

# Primi rilievi EGG sulla voce cantata

Antonio Romano<sup>1</sup>, Valentina De Iacovo<sup>1</sup>, Dario Strangis<sup>2</sup>, Silvestro Roatta<sup>3</sup>

<sup>1</sup>LFSAG - <sup>2</sup>Voce InForma–Queen Margaret Univ. - <sup>3</sup>Ist. Fisiologia, Dip. Neuroscienze, UniTo

## Introduzione

Nell'ambito del progetto «fisiologia emissione vocale parlata e cantata» il LFSAG ha finalmente ottenuto la possibilità di acquistare un sistema di rilevamento elettroglottografico (EGG).

Dopo mesi di attesa, il dispositivo fornito dall'azienda *Laryngograph* Ltd. (Sutton, UK) è stato ricevuto e immediatamente testato<sup>1</sup>. I primi rilievi sono stati resi possibili grazie alla pazienza di Xinghui Hu, *general manager* dell'azienda nata dalla collaborazione con alcuni dei pionieri degli studi basati sulla tecnica della glottografia impedenziometrica: Evelyn Abberton & Adrian Fourcin<sup>2</sup>.

1. Apprezziamo per ringraziare il personale di UniTo che ha garantito le condizioni per concludere l'acquisto. Sul piano amministrativo questo è avvenuto con RDA 51865 (<https://www.unito.it/ugov/trasparenza/2406578>).

2. V. Abberton & Fourcin (1997). L'incontro dell'autore AR con le tecniche sviluppate da questi ricercatori era già avvenuto nell'ambito della realizzazione della sua tesi di laurea in telecomunicazioni (*Analyse de la fonction de transfert du conduit vocal dans les intervalles de fermeture de la glotte et extraction de la forme d'onde d'excitation glottale par filtrage inverse*, 1994). In particolare in quella sede i riferimenti erano stati offerti da Fourcin & Abberton (1971) e Fourcin A. (1974). Gli autori dei numerosi studi condotti con questi

La tecnica, generalmente nota come *Impedance Glottography*, recentemente illustrata da Celata & Ricci (2021), trova riscontro in numerosi studi ed è stata implementata in diversi prodotti commerciali. La nostra esperienza era infatti partita dalla valutazione preliminare di altri dispositivi simili (quello della *F-J Electronics*, appartenente al fondo A. Genre del LFSAG, o l'estensione EGG CSL 4400 della Kay Elemetrics, testato presso il Laboratorio di Fonetica del DLLC dell'Univ. di Aveiro, Portogallo, diretto da Lurdes de Castro Moutinho, su cui si basano le considerazioni riportate nel §1).

Dopo un primo collaudo, il dispositivo EGG in dotazione del LFSAG è stato collocato provvisoriamente nella sede dove sono avvenuti i rilievi discussi in questo breve resoconto<sup>3</sup>.

metodi accreditano come precursore quello illustrato da P. Fabre nel 1957 ("Un procédé électrique percutané d'inscription de l'accollement glottique au cours de la phonation: glottographie de haute fréquence. Premiers résultats", *Bulletin de l'Académie Nationale de Médecine*, 141, 66-99); cfr. anche la testimonianza di K. Marasek, nell'ormai sito storico pluridecennale EGG dell'università di Stoccarda, 1997, e, più recentemente, Henrich et al. (2004).

3. Si tratta del laboratorio che si trova sotto la responsabilità dell'autore SR nella strut-

Nel §2 si dettagliano in particolare le caratteristiche del segnale EGG che è stato possibile analizzare sulla base di alcune emissioni vocali controllate, in questa prima fase ancora solo in associazione con una presa di suono microfónica e a una grossolana stima fonometrica<sup>4</sup>.

## 1. I cicli glottidali

### 1.1. Velocità di vibrazione delle pliche, frequenza fondamentale, altezza e funzioni linguistiche

Secondo il modello sorgente-filtro proposto sin da Fant (1960), l'onda di pressione acustica, il cui andamento può essere rappresentato per mezzo dell'ampiezza istantanea rilevata da un microfono (e raffigurata come oscillogramma), è il risultato di effetti risonanziali le cui caratteristiche sono ascrivibili a una sorgente sonora (vibratoria o rumorosa) e alle attività filtranti delle cavità di risonanza e irraggiamento del suono. Assumiamo per tutti i suoni sonori (*voiced*) che la

tura di c.so Raffaello, 30 (*Fisiologia*, Dip. di Neuroscienze dell'Univ. di Torino).

4. I rilievi sono avvenuti a cura dell'autrice VD sulle produzioni controllate dell'autore DS (a sua volta già coinvolto in simili esperimenti presso il *Centro di Foniatria y Logopedia* di Santander, Spagna, diretto dal Medico Foniatra Alfonso Borragán Torre). Il fonometro è un *Sound Level Meter* acquistato sugli stessi fondi dal collega Massimo Spadola Bisetti che ringraziamo in quest'occasione.

sorgente sia localizzata essenzialmente nella laringe (in particolare nella glottide, la superficie tra le pliche vocali) e che gli effetti filtranti, di caratterizzazione del timbro, siano imputabili alle configurazioni assunte dalle cavità epilaringee (condotto vocale, labbra, cavità nasali).

La responsabilità della presenza di un'armonica fondamentale (che fissa il passo di definizione dello spettro acustico del suono) è riconducibile alla ciclicità con cui si alternano configurazioni di apertura e chiusura delle rime, interlegamentosa e interaritenoidica, che definiscono lo spazio glottidale<sup>5</sup>.

La periodicità con cui si presentano questi fenomeni ciclici è dunque alla base della definizione della variabile acustica detta frequenza fondamentale ( $f_0$ ), a sua volta ritenuta all'origine delle sensazioni di altezza dei suoni sonori<sup>6</sup>.

