

I eventually developed a number of software instruments for the spectral transformation of sounds which were then used to compose *Vox 5*. These instruments massaged the data in the analysis files produced by the Phase Vocoder. The most significant of these were **stretching the spectrum** and **spectral morphing** - creating a seamless transition between two different sounds *which are themselves in spectral motion*¹¹⁷.

Possiamo vedere la ‘partitura’ per la diffusione redatta da Wishart in Fig. 12.

La novità fu dunque l’introduzione del PV, che Wishart studiò approfonditamente e con il quale realizzò parte del software convogliato in seguito nel Composer Desktop Project¹¹⁸. La conoscenza approfondita dei dati ricavati dall’analisi e la possibilità di manipolarli, modificarli, controllarli a piacere rendono per Wishart il PV uno strumento di lavoro adatto alla sua poetica. Ma soltanto dopo la programmazione degli ambienti di controllo dei dati sarà per lui possibile un uso veramente creativo di uno strumento così versatile come il PV. Wishart insiste sul continuo *feedback* fra dati ricavati dall’analisi, ascolto, dati generati per la sintesi. Sono momenti distinti di un processo teso alla composizione. Il sistema da lui utilizzato per realizzare *Vox 5* “was an analysis-synthesis tool. You could run analysis on its own, synthesis on its own, or analysis+synthesis. The possibility to run analysis and synthesis separately allowed me to write software to manipulate the analysis data, then to resynthesize the output”¹¹⁹. La continua manipolazione dei dati e la continua verifica percettiva pongono questioni quali la tendenza alla categoricità della nostra percezione. Illustrando un esempio di *morphing* fra suono umano e suono animale (voce -> ronzio di api), Wishart sottolinea che “we hear sounds either as a voice or as a sound of bees. Hence there is a moment in the transformation, no matter how smooth, where aural recognition tends to suddenly switch from ‘It’s a voice’ to ‘No, it’s a bees!’”¹²⁰. Calibrando e modificando adeguatamente i dati durante l’interpolazione fra i due spettri (quello della voce e quello del ronzio dell’ape), Wishart ottiene l’effetto di *morphing* desiderato. “In the *Voice -> Bees* morph the

¹¹⁷Wishart, *Computer Sound Transformation*.

¹¹⁸Si veda *Composers Desktop Project*. Per un riepilogo delle operazioni sullo spettro possibili con il CDP si veda inoltre Roads, *Microsound*, pp. 264-267. Sull’uso del PV da parte di Wishart si possono leggere, oltre al già citato contributo su *Vox 5*, i suoi *Audible Design, Computer Sound Transformation* e l’articolo qui presente alle pp. 61-65.

¹¹⁹Comunicazione personale, e-mail del 4 aprile 2013.

¹²⁰Wishart, *Sound Composition*, p. 65.

5 Conclusioni

La separazione – sempre più intangibile – fra controllo (‘gesto’, partitura) e strumento/orchestra (‘sintesi’, generazione del suono), e soprattutto la possibilità di intervenire contemporaneamente su *entrambi*, sono le novità che il compositore di musica elettroacustica deve affrontare. La sfida di comporre *sia* lo strumento *sia* il controllo dello strumento fanno del compositore un demiurgo in grado di costruire la macchina e di farla funzionare. Infondere, tramite la propria arte (tecnica), un’anima allo strumento/orchestra gli sarà possibile solo dopo averlo costruito. Lunghi dal voler dare all’anima significati metafisici, le si vuole attribuire, in questo contesto, il carattere della ‘creatività’, proprio di ogni artista. Il vocoder – che permette al compositore di manipolare il controllo modificando i dati ricavati dall’analisi –, e tutta la famiglia di strumenti da esso derivata, è paradigmatico di questa caratteristica, peculiare a tutta la musica elettroacustica. 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.”¹²³.

