

PROCEEDINGS E REPORT

ISSN 2704-601X (PRINT) | ISSN 2704-5846 (ONLINE)

**MODELS AND ANALYSIS OF VOCAL
EMISSIONS FOR BIOMEDICAL
APPLICATIONS**

11th INTERNATIONAL WORKSHOP

**December 17-19, 2019
Firenze, Italy**

**Edited by
Claudia Manfredi**

Firenze University Press
2019

Models and Analysis of Vocal Emissions for Biomedical Applications : 11th International Workshop, December, 17-19, 2019 / edited by Claudia Manfredi. – Firenze : Firenze University Press, 2019.
(Proceedings e report ; 122)

<https://www.fupress.com/isbn/9788864539454>

ISSN 2704-601X (print)

ISSN 2704-5846 (online)

ISBN 978-88-6453-951-5 (print)

ISBN 978-88-6453-961-4 (online)

Cover: designed by CdC, Firenze, Italy.

Peer Review Process

All publications are submitted to an external refereeing process under the responsibility of the FUP Editorial Board and the Scientific Committees of the individual series. The works published in the FUP catalogue are evaluated and approved by the Editorial Board of the publishing house. For a more detailed description of the refereeing process we refer to the official documents published on the website and in the online catalogue (www.fupress.com).

Firenze University Press Editorial Board

M. Garzaniti (Editor-in-Chief), M. Boddi, A. Bucelli, R. Casalbuoni, A. Dolfi, R. Ferrise, M.C. Grisolia, P. Guarnieri, R. Lanfredini, P. Lo Nostro, G. Mari, A. Mariani, P.M. Mariano, S. Marinai, R. Minuti, P. Nanni, G. Nigro, A. Perulli.

📄 The on-line digital edition is published in Open Access on www.fupress.com.

The present work is released under Creative Commons Attribution 4.0 International license (CC BY 4.0: <http://creativecommons.org/licenses/by/4.0/legalcode>). This license allows you to share any part of the work by any means and format, modify it for any purpose, including commercial, as long as appropriate credit is given to the author, any changes made to the work are indicated and a URL link is provided to the license.

© 2019 Firenze University Press
Published by Firenze University Press
Firenze University Press
Università degli Studi di Firenze
via Cittadella, 7, 50144 Firenze, Italy
www.fupress.com

*This book is printed on acid-free paper
Printed in Italy*



MAVEBA 2019

Firenze, Italy

The MAVeBA 2019 Workshop is sponsored by:

Università degli Studi di Firenze
Department of Information Engineering - DINFO



Int. Journal Biomedical Signal Processing and Control, Elsevier I.t.d

and is supported by:



UNIVERSITÀ
DEGLI STUDI
FIRENZE
DINFO
DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE

Department of Information Engineering - DINFO



FONDAZIONE
CR FIRENZE

Fondazione Cassa di Risparmio di Firenze - FCRF

REGIONE TOSCANA



Consiglio Regionale

Regione Toscana - Consiglio Regionale



CoMeT Collegium Medicorum Theatri

CONTENTS

Foreword	XIII
----------------	------

SESSION I – PARKINSON’S DISEASE AND SPEECH

FUNDAMENTAL FREQUENCY AND DURATION CUES FOR SENTENCE FOCUS IN MANDARIN: PARKINSON’S AND HEALTHY SPEAKERS	17
Xi Chen, Diana Sidtis	

TEMPORAL REVERSION OF PHONATION INSTABILITY IN PARKINSON’S DISEASE BY NEUROACOUSTICAL STIMULATION	21
Pedro Gómez-Vilda, Gerardo Gálvez-García, Andrés Gómez-Rodellar, Daniel Palacios-Alonso, Guillermo de Arcas-Castro	

A NEUROMECHANICAL MODEL OF JAW-TONGUE ARTICULATION IN PARKINSON’S DISEASE SPEECH	25
A.Gómez, A. Tsanas, P. Gómez, D. Palacios, A. Álvarez, R. Martínez	