5. In determinati assetti di fonazione, queste possono tuttavia partecipare anche alla definizione del timbro ed essere responsabili di sorgenti secondarie (rumori, sfiati etc.) che contribuiscono a definire la qualità della voce (riferimenti vari in Romano et al. 2011).

6. Ovviamente, sul piano uditivo, tutte le armoniche contribuiscono a quest'effetto e, anche in assenza della prima, se ne potrebbe avere una percezione illusoria (cfr. varie considerazioni e riferimenti in Albera 2019). Ricordiamo che in molta letteratura di stampo anglosassone, tanto la variabile acustica quanto i suoi effetti in termini di percezione sono generalmente confusi nel ricorso al termine iperonimico di *pitch*. Al lettore inesperto ricordiamo anche che  $f_0$  indica l'armonica fondamentale, mentre si usa la lettera

A differenza del canto, in cui l'altezza viene regolata in modo da raggiungere e sostenere nel tempo le note (cioè livelli culturalmente quantizzati) di una melodia prefissata con determinate qualità vibratorie, nel parlato la periodicità (e la stessa qualità) di questi cicli subisce variazioni notevoli nel passaggio da una porzione all'altra degli enunciati ed evolve solitamente nell'arco di singoli segmenti sonori e loro fasi. Questo contribuisce localmente alla categorizzazione di qualità linguistiche (fonologiche) dei segmenti sonori e alla caratterizzazione ritmico-intonativa degli enunciati: le variazioni di  $f_0$  sono infatti suscettibili di codificare relazioni sintattiche e veicolare informazioni semantiche, emotive e caratteriali del parlante.

Contrariamente alla visione generale di questa variabile che può emergere nell'immaginario del non specialista a causa delle scarse informazioni che se ne danno nella manualistica corrente, uno studio rigoroso dei caratteri linguistici dell'intonazione si scontra con le difficoltà della sua misurazione partendo dal segnale acustico, soprattutto in condizioni di basso rapporto segnale/rumore e, appunto, nelle varie condizioni che inducono il parlante a modificare non solo la velocità di

maiuscola ( $F_0$ ) per indicare l'eventuale formante associata a questa nella parte bassa dello spettro (anche se questa, nella maggior parte dei casi, coincide con una sola armonica:  $f_0$  appunto).

vibrazione delle pliche, ma gli assetti fonatori e le modalità di vibrazione all'interno di uno stesso segmento. Questo è causa di irregolarità vibratorie che inducono in errore gli algoritmi di misurazione e mettono in difficoltà lo specialista che deve decidere da un lato se i valori rilevati si riferiscano a bersagli sovrasegmentali prefissati dal parlante e dall'altro se il ruolo funzionale, in termini linguistici, dipenda dall'altezza percepita o da quella indicata dalla misurazione (indecidibile nei casi di diplofonia e in diverse condizioni di cricchiato)<sup>7</sup>.

### 1.2. Definizione del ciclo glottidale

Il monitoraggio delle caratteristiche del ciclo glottidale può avvenire ovviamente con mezzi fibroscopici, ma sono invalsi e diffusi da decenni metodi indiretti non invasivi come l'elettroglottografia (EGG) che contribuisce tradizionalmente anche una misurazione più accurata, direttamente alla sorgente e, quindi, maggiormente esente da errori indotti da effetti di filtraggio.

Di questo metodo daremo qui una dimostrazione, applicandolo a modalità fonatorie tipiche di alcuni stili di

7. In questi casi è inoltre difficile stabilire se l'altezza percepita sia quella di una prima armonica irregolarmente raccolta/rappresentata nel segnale sonoro o dell'insieme di armoniche superiori che attivano determinati *pattern* cocleari (v. n. prec.).

canto diffusi oggi nella musica commerciale occidentale. L'argomento non è nuovo, ma sono sperimentali le condizioni di osservazione e descrizione.

Pare opportuno, però, anticipare le variabili in gioco e i grafici convenzionali su cui si basano le osservazioni tradizionali in riferimento a una rappresentazione ipotetica delle condizioni fisiologiche in cui avviene un passaggio d'aria attraverso la glottide (Fig. 1, in cui le dimensioni delle frecce indicano arbitrariamente flussi d'aria, in verde, e forze di pressione, in blu).

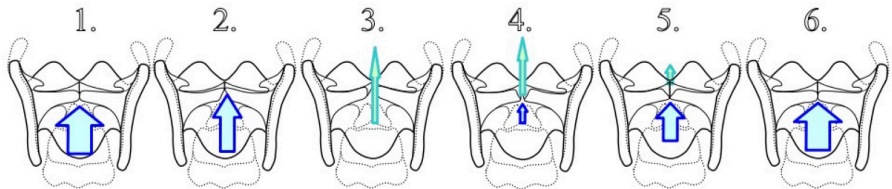


Figura 1. Fasi di vibrazione delle pliche vocali (tratto da Romano 2008 IS).

Il ciclo glottidale è ben noto sin dagli anni '50 (Fant 1960, Hirano 1989) ed è riprodotto in numerose opere scientifiche anche a carattere didattico o divulgativo.

Alle distinte fasi ricostruite schematicamente si possono associare i valori istantanei di diverse variabili fisiche rilevabili nelle forme d'onda raccolte ed elaborate con varie procedure analitiche.

A ciascuna fase del ciclo di apertura-chiusura di Fig. 1 corrisponde una variazione che si ripercuote su diverse variabili tra le quali le più studiate sono

il volume d'aria istantaneo che attraversa la glottide, forma d'onda glottidale o *fdog* (misurabile solo attraverso sonde flussometriche, ricostruibile dal segnale microfonic per *inverse filtering* v. Fig. 2a; v. Romano 1994; cfr. Ferrero et al. 1994). In corrispondenza di questo grafico proponiamo quello di Fig. 2b relativo a una tipica curva di variazione istantanea (velocità) del segnale EGG, associato a un segnale reale EGG (Fig. 2c), ottenuto in un rilievo pilota eseguito tramite EGG CSL 4400 della Kay Elemetrics, testato presso il Laboratorio di Fonetica

del DLLC dell'Univ. di Aveiro, Portogallo (diretto da Lurdes de Castro Moutinho).