Bibliografia

- Aa. Vv. *Composers Desktop Project*. 2010. URL: <http://www.composersdesktop.com/> (visited on 03/19/2013).
- *Computer Music Currents 5*. Computer Music Currents 5; WER 2025-2. CD. Mainz: Wergo, [1990].
- Académie des sciences (France). “Procès-verbaux. T102 (1783) / Académie royale des sciences”. Manoscritto. 1783. URL: <http://gallica.bnf.fr/ark:/12148/bpt6k557782/f408.image>.
- Backhaus, H. “Über die Bedeutung der Ausgleichsvorgänge in der Akustik”. In: *Zeitschrift für Technische Physik* 13.1 (1932), pp. 31–46.
- “Über Geigenklänge”. In: *Zeitschrift für Technische Physik* 8 (1927), p. 510.
- Bell, C. G. et al. “Reduction of Speech Spectra by Analysis-by-Synthesis Techniques”. In: *The Journal of the Acoustical Society of America* 33.12 (Dec. 1961), pp. 1725–1736.
- Bennett, William R. “Secret Telephony as a Historical Example of Spread-Spectrum Communications”. In: *IEEE Transactions on Communications* COM-31.1 (Jan. 1983), pp. 98–104.

¹²³Roads, *Microsound*, p. 258.

- Boone, J. V. and R. R. Peterson. *Sigsaly - The Start of the Digital Revolution*. URL: {https://www.nsa.gov/about/cryptologic_heritage/center_crypt_history/publications/sigsaly_start_digital.shtml}.
- Bowers, Brian. *Sir Charles Wheatstone FRS. 1802-1875*. IEE history of technology 29. London: The Institution of Electrical Engineers, 2001.
- Cordemoy, Géraud de. *Discours physique de la parole*. Paris: Florentin Lambert, 1668. URL: <http://gallica.bnf.fr/ark:/12148/bpt6k50506f>.
- Descartes, René. *Discorso sul metodo*. Ed. by Gianluca Mori. Firenze: La Nuova Italia, 2000.
- *Opere filosofiche*. Torino: UTET, 1969.
- Dirks, Patricia Lynn. *An Analysis of Jonathan Harveys "Mortuos Plango, Vivos Voco"*. 2007. URL: http://cec.sonous.ca/econtact/9_2/dirks.html (visited on 06/24/2012).
- Dolson, Mark. "Refinements in Phase-Vocoder-Based Modification of Music". In: *ICMC Proceedings*. 1984, pp. 65–66.
- "The Phase Vocoder: A Tutorial". In: *Computer Music Journal* 10.4 (Feb. 1986), pp. 14–27.
- Dudley, Homer. "Fundamentals of Speech Synthesis". In: *Journal of the Audio Engineering Society* 3.4 (Oct. 1955), pp. 170–185.
- "Remaking Speech". In: *The Journal of the Acoustical Society of America* 11.2 (Oct. 1939), pp. 169–177.
- "Synthesizing Speech". In: *Bell Laboratories Record* 15 (Dec. 1936), pp. 159–177.
- "The Carrier Nature of Speech". In: *Bell System Technical Journal* 19.4 (Oct. 1940), pp. 495–515.
- "The Vocoder". In: *Bell Laboratories Record* 17.0 (Dec. 1939), pp. 122–126.
- Dudley, Homer and Otto O. Jr. Gruenz. "Visible Speech Translators with External Phosphors". In: *Journal of Acoustical Society of America* 18.1 (July 1946), pp. 62–73.
- Dudley, Homer and T. H. Tarnoczy. "The Speaking Machine of Wolfgang von Kempelen". In: *The Journal of the Acoustical Society of America* 22.2 (Mar. 1950), pp. 151–185.
- Dudley, Homer W. "Secret Telephony". Pat. US 3,985,958. Oct. 1976. URL: <http://www.google.com/patents/US3985958>.
- "Signal Transmission". Pat. US 2,151,091. URL: <http://www.google.com/patents/US2151091>.
- "Signaling System". Pat. US 2,098,956. URL: <http://www.google.com/patents/US2098956>.
- "System for the artificial production of vocal or other sounds". Pat. US 2,121,142. URL: <http://www.google.com/patents/US2121142>.
- Dudley, Homer W., Robert R. Riesz, and S. S. A. Watkins. "A Synthetic Speaker". In: *Journal of The Franklin Institute* 227.6 (June 1939), pp. 739–764.