PARKINSON’S DISEASE CLASSIFICATION BASED ON VOWEL SOUND	29
Daria Hemmerling, David Sztaho	

ANALYSIS OF PHONATORY FEATURES FOR THE AUTOMATIC DETECTION OF PARKINSON’S DISEASE IN TWO DIFFERENT CORPORA	33
Laureano Moro-Velázquez, Jorge Andrés Gómez-García, Najim Dehak, Juan Ignacio Godino-Llorente	

JOINT ANALYSIS OF VOCAL JITTER, FLUTTER AND TREMOR IN VOWELS SUSTAINED BY NORMOPHONIC AND PARKINSON SPEAKERS	37
J. Schoentgen, A. Kacha, F. Grenez	

THE EFFECTS OF DEEP BRAIN STIMULATION ON SPEECH ARTICULATION AND VOCALIZATION IN PARKINSON’S DISEASE	41
John J. Sidtis, Diana Van Lancker Sidtis	

BIOMEDICAL SPEECH SIGNAL INSIGHTS FROM A LARGE SCALE COHORT ACROSS SEVEN COUNTRIES: THE PARKINSON’S VOICE INITIATIVE STUDY	45
Athanasios Tsanas, Siddharth Arora	

SESSION II – SINGING VOICE

TOWARDS A SOMATOSENSORY TRAINING DIGITAL ENVIRONMENT FOR LYRIC SINGING PEDAGOGY	51
Angelakis Evangelos, Georgaki Anastasia	

ASSESSING ARTICULATION IN SINGING DURING A VOWEL MATCHING EXERCISE USING ARTICULOGRAPHY: A PILOT STUDY	55
Helena Daffern, Amelia Gully	

VIII

TUNING TENDENCIES IN A SINGING QUINTET: EVOLUTION ACROSS REHEARSALS 59
Sara D'Amario, David M. Howard

FACE VIBRATIONS MEASUREMENT IN SINGING — PILOT STUDY 63
Marek Frič, Pavel Dlask, Viktor Hruška

COMPARISON OF SOUND RADIATION BETWEEN POP AND CLASSICAL SINGERS 67
Iva Podzimková, Marek Frič

DISSOCIATION OF SPOKEN AND SUNG VOCAL PRODUCTION 71
Diana Van Lancker Sidtis, Y-J. Kim, John J. Sidtis

SHOULD I OPEN MY MOUTH MORE TO SING LOUDER? 75
Allan Vurma

SESSION III - SPECIAL SESSION - EDUCATION AND REHABILITATION OF THE ARTISTIC VOICE

SOVTE: BETWEEN MYTH AND REALITY 81
F. Fussi

THE SOVTE PROTOCOLS: INDICATIONS AND CRITICISM IN VARIOUS SINGING STYLES 81
E. Biavati, E. Bruni

G.E.M.M.A. TRAINING: A LINK BETWEEN PSYCHOLOGY AND LOGOPAEDICS IN THE TREATMENT OF SINGING VOICE DISORDERS 81
E. Rosa

SESSION IV - VOICE AND EMOTIONS

THE CORRELATION BETWEEN POETIC RHYTHM AND HEART RATE VARIABILITY IN SUBJECTS READING AND PERCEIVING RUSSIAN POETRY 85
Karina Evgrafova, Pavel Skrelin, Vera Evdokimova, Tatjana Chukaeva

EMOTIONALLY EXPRESSED VOICES ARE RETAINED IN MEMORY FOLLOWING A SINGLE EXPOSURE 89
Y-J. Kim, J. J. Sidtis, D. Van Lancker Sidtis

ELECTRODERMAL ACTIVITY AND SPEECH FEATURES AS PREDICTORS FOR AROUSAL LEVEL CHANGES AFTER AFFECTIVE WORD PRONUNCIATION" 93
Nicola Vanello, Alberto Greco, Claudia Marzi, Enzo Pasquale Scilingo

A COMMONLY SHARED CODE FOR SADNESS IN HUMAN VOCALIZATIONS AND MUSIC REVEALED BY HUMAN INFANT CRIES 97
G. Zeloni, F. Pavani