Il grafico di «velocità della glottide» (Fig. 2b), basato sulla funzione  $u_g(t)$  per la quale si rimanda ora a Henrich et al. (2004) e Roubeau et al. (2009), mostra un insieme di valori relativamente stabili all'inizio del ciclo evidenziato di durata  $T_0$ ; la fase di glottide chiusa (1-2), il cui inizio è qui indicato con  $t_p$ , corrisponde infatti a velocità nulla. Si assume positiva invece nella fase di apertura, che avviene piuttosto gradualmente fino a un massimo raggiun-

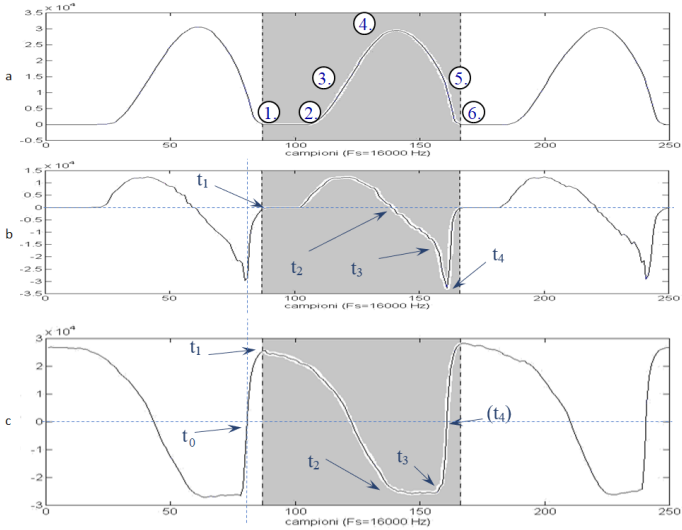


Figura 2: a) forma d'onda glottidale ( $fdog$ ) di tre cicli ricostruita per filtraggio inverso (stima del glottal flow, volume d'aria istantaneo che attraversa la glottide). L'area evidenziata stabilisce le corrispondenze con il ciclo di cui si dettagliano varie fasi in Fig. 1; b) diagramma di velocità della glottide di tre cicli (funzione  $u_g(t)$ , generalmente derivata dal segnale EGG; c) segnale elettroglottografico (EGG,  $g(t)$ ) rilevato per tre cicli simili.

to in un punto di flesso della  $fdog$  (3) e tornare a zero nel momento in cui si ha una glottide massimamente aperta ( $4 \approx t_2$ )<sup>8</sup>. Dopo un certo tempo di latenza inizia la fase di riavvicinamento con velocità moderata all'incirca costante nel verso opposto (durante una fase in cui si può considerare la glottide ancora relativamente aperta) fino a un momento in cui subisce un'accelerazione è più decisa (dall'istante  $t_3$ ) e raggiunge il suo massimo nell'ultima fase, impul-

8. I momenti in cui le pliche vocali sono ferme sono quindi due: quello, relativamente più lungo, in cui restano accollate e quello, di durata più breve e variabile, in cui raggiungono il massimo distanziamento nella porzione mediana (massima apertura della glottide).

siva, di chiusura ( $5 \approx t_4$ ) che la riporta a zero nel punto in cui (6) si torna al  $t_1$ .

Nonostante sia più promettente l'analisi della curva di velocità della glottide (funzione  $u_g(t)$ , Henrich et al. 2004) in questa prima fase di studio ci siamo concentrati sul segnale elettroglottografico (EGG),  $g(t)$ , di cui sono riprodotti tre cicli in Fig. 2c. Per un trattamento quantitativo (automatizzato) sono stati infatti proposti vari metodi di parametrizzazione proprio di questa forma d'onda (Rothenberg 1992, Marasek 1995, Vieira et al. 1996) e ci riferiamo quindi qui in particolare a questi ultimi, evidenziando nei nostri rilievi i tre punti (tre tempi) che definiscono le durate delle principali fasi osservabili

sul segnale EGG di una voce modale ( $t_1$  corrisponde al punto 3 di Marasek,  $t_2$  al 6 e  $t_3$  al punto 1; il  $t_0$  di riferimento è l'intersezione del segnale  $g(t)$  con l'asse delle ascisse)<sup>9</sup>.

## 2. Dispositivi e materiali analizzati

Come anticipato nell'introduzione, per i nostri rilievi abbiamo usato un dispositivo di rilevamento elettroglottografico (EGG), acquistato presso l'azienda *Laryngograph*. L'elettroglottografo della *Laryngograph* comprende tre collari con sensori laringometrici (elettrodi di taglia diversa) e un microfono a cravatta per il rilevamento dell'onda di pressione acustica che si diffonde nell'ambiente. Si collega direttamente tramite porta USB a un PC su cui sia installata una *Voice Suite* che prevede vari software tra cui *Speech Studio* che consente la visualizzazione in tempo reale (e differito) della forma d'onda istantanea catturata dal microfono ( $S_p$ ) e di quella EGG ( $Lx$ )<sup>10</sup>. Gli elettrodi, opportunamente posizionati sul col-

9. Alle valutazioni del quoziente di chiusura (*closing quotient*), una cui stima è ricavabile dal rapporto tra le differenze  $t_4 - t_2$  e  $t_2 - t_1$ , contiamo di aggiungere misure del quoziente di *skewing* (in riferimento "all'inclinazione" dell'onda) basate sul rapporto tra il tempo di incremento (stimato da  $t_2 - t_1$ ) e tempo di diminuzione del flusso (stimato da  $T_0 - t_2$ ).