- Ellis, D. P. W. *A Phase Vocoder in Matlab*. 2002. URL: <http://www.ee.columbia.edu/~dpwe/resources/matlab/pvoc/>.
- Fagen, D. et al., eds. *A History of Engineering and Science in the Bell System*. The Laboratories, 1975-1984.
- Flanagan, James Loton. *Speech Analysis Synthesis and Perception*. 2nd ed. Kommunikation und Kybernetik in Einzeldarstellungen 3. Berlin, New York: Springer-Verlag, 1972.
- Flanagan, James Loton and R. M. Golden. "Phase Vocoder". In: *Bell System Technical Journal* 45.9 (Nov. 1966), pp. 1493-1509.
- Flanagan, James Loton et al. "Phase Vocoder". In: *The Journal of the Acoustical Society of America* 38 (Nov. 1965). Abstract, pp. 939-940.
- Gordon, John William and John Strawn. "An Introduction to the Phase Vocoder". In: *Digital Audio Signal Processing: An Anthology*. Ed. by John Strawn and F. R. Moore. W. Kaufmann, 1985, pp. 221-270.
- Grey, John M. "Multidimensional perceptual scaling of musical timbres". In: *The Journal of the Acoustical Society of America* 61 (May 1977), pp. 1270-1277.
- Grey, John M. and James A. Moorer. "Perceptual evaluations of synthesized musical instrument tones". In: *The Journal of the Acoustical Society of America* 62 (Aug. 1977), pp. 454-462.
- Guanella, Gustav. "Method and Apparatus for Scrambling Signals". Pat. US 2,364,210. URL: <http://www.google.com/patents/US2364210>.
- "Methods for the automatic scrambling of speech". In: *Brown Boveri Rev.* 28 (Dec. 1941), pp. 397-408.
- "Speech Transmission System". Pat. US 2,439,293. URL: <http://www.google.com/patents/US2439293>.
- Hankins, Thomas L. and Robert J. Silverman. *Instruments and the Imagination*. 2nd ed. Princeton: Princeton University Press, 1999.
- Harvey, Jonathan. "*Mortuos Plango, Vivos Voco*: A Realization at IRCAM". In: *Computer Music Journal* 5.4 (1981), pp. 22-24.
- "The Mirror of Ambiguity". In: *The Language of Electroacoustic Music*. Ed. by Simon Emmerson. Houndmills and London: The Macmillan Press Ltd., 1986, pp. 175-190.
- Haynes, Stanley. "The Computer as a Sound Processor: A Tutorial". In: *Computer Music Journal* 6.1 (1982), pp. 7-17.
- Henrici, O. "On a New Harmonic Analyzer". In: *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*. 5th ser. 38 (1894), pp. 110-121.
- Kempelen, Wolfgang von. *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine*. Wien: Degen, 1791.
- Koenig, Rudolph. *Quelques Expériences d'Acoustique*. Paris: Quai d'Anjou, 27, 1882. URL: <http://archive.org/details/quelquesexprie00koen> (visited on 03/27/2012).
- Koenig, W., H. K. Dunn, and L. Y. Lacy. "The Sound Spectrograph". In: *The Journal of the Acoustical Society of America* 18.1 (1946), pp. 19-49.