SESSION V - VOICE QUALITY

TOWARDS ROBUST FEATURES IN VOICE DISEASE DIAGNOSIS: MFCCs VS. PNCCs 103
M. Madruga, Y. Campos-Roca, C. J. Pérez

ELECTROGLOTTOGRAPHIC VOICE MAPS OF UNTRAINED VOCALLY HEALTHY ADULTS, WITH GENDER DIFFERENCES AND GRADIENTS	107
Sten Ternström, Rita R. Patel	

ACOUSTIC AND ELECTROGLOTTOGRAPHIC PARAMETRISATION OF PHONATORY QUALITY PROVIDE VOICE PROFILES OF PATHOLOGICAL SPEAKERS	111
Manfred Pützer, Wolfgang Wokurek	

VOCAL QUALITIES OF SARCASTIC UTTERANCES: CROSS-LINGUISTIC STUDY OF ENGLISH AND KOREAN	115
Seung-yun Yang	

SESSION VI - VOCAL FOLDS DYNAMICS

PERTURBATION OF CYCLE LENGTHS AND CYCLE PEAK AMPLITUDES IN DIPLOPHONIC VOICES	121
Philipp Aichinger	

TRACKING OF MULTIPLE FUNDAMENTAL FREQUENCIES IN STANDARD TEXT READINGS OF DIPLOPHONIC SPEAKERS.....	125
Philipp Aichinger	

AERODYNAMICS OF GLOTTAL VIBRATION ONSET	129
P. H. DeJonckere, J. Lebacqz	

A GLOTTAL AREA WAVEFORM MODEL FOR MULTI-PULSED VOCAL FRY.....	133
Vinod Devaraj, P. Aichinger	

MODELLING LONGITUDINAL PHASE DIFFERENCES IN A LUMPED AND DISTRIBUTED ELEMENTS VOCAL FOLD MODEL	137
Carlo Drioli, Philipp Aichinger	

EXTRACTING VOCAL FOLD KINEMATIC PARAMETERS FROM VIDEOKYMOGRAMS VIA SIMULATION OF CLINICALLY OBSERVED DATA.....	141
Sridhar Bulusu, S. Pravin Kumar, Jan G. Svec, Philipp Aichinger	

VOCAL FOLD OSCILLATORS AT LARGE ASYMMETRIES	145
Jorge C. Lucero, Xavier Pelorson, Annemie Van Hirtum	

PHYSICAL STUDY OF THE INFLUENCE OF LEFT-RIGHT VOCAL FOLDS ANGLE ASYMMETRY ON PHONATION.....	149
Annemie Van Hirtum, Anne Bouvet, Xavier Pelorson, Isao Tokuda	

THEORETICAL AND EXPERIMENTAL MODELING OF LESIONS OF THE VOCAL FOLDS.....	153
Xavier Pelorson, Anne Bouvet, Annemie Van Hirtum	

VOICE SOURCE EFFECTS OF FUNDAMENTAL FREQUENCY VARIATION.....	157
Johan Sundberg	

SESSION VII - COMET SESSION - ACTOR'S AND ACTRESS' VOICES

SPECTRAL SPECIFICITIES OF ACTING VOICE IN PROFESSIONAL ACTRESSES	163
P.H. DeJonckere, H. Stoffels	
SHORT TERM EFFECT OF 'SEMI-OCCLUDED VOCAL TRACT EXERCISES' ON HEALTHY ACTORS' VOICES	167
Valentina Di Natale, Giovanna Cantarella, Claudia Manfredi, Annaclara Ciabatta, Cosimo Becherini, Philippe DeJonckere	
SINGING WHILE ACTING AND VICE-VERSA	171
Orietta Calcinoni	
THE SHOUTING VOICE AND THE IMPORTANCE OF SHOUTING ABILITY AS A 'FITNESS' PARAMETER FOR ALL VOICE USERS.....	175
Josef Schlömicher - Thier, Hannes Tropper, Ingolf Franke	
THE ACTORS VOICE: LAUGHING, CRYING AND SHOUTING IN THE MRI	179
Bernhard Richter, Louisa Traser, Michael Burdumy, Matthias Echternach, Claudia Spahn	
DISTORTED VOCALITY: SUPRAGLOTTIC MANAGEMENT IN ACTING AND DUBBING	181
Franco Fussi, Eleonora Bruni	