10. Un'altra utile interfaccia operante su dati precedentemente salvati è QA che consente l'esecuzione di un'analisi quantitativa complessiva ( $QA = Quantitative Analysis$ ).

lo, in corrispondenza dei due lati della cartilagine tiroidea, registrano le rapide, minime variazioni di conduttanza tra di essi che dipendono dai cambiamenti in natura e quantità del contatto che si verifica istantaneamente tra le pliche vocali durante la fonazione.

### 2.1. Materiali

Sono state eseguite 18 prese di suono con registrazione del segnale glottografico e valutazione fonometrica. Il soggetto all'origine delle produzioni vocali è un professionista con una buona estensione vocale e con particolari abilità nell'esecuzione canora in vari assetti fonatori<sup>11</sup>. Con l'aiuto di un diapason<sup>12</sup>, nel corso di una prima registrazione, ha prodotto in succes-

11. Il soggetto riferisce tuttavia la presenza di solcature nella porzione anteriore delle sue pliche vocali, descritta per mezzo di rilievi laringostroboscopici.

12. Assumiamo qui come riferimento il LA<sub>3</sub> a 440 Hz (Anche in Italia la legge 170 del 3/5/1989 stabilisce che "Il suono di riferimento per l'intonazione di base degli strumenti musicali è la nota la<sub>3</sub>, la cui altezza deve corrispondere alla frequenza di 440 Hertz (Hz), misurata alla temperatura ambiente di 20 gradi centigradi" - *Gazzetta Ufficiale* [www.gazzettaufficiale.it/eli/id/1989/05/12/089G0200/sg]). In base a questa convenzione: RE<sub>1</sub> = 73 Hz, RE<sub>2</sub> = 147 Hz, RE<sub>3</sub> = 294 Hz, RE<sub>4</sub> = 587 Hz e RE<sub>5</sub> = 1175 Hz. I valori di frequenza associati alle singole note sono quindi quelli di Romano et al. (2012) in relazione a una classificazione novecentesca delle voci della tradizione europea del *bel canto*.

sione i seguenti 11 suoni con caratteristiche timbriche di [a] (per una durata media di 5 s ciascuno): un RE<sub>1</sub> tenuto, un RE<sub>1</sub> messa di voce, un RE<sub>2</sub> tenuto, un RE<sub>2</sub> messa di voce, un RE<sub>3</sub> tenuto, un RE<sub>3</sub> messa di voce, un RE<sub>4</sub> tenuto, un RE<sub>4</sub> messa di voce, un segmento di voce fischiata (FA#<sub>4</sub>), un glissando RE<sub>1</sub>-RE<sub>4</sub>-RE<sub>2</sub> e un glissando RE<sub>1</sub>-RE<sub>4</sub>-RE<sub>1</sub><sup>13</sup>. Tutte le registrazioni sono avvenute in presenza di un rumore di fondo stimato fonometricamente a circa 45 dB.

## 2.2. Osservazioni sperimentali

Sono state rilevate le forme d'onda riprodotte nelle figure 3-21 con le seguenti caratteristiche.

Il primo RE<sub>1</sub> tenuto ha presentato un'altezza media di 72 Hz e 62 dB<sup>14</sup>, con una forma d'onda elementare EGG caratterizzata come in Fig. 3 (con  $T_0$  pari a poco più di 13 ms). La fase di chiusura è risultata relativamen-

13. Nel corso di una seconda registrazione, il soggetto ha invece prodotto 7 modalità fonatorie: un *(vocal) fry* espiratorio, un *(vocal) fry* inspiratorio, un [a] modale (MI<sub>3</sub>), un [a] soffiato (MI<sub>3</sub>), un [a] nello stile *grunt*, un [a] nello stile *rattle* e un [a] nello stile *growl*. In questa sede dettagliamo solo gli 11 campioni modali, rinviando la discussione delle 7 produzioni non-modali a un lavoro a venire che sarà svolto in riferimento alla tassonomia proposta da Gerratt & Kreiman (2001) e alla luce delle riflessioni di Aaen et al. (2020).

14. Nel corso di tre osservazioni gli intervalli di intensità sonora sono stati di 50÷61,6; 50,3÷61,4; 49,4÷63,4 dB.

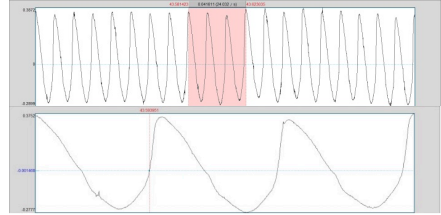


Figura 3. Segmento di forma d'onda EGG rilevata per un RE1 tenuto (circa 250 ms) con zoom su tre cicli (in basso).

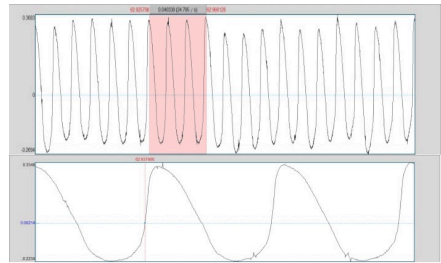


Figura 4. Segmento di forma d'onda EGG rilevata per una messa di voce su RE1 (ca. 250 ms) con zoom su tre cicli (in basso).

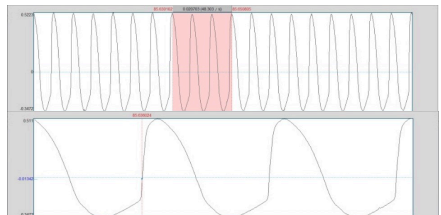


Figura 5. Segmento di forma d'onda EGG rilevata per un RE2 tenuto (circa 134 ms) con zoom su tre cicli (in basso).