- Kratzenstein, Christian Gottlieb. *Tentamen resolvendi problema ab Academia Scientiarum Imperiali Petropolitana ad annum 1780 publice propositu*. Petropoli: Academiae Scientiarum, 1781.
- Kyunnap, E. "Speech Commands in Control Systems". In: *NASA Technical Translation NASA TT F-11,252* (Apr. 1968). Trad. en. di "Ustnyye Komandy v Sistemakh Upravleniya", pp. 1–32. URL: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19680016296_1968016296.pdf.
- "Ustnyye Komandy v Sistemakh Upravleniya". In: *Izvestiya Akademii Nauk Estonskoy SSR. Seriya Fiziko-Matematicheskikh i Tekhnicheskikh Nauk* 15.3 (1966), pp. 377–399.
- La Mettrie, Julien Offray de. *L'homme machine*. 1748.
- Lindsay, David. "Talking Head". In: *American Heritage of Invention & Technology* 13.1 ().
- Manning, Peter. *Electronic & Computer Music*. Revised and expanded edition. Oxford: OUP, 2004.
- Maxey, David H., ed. *Smithsonian Speech Synthesis History Project (SSSHP) 1986 - 2002*. URL: http://amhistory.si.edu/archives/speechsynthesis/ss_home.htm (visited on 03/05/2013).
- Michelson, A. A. and S. W. Stratton. "A New Harmonic Analyzer". In: *The American Journal of Science*. 4th ser. 5.25 (Jan. 1898), pp. 1–13.
- Miller, Dayton Clarence. *The Science of Musical Sounds*. New York: The Macmillian Company, 1916. URL: <http://www.archive.org/details/scienceofmusical00mill>.
- Miller, Ralph L. "Signaling System". Pat. US 2,117,739. URL: <http://www.google.com/patents/US2117739>.
- Millikan, Frank Rives. *Joseph Henry and the Telephone*. Ed. by Smithsonian Institution. URL: <http://siarchives.si.edu/history/jhp/joseph23.htm#1> (visited on 03/11/2013).
- Montmignon, Jean-Baptiste. *Système de prononciation figurée: applicable à toutes les langues et exécuté sur les langues françoises et angloise*. Paris: Royez, 1785. URL: <http://gallica.bnf.fr/ark:/12148/bpt6k520981>.
- Moore, F. Richard. *Elements of Computer Music*. Englewood Cliffs: Prentice Hall, 1990.
- Moorer, James A. *On the Segmentation and Analysis of Continuous Musical Sound by Digital Computer*. STAN-M-3. Center for Computer Research in Music and Acoustics - Stanford University, May 1975. 171 pp. URL: <https://ccrma.stanford.edu/files/papers/stanm3.pdf> (visited on 01/22/2011).
- Moorer, James Andy. "The Use of the Phase Vocoder in Computer Music Applications". In: *Journal of the Audio Engineering Society* 26.112 (1978), pp. 42–46.
- Moro, Elena Battaner. "A 19th-Century Speaking Machine: The Tecnefón of Severino Perez y Vazquez". In: *Historiographia Linguistica* 34.1 (2007), pp. 19–36.