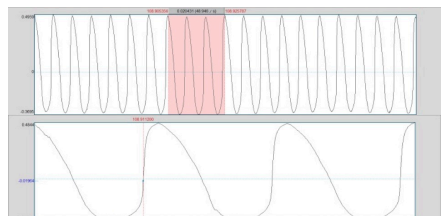


Figura 6. Segmento di forma d'onda EGG rilevata per una messa di voce su RE2 (ca. 134 ms) con zoom su tre cicli (in basso).



te veloce ( $t_1 = 1,1$  ms) con una tenuta della chiusura istantanea e una fase di apertura più lunga ( $t_2 = 9,3$  ms).

La messa di voce sul RE<sub>1</sub> ha presentato un'altezza con maggiori oscillazioni (tra 66 e 82 Hz con *jitter* = 1,48%) e una media di 74 Hz, ma con energia crescente (da 56 a 66 dB nelle fasi di maggiore stabilità, *shimmer* = 4,46%)<sup>15</sup>. In una porzione centrale, la forma d'onda elementare EGG è caratterizzata, come si può vedere in Fig. 4 da un  $T_0$  pari a poco più di 13 ms (cfr. Fig. 3). Anche in questo caso la fase di chiusura è risultata relativamente veloce ( $t_1 = 0,9$  ms) con una tenuta della chiusura istantanea e una fase di apertura che si esaurisce a  $t_2 = 8,4$  ms.

Il primo RE<sub>2</sub> tenuto ha presentato un'altezza media di 144 Hz (126÷158) e intensità 42÷72,2 dB, con una forma d'onda elementare EGG caratterizzata come in Fig. 5 (con  $T_0$  pari a 7 ms). La fase di chiusura presenta in questo

caso due diverse velocità ( $t_1 = 0,85$  ms, ma in presenza di un  $t_3 = 1,8$  ms e fase di massimizzazione della velocità di durata inferiore a 1 ms) con una tenuta della chiusura istantanea e una fase di apertura più lunga ( $t_2 = 4,8$  ms).

La messa di voce del RE<sub>2</sub> ha presentato un'altezza media di 146 Hz (nell'intervallo 139÷158 Hz) e intensità variabile (nell'intervallo 49÷84,3 dB), con una forma d'onda elementare EGG caratterizzata come in Fig. 6 (con  $T_0$  pari a circa 7 ms e  $t_3 \approx t_4$ ). La fase di glottide chiusa si presenta come nel caso di Fig. 4 con fasi di chiusura e apertura relativamente veloci (con tempi dimezzati:  $t_1 = 0,85$  ms,  $t_2 = 5$  ms e fase di massimizzazione della velocità da  $t_3 = 6,8$  ms)<sup>16</sup>.

Il RE<sub>3</sub> tenuto ha presentato un'altezza media di 295 Hz (min 261 max 303, per via di un transitorio iniziale più instabile, con intensità che si stabilizza su circa 55 dB). La forma d'onda elementare EGG si caratterizza come in Fig. 7 (con  $T_0$  pari a 3,5 ms). Le fasi di chiusura e apertura tendono a presentare durate simili (con  $t_1 = 0,6$  ms e  $t_2 = 2,5$  ms). Seppur visibile, la

15. Le misure di *jitter* (microperturbazioni della frequenza fondamentale) e di *shimmer* (microperturbazioni dell'ampiezza dell'onda) sono state rilevate grazie al *Voice Report* di PRAAT (lo stesso software usato per ottenere i grafici qui riprodotti). I valori rilevati sono generalmente al di sotto delle soglie tipiche rilevate per voci eufoniche di altri spazi linguistici (Brockmann et al. 2009). Un confronto sarebbe auspicabile con i valori riportati da fonti italiane. Quelle da noi reperite si basano però su altre variabili e offrono un riferimento meno agevole che rinviamo a studi più approfonditi.

16. Il confronto tra tenuta e messa di voce in questo caso non mostra le stesse differenze osservabili per RE<sub>1</sub> e RE<sub>3</sub> (v. dopo): la riduzione della fase di apertura è minima e non consente di osservare l'aumento della velocità della chiusura che sarebbe stato in accordo con l'aumento di pressione (e flusso) che si presume avvenga durante la messa di voce.



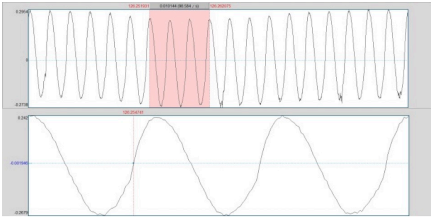


Figura 7. Segmento di forma d'onda EGG rilevata per un RE3 tenuto (circa 68 ms) con zoom su tre cicli (in basso).

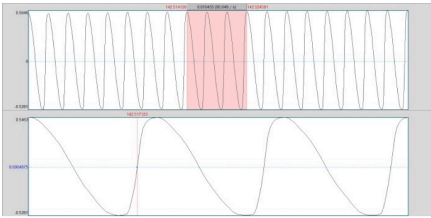


Figura 8. Segmento di forma d'onda EGG rilevata per una messa di voce su RE3 (ca. 66 ms) con zoom su tre cicli (in basso).

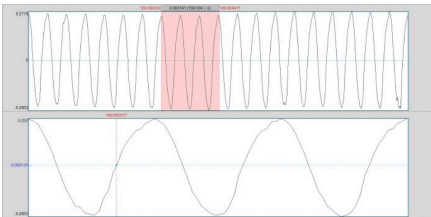


Figura 9. Segmento di forma d'onda EGG rilevata per un RE4 tenuto (ca. 33 ms) con zoom su tre cicli (in basso).

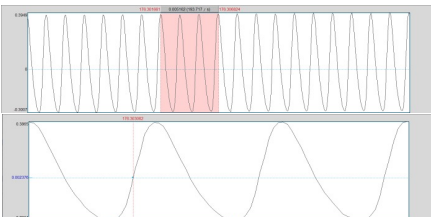


Figura 10. Segmento di forma d'onda EGG rilevata per una messa di voce su RE4 (ca. 34 ms) con zoom su tre cicli (in basso).

presenza di un punto di aumento di velocità nella fase di chiusura ( $t_3$ ) risulta poco rilevante rispetto ai casi visti sopra. Tuttavia il secondo rilevamento di questo RE<sub>3</sub>, in una messa di voce (con intensità variabile 49,1÷88,2 dB), ha prodotto un risultato lievemente diverso (v. Fig. 8;  $T_0 \sim 3,5$  ms): anche a queste frequenze, infatti, a un tempo di chiusura ridotto ( $t_1 \sim 0,6$  ms) si associano tempi di apertura più lunghi ( $t_2$  dell'ordine dei 2,9 ms).

Un'osservazione variabile si conferma anche tra RE<sub>4</sub> tenuto e messa di voce. Se nel caso del primo si ha un transitorio iniziale particolarmente mosso (altezza media 588 Hz, min 499 max 596, e intensità 56,1÷83,5 dB), la forma d'onda si caratterizza come in Fig. 9 e differisce da quella del secondo (Fig. 10) che ha una media 586 Hz e minimi e massimi (rispettivamente 563 e 656) distanti anche per via di particolari instabilità cicliche (associate a uno *shimmer* più elevato, 1,33%<sup>17</sup>). In entrambi i casi si ha  $T_0$  pari a 1,6 ms ma nel primo caso la fase di chiusura si conclude a  $t_1 \sim 0,5$  ms mentre quella di apertura culmina a  $t_2 \sim 1,3$  ms; nel secondo – compatibilmente con quanto osservato nel caso di RE<sub>1</sub> e RE<sub>3</sub> – la proporzione cambia, con un tempo di chiusura più corto ( $t_1 = 0,6$  ms e  $t_2 = 1,4$  ms circa).

17. L'intensità rilevata dal fonometro è infatti nell'intervallo 75÷91,2 dB.

Una voce fischiata è infine stata osservata a circa 734 Hz (FA#<sub>4</sub>,  $T_0 = 1,4$  ms circa, v. Fig. 11) con intensità variabile ma generalmente piuttosto elevata (70,1÷78 dB). A un tempo di chiusura piuttosto breve ( $t_1 = 0,5$  ms) si associa un tempo di apertura più lungo ( $t_2 = 1,1$  ms) con proporzioni simili a quelle già descritte per l'esempio di Fig. 10 (v. §1.2).

In merito alle osservazioni svolte sui glissando, ci riferiamo qui (Figg. 12 e

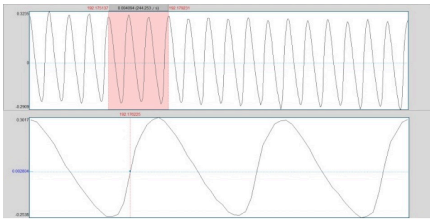


Figura 11. Segmento di EGG rilevato per una voce fischiata su FA#4 (ca. 26 ms) con zoom su tre cicli (in basso).

13) solo alla fase crescente del primo e a quella decrescente del secondo. Il primo glissando presenta un progressivo aumento di intensità sonora da 48,3 a 83,7 dB e una successiva diminuzione fino a 65,2 dB, con rilevanti perturbazioni nella cosiddetta fase di *passaggio*<sup>18</sup>.

Queste instabilità sono state studiate recentemente proprio su voci teno-

18. Si noti che, malgrado l'energia sonora rilevata dal fonometro presenti questi valori, la misura stimata da PRAAT sul segnale microfonico varia da circa 60 a circa 82 dB, mentre quella rilevata sull'EGG resta all'incirca co-

rili da Echternach et al. (2017) e Selamtzis (2018) che le hanno osservate e descritte a partire dal FA<sub>4</sub> (circa 350 Hz) fino a una frequenza individuale che stabilisce un altro intervallo stabile (almeno in termini di intensità) nel quale si definisce un tipo vocale detto *SVaP* (*Stage Voice above the Passaggio*)<sup>19</sup>. Si riportano in Fig. 12 due campioni di cicli raccolti rispettivamente nella fase grave iniziale ( $T_0 = 13$  ms circa, con  $t_1 = 1$  ms,  $t_2 = 8,6$  e  $t_3 = 12$  ms) e di tenuta centrale ( $T_0 = 1,7$  ms circa, con  $t_1 = 0,6$  ms e  $t_2 = 1,1$  ms).

La parte decrescente del secondo glissando ha presentato una progressiva diminuzione di intensità sonora da

stante, sugli 82 dB, con un'unica flessione (min. 75 dB) in corrispondenza del *passaggio*. Discrepanze simili si osservano tra i valori rilevati per il secondo glissando.

19. Le qualità performative degli interpreti a questo riguardo sono oggetto delle preoccupazioni di C. Ambrosini che, in riferimento alla tradizione colta europea, nel suo contributo dal titolo “Voce cantata e condizionamenti linguistici” (in Ambrosini et al. 2013) si pone il problema di comporre un'aria tenendo conto delle sue caratteristiche promeliche o antimeliche, in funzione delle possibilità dell'interpretazione (definendo condizioni in cui questa possa definirsi comoda e congeniale alla vocalità del cantante, proprio in termini di tessitura e aree di passaggio tra i registri). Nella stessa sede si veda anche l'intervento di G. Tisato che discute della non-linearità dei meccanismi fonatori e analizza le possibilità classificatorie delle modalità vocali documentate in alcune tradizioni ed esplorate in applicazioni sperimentali (soprattutto sin dai primi contributi di D. Stratos).

circa 82 dB a 48 dB. Si riportano in Fig. 13 due campioni di cicli raccolti rispettivamente in corrispondenza della nota acuta iniziale ( $T_0 = 1,6$  ms circa, con  $t_1 = 0,5$  ms,  $t_2 = 1,1$  ms) e di quella finale ( $T_0 = 13,6$  ms circa, con  $t_1 = 1,7$  ms,  $t_2 = 8,5$  ms e  $t_3 = 11,4$  ms). In questo caso, nel *passaggio*, nel corso della lunga fase d'instabilità di ampiezza/intensità sonora, l'algoritmo di stima di  $f_0$  usato in PRAAT ha prodotto anche alcuni errori di rilevamento nell'ultima fase di transizione verso RE<sub>1</sub>.