- URL: <http://www.ingentaconnect.com/content/jbp/hl/2007/00000034/00000001/art00002>.
- Normandeau, Robert. "The visitors and the residents". In: *Musica/Tecnologia* 4.4 (2010). URL: <http://www.fupress.net/index.php/mt/article/view/9108>.
- Ohala, John J. "Christian Gottlieb Kratzenstein: pioneer in speech synthesis". In: *International Congress of Phonetic Sciences* (2011), pp. 156–159. URL: <http://www.icphs2011.hk/resources/OnlineProceedings/SpecialSession/Session7/Ohala/Ohala.pdf> (visited on 03/02/2012).
- Pantalony, D. *Altered sensations: Rudolph Koenig's acoustical workshop in nineteenth-century Paris*. Archimedes Series. Dordrech Heidelberg London New York: Springer, 2009.
- Pickett, J. M. "Tactual vocoder as an aid in speech transmission to the deaf". In: *STL-Quarterly Progress and Status Report* 2.4 (1961), pp. 14–17. URL: http://www.speech.kth.se/prod/publications/files/qpsr/1961/1961_2_4_014-017.pdf (visited on 04/08/2013).
- Pisko, Fr. Jos. *Die neueren apparate der akustik: Für freunde der naturwissenschaft und der tonkunst*. Wien: Carl Gerold's sohn, 1865. Google Books: vj0IAAAAI AAJ.
- Pompino-Marschall, Bernd. "Von Kempelen et al. Remarks on the history of articulatory-acoustic modelling". In: *ZAS Papers in Linguistics* 40 (2005), pp. 145–159. URL: http://www.zas.gwz-berlin.de/fileadmin/material/ZASPiL_Volltexte/zp40/zaspil40-pompino-marschall.pdf (visited on 12/15/2012).
- Portnoff, Michael R. "Implementation of the Digital Phase Vocoder Using the Fast Fourier Transform". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* ASSP-24.3 (June 1976), pp. 243–248.
- Potter, Ralph K. "Introduction to Technical Discussions of Sound Portrayal". In: *Journal of Acoustical Society of America* 18.1 (July 1946), pp. 1–3.
- Potter, Ralph. K. "Transmission Characteristics of a Short-Wave Telephone Circuit". In: *Proceedings of IRE*. Vol. 18. 4. Apr. 1930, pp. 581–648.
- Potter, Ralph K., George A. Kopp, and Harriet C. Green. *Visible speech*. New York: D. Van Nostrand Inc., 1947.
- Preece, William Henry and J. M. Augustus Stroh. "Studies in Acoustics. I. On the Synthetic Examination of Vowel Sounds". In: *Proceedings of the Royal Society of London; Philosophical Transactions of the Royal Society*. Vol. 28. Feb. 1879, pp. 358–367. URL: <http://archive.org/details/philtrans07329363>.
- Prieberg, Fred K. *Musica ex machina. Über das Verhältnis von Musik und Technik*. Berlin-Frankfurt-Wien: Verlag Ullstein, 1960.
- Ramsay Gordon, J. "L'Abbé Mical et les Têtes Parlantes: L'Histoire de Sa Vie, l'Histoire de Son Oeuvre". In: *10ème Congrès Français d'Acoustique*. Ed.

- by Société Française d'Acoustique SFA. Lyon, France, 2010. URL: <http://hal.archives-ouvertes.fr/hal-00549594>.
- Ranostaj, Frank. "Untersuchungen zur Sprechtraktakustik". PhD thesis. Frankfurt am Main: Johann Wolfgang Goethe-Universität, 2012. URL: <http://arxiv.org/pdf/1302.1619v1.pdf>.
- Rees, Torben. *Helmholtz's apparatus for the synthesis of sound: an electrical 'talking machine'*. Ed. by Explore Whipple Collections, Whipple Museum of the History of Science, University of Cambridge. 2010. URL: <http://www.hps.cam.ac.uk/whipple/explore/acoustics/hermanvonhelmholtz/helmholtzssynthesizer/> (visited on 04/04/2013).
- Riesz, Robert R. "Artificial larynx". Pat. US 1,836,816. URL: <http://www.google.com/patents/US1836816>.
- "Description and Demonstration of an Artificial Larynx". In: *Journal of Acoustical Society of America* 1.2 (Jan. 1930), pp. 273–279.
- "Electrical Artificial larynx". Pat. US 2,056,295. URL: <http://www.google.com/patents/US2056295>.
- Riesz, Robert R. and L. Schott. "Visible Speech Cathode-Ray Translator". In: *The Journal of the Acoustical Society of America* 18.1 (July 1946), pp. 50–61.
- Rivarol, Antoine de. *Oeuvres complètes de Rivarol, précédées d'une notice sur sa vie; ornées du portrait de l'auteur*. Vol. 2. Paris: Léopold Collin, 1808. URL: <http://gallica.bnf.fr/ark:/12148/btv1b86264329>.
- Roads, Curtis. *Microsound*. Cambridge, London: The MIT Press, 2001.
- Rodet, Xavier, Yves Potard, and Jean-Baptiste Barrière. *Chant. De la synthèse de la voix à la synthèse en général*. Rapport 35. Paris: IRCAM, 1985.
- Rubin, Philip and Eric Vatikiotis-Bateson. *The Science of the Spoken and Written Word*. Ed. by Haskins Laboratories. URL: <http://www.haskins.yale.edu/featured/heads/simulacra.html> (visited on 02/26/2013).
- Scha, Remko. *Virtual Voices (1)*. 1992. URL: <http://www.mediamatic.net/8635/en/virtual-voices-1> (visited on 03/01/2013).
- *Virtual Voices (2)*. 1992. URL: <http://www.mediamatic.net/8340/en/virtual-voices-2> (visited on 03/01/2013).
- Schroeder, M. R. "Vocoders: Analysis and Synthesis of Speech. A Review of 30 Years of Applied Speech Research". In: *Proceedings of the IEEE*. Vol. 54. 5. May 1966, pp. 720–733.
- Schuck, O. H. and R. W. Young. "Observations on the Vibrations of Piano Strings". In: *Journal of Acoustical Society of America* 15.1 (July 1943), pp. 1–11.
- Steinberg, J. C. and N. R. French. "The Portrayal of Visible Speech". In: *Journal of Acoustical Society of America* 18.1 (July 1946), pp. 4–18.
- Steinberg, John C. "Application of Sound Measuring Instruments to the Study of Phonetic Problems". In: *Journal of Acoustical Society of America* 6.0 (July 1934), pp. 16–24.
- Stewart, John Q. "An Electrical Analogue of the Vocal Organs". In: *Nature* 110 (1922), pp. 311–312.