Alla somiglianza tra le forme d'onda elementari dell'EGG nel caso dei due RE<sub>1</sub> si contrappone la marcata differenziazione di quelle dei due RE<sub>4</sub>.

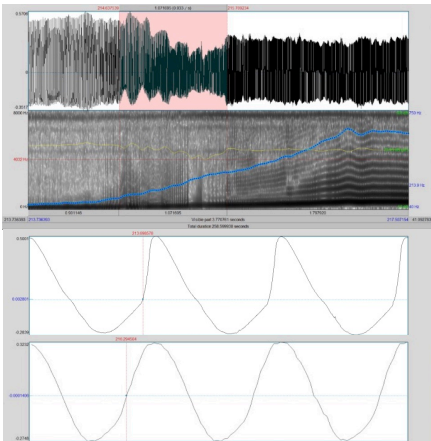


Figura 12. Oscillogramma e spettrogramma dell'onda EGG della fase crescente del primo glissando (in alto). Si noti la lunga instabilità di ampiezza/energia in corrispondenza del passaggio (evidenziato). Al centro, zoom su tre cicli rilevati nel corso del RE1 iniziale; in basso, zoom su tre cicli del RE4 finale.

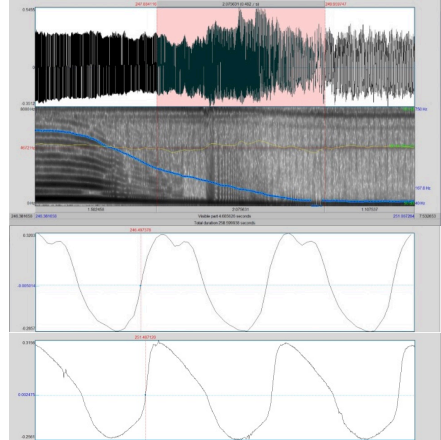


Figura 13. Oscillogramma e spettrogramma della fase decrescente del secondo glissando (in alto). Si noti la lunga instabilità di ampiezza/intensità sonora in corrispondenza del passaggio (evidenziato). Al centro, uno zoom su tre cicli di EGG in corrispondenza della nota più alta (RE4); in basso, zoom su tre cicli in corrispondenza di quella più bassa (RE1).

### 3. Interpretazione e conclusioni parziali

Sulle emissioni vocali controllate di voce modale qui presentate è stata eseguita una prima accurata osservazione dei tempi con cui si organizza la forma d'onda EGG alle varie frequenze a cui abbiamo associato una prima valutazione impressionistica delle possibilità di ritrovare una stessa forma d'onda elementare EGG per una stessa nota prodotta da uno stesso cantante in condizioni diverse.

Nonostante la disponibilità delle misure di una selezione dei tempi parametrizzati in letteratura, non è stata ancora condotta la discussione sui rapporti tra fasi di durata e di chiusura

(sulla base di soglie che restano da stabilire). Data però la presenza di profili con un diverso distanziamento degli eventi rilevabili nelle forme d'onda elementari analizzate (cfr. ad es. Fig. 5 e Fig. 12 destra) sembrerebbe evidente che si tratti di una variabilità tale da rendere necessario il supporto di altre variabili, come i volumi e le strategie respiratorie (Folino et al. 2017) o l'ispezione di curve derivate più facilmente caratterizzabili (Henrich et al. 2004). Una rappresentazione multidimensionale potrà includere anche osservazioni svolte con modalità di presentazione dei dati che prevedano il ricorso al fonetogramma (Ginepro 2022) in grado di tracciare le dinamiche che consentono il raggiungimento e la tenuta dell'altezza osservata.

Anche tenendo conto di alcune anomalie rilevate (come quella dell'oscillazione nei valori della fase di glottide chiusa di Fig. 12 sinistra, ancora da quantificare adeguatamente in base a un'opportuna definizione delle soglie), resta comunque ancora impressionistica l'osservazione di un principio di minor tempo di chiusura (cioè valori inferiori del *closing quotient*) associato alle tonalità più alte.

Un altro fatto notevole riguarda le proporzioni generali che regolano i rapporti tra le misurazioni della qualità EGG di una nota e quelle della nota corrispondente tre ottave più in alto. Passando ad es. da una forma d'onda

elementare con un  $T_0 = 13$  ms a una con un  $T_0 = 1,6-1,7$  ms si arriva a condizioni in cui l'errore di misura non è più trascurabile.

Un tentativo di interpretazione più organica delle variazioni osservate è quindi rimandato alle successive fasi di lavoro in cui, a un'osservazione più accurata dei *passaggi* (da un meccanismo all'altro), sarà associata l'analisi dei campioni di voce cantata prodotta in condizioni non-modal di fonazione tipiche delle voci estreme (sul modello di Aaen et. al. 2020).

## Bibliografia

Aaen M., McGlashan J. & Sadolin C. (2020). «Laryngostroboscopic Exploration of Rough Vocal Effects in Singing and their Statistical Recognizability: An Anatomical and Physiological Description and Visual Recognizability Study of Distortion, Growl, Rattle, and Grunt using laryngostroboscopic Imaging and Panel Assessment». *Journal of Voice*, 34(1), 162.e5-14.

Abberton E. & Fourcin A. (1997), «Electrolaryngography». In Martin J. Ball & Chris Code (eds.), *Instrumental Clinical Phonetics*, London: Whurr, 119-148.

Ambrosini C., Bravi P., Proto T., Tisato G. & Romano A. (2013). «Speaking voice, singing voice, and performance». In: V. Galatà (a cura di), *Multimodalità e multilinguallità: la sfida più avanzata della comunicazione orale* (Atti del IX Convegno Nazionale dell'Asso-

ciazione Italiana di Scienze della Voce, Venezia 21-23/01/2013), Roma: Bulzoni, 3-11 (abstract p. 3).

Brockmann M., Drinnan M.J., Carding P.N. (2009). «Comparison of voice intensity effects on electroglottographic versus acoustic jitter and shimmer». In: *Proc. of the 157<sup>th</sup> meeting of the Acoustical Society of America* (Portland, Oregon, USA, April 2009), 2532.

Celata C. & Ricci I. (2021). «Electrolaryngography/Electroglottography». In Martin J. Ball (ed.), *Manual of Clinical Phonetics*, London: Routledge, 306-321.

Echternach, M., Burk, F., Köberlein, M., Herbst, C. T., Döllinger, M., Burdumy, M., and Richter, B. (2017). Oscillatory Characteristics of the Vocal Folds Across the Tenor Passaggio. *Journal of Voice*, 31(3), 381.e5-e14.

Fant G. (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton (new ed. Dordrecht: Kluwer, 2004).

Ferrero F., Pierucci P. & Mesirca A. (1994). «Determinazione automatica degli istanti di chiusura glottale», *Rivista Italiana di Acustica*, 18, 19-26.

Folino A., Benzo M., Pasquero P., Laguzzi A., Mesin L., Messere A., Porta M., Roatta S. (2017). Vena cava responsiveness to controlled isovolumetric respiratory efforts, *J. Ultras Med.*, 36, 2113-2123 (Q2 Scopus Radiology, Nuclear medicine).

Fourcin A. (1974). «Laryngographic Examination of Vocal Fold Vibration, Ventilatory and Phonatory Control

Mechanisms». In B. Wyke (ed.), *Ventilatory and Phonatory Control Systems* (International Symposium, Papers 19), London: OUP, 316-333.

Fourcin A. (1986). «Electrolaryngographic Assessment of Vocal Fold Function», *Journal of Phonetics*, 14, 435-442.

Fourcin A. (2005). «Clinical Voice Measurement Using Egg/Lx Signals». In C. Manfredi (ed.), *Models and Analysis of Vocal Emissions for Biomedical Applications. Proc. of the 4<sup>th</sup> International Workshop* (Firenze, 29-31 Oct. 2005), 111-114.

Fourcin A. & Abberton E. (1971). «First application of a new laryngograph». *Med. Biol. Ill.*, 21, 172-182,

Gerratt B. & Kreiman J. (2001). Toward a taxonomy of nonmodal phonation, *Journal of Phonetics*, 29, 365-381.

Ginepro R. (2022). «L'impiego dei fonetogrammi nell'analisi di voci cantate in molteplici lingue». Univ. di Torino: Tesi di Laurea in Scienze linguistiche (rel. Antonio Romano).

Henrich Bernardoni N. (2021). La voce umana, dal respiro al canto. *Bollettino del LFSAG*, 7, 43-57 [www.lfsag.unito.it/ricerca/phonews/07/7\_3.pdf]

Henrich N., d'Alessandro Chr., Doval B. & Castellengo M. (2004). On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *The Journal of the Acoustical Society of America*, 115, 1321-1332.

Marasek K. (1995). «A semi-automated Lx-based method for the measurement of voice onset time», *Proc. of XIIIth ICPHS* (Stockholm, 14-19 August 1995), vol. 3, 310-313.

Marasek K. (1997) *EGG & Voice quality*: <https://www2.ims.uni-stuttgart.de/EGG/>

Romano A. (1994). Analyse de la fonction de transfert du conduit vocal dans les intervalles de fermeture de la glotte et extraction de la forme d'onde d'excitation glottale par filtrage inverse / Analisi della funzione di trasferimento del condotto vocale durante gli intervalli di chiusura della glottide ed estrazione della forma d'onda glottidale per filtraggio inverso, Grenoble: Institut de la Communication Parlée/Politecnico di Torino (resp. G. Feng & E. Biglieri).

Romano A., Cesari U., Mignano M., Schindler O. & Vernerio I. (2012). «Voice Quality» / «La qualità della voce». In: A. Paoloni & M. Falcone (a cura di), *La voce nelle applicazioni* (Atti

dell'VIII Convegno dell'Associazione Italiana Scienze della Voce, Roma, 25-27 gennaio 2012), Roma: Bulzoni, 75 (art. int. CD 35 pp.)

Rothenberg M. (1992). «A multi-channel electroglottograph», *Journal of Voice*, 6 (1), 36-43.

Roubeau B., Henrich N. & Castellingo M. (2009). Laryngeal Vibratory Mechanisms: The Notion of Vocal Register Revisited. *Journal of Voice*, Vol. 23, No. 4, 425-438.

Selamtzis A. (2018). Analyses of voice and glottographic signals in singing and speech. Doctoral Thesis in Speech and Music Communication, Stockholm: KTH Royal Institute of Technology School of Electrical Engineering and Computer Science

Vieira M.N., McInnes F.R. & Jack M.A. (1996). «Analysis of the effects of electroglottographic baseline fluctuation on the F0 estimation in pathological voices», *The Journal of the Acoustical Society of America*, 99, 3171-3178.