- Stumpf, Carl. *Die Sprachlaute*. Berlin: Springer, 1926.
- Thomson, sir William. "Harmonic Analyzer". In: *Proceedings of the Royal Society* 27 (May 1878), pp. 371–374.
- Traunmüller, Hartmut. *Wolfgang von Kempelen's speaking machine and its successors*. 2000. URL: <http://www2.ling.su.se/staff/hartmut/kemplne.htm> (visited on 03/01/2013).
- Vandenheede, Jan and Jonathan Harvey. "Identity and Ambiguity: The Construction and Use of Trimbral Transitions and Hybrids". In: *ICMC Proceedings*. Ann Arbor, MI: MPublishing, 1985, pp. 97–102. URL: <http://hdl.handle.net/2027/spo.bbp2372.1985.016>.
- Vázquez, Severino Pérez y. *El Tecnefón, y los medios representativos de la palabra*. 1868. URL: <http://books.google.it/books?id=-PEEcgAACAAJ>.
- Wheatstone, Charles. "Wind Musical Instruments". Pat. EN 5,803.
- Wiener, Norbert. *Introduzione alla cibernetica. L'uso umano degli esseri umani*. Universale Bollati Boringhieri 14. Torino: Bollati Boringhieri, 1966.
- "Sound Communication with the Deaf". In: *Philosophy of Science* 16.3 (1949), pp. 260–262.
- *The human use of human beings. Cybernetics and society*. Houghton Mifflin, 1950.
- Williams, Tom. "Vox 5 by Trevor Wishart. The analysis of an electroacoustic tape piece". In: *Journal of Electroacoustic Music* 7 (1993), pp. 6–13.
- Willis, W. "On vowel sounds, and on reed-organ pipes". In: *Trans. Camb. Phil. Soc.* 3 (1838), p. 231.
- Wishart, Trevor. *Audible Design*. CD allegato. Orpheus the Pantomime, 1994.
- *Computer Sound Transformation. A personal perspective from the U.K.* Oct. 2000. URL: <http://www.composersdesktop.com/trnsform.htm> (visited on 02/16/2012).
- *Sound Composition*. 2 CD allegati. In proprio, [2012].
- "The Composition of Vox 5". In: *Computer Music Journal* 12.4 (1988), pp. 21–27.
- *The Vox Cycle*. OPT 904. CD. s. d.
- Zattra, Laura. *Studiare la computer music. Definizioni, analisi, fonti*. Padova: libreri-auniversitaria.it, 2011.
- Zavagna, Paolo. "Trascrivere documenti sonori". In: *Musica/Tecnologia* 6.0 (2012). URL: <http://www.fupress.net/index.php/mt/article/view/11908>.