SESSION VIII – VOCAL FOLDS PARALYSIS/ABNORMALITIES

PHONOSURGICAL TREATMENT OF BILATERAL LARYNGEAL PARALYSES.....	185
Giovanna Cantarella	
SELECTIVE SURFACE STIMULATION IN UNILATERAL VOCAL FOLD PARALYSIS (UVFP)...	189
Berit Schneider-Stickler, Matthias Leonhard, Matthias Krenn, Winfried Mayr	
ULTRAHIGH RESOLUTION OPTICAL COHERENCE TOMOGRAPHY FOR DETECTING TISSUE ABNORMALITIES OF THE ORAL AND LARYNGEAL MUCOSA: A PRELIMINARY STUDY	195
Niels Møller Israelsen, Anders Overgård Jønsson, Mette Pedersen	

SESSION IX - KEYNOTE LECTURE

DEVELOPING NEW SPEECH SIGNAL PROCESSING ALGORITHMS FOR BIOMEDICAL AND LIFE SCIENCES APPLICATIONS: PRINCIPLES, FINDINGS, CHALLENGES, AND A VIEW TO THE FUTURE.....	201
Athanasios Tsanas	

SESSION X - BIOMECHANICS/DEVICES

EXPERIMENTAL MODELLING OF GLOTTAL AREA DECLINATION RATE IN VOWEL AND RESONANCE TUBE PHONATION.....	205
Jaromír Horáček, Radolf Vojtěch, Bula Viteslav, Anne-Maria Laukkanen	

DEVELOPMENT AND USE OF AN ANECHOIC SUBGLOTTAL TRACT FOR EXCISED LARYNX EXPERIMENTS.....	209
Hugo Lehoux, Vít Hampala, Jan G. Švec	

SHAKER: PRELIMINARY OBSERVATIONS OF A POTENTIAL DEVICE FOR VOICE TRAINING AND THERAPY.....	213
Anne-Maria Laukkanen , Jaromír Horáček , Vojtěch Radolf	

SESSION XI - SPEECH

BEATBOX SOUNDS RECOGNITION USING A SPEECH-DEDICATED HMM-GMM BASED SYSTEM.....	219
Solene Evain, Adrien Contesse, Antoine Pinchaud, Didier Schwab, Benjamin Lecouteux, Nathalie Henrich Bernardoni	

AM-FM DECOMPOSITION OF SPEECH SIGNAL: APPLICATIONS FOR SPEECH PRIVACY AND DIAGNOSIS.....	223
Petr Motliceck, Hynek Hermansky, Srikanth Madikeri, Amrutha Prasad, Sriram Ganapathy	

THE TOMATIS ELECTRONIC EAR EFFECTS ON SIMPLE VOCALIZATIONS	227
Marco Prenassi, Walter Coppola, Giovanni Ramponi, Tiziano Agostini, Sara Marceglia	

SESSION XII - VOICE VS OTHER PHYSIOLOGICAL SIGNALS/DISEASES

THE INTELLIGIBILITY OF SPEECH IN POLISH SPEAKERS BORN WITH CLEFT LIP AND/OR PALATE.....	233
Wiktor Gonet, Maria Hortis-Dzierzbicka, Edyta Zomkowska	

THE EFFECT OF PHYSICAL ACTIVITY ON THE PHONETIC CHARACTERISTICS OF SPEECH.....	237
V. V. Evdokimova, E. A. Zakharchenko	

CHANGES TO VOICE PRODUCTION CAUSED BY LONG-TERM HEARING LOSS (HL).....	241
I.V. Kastyro, A.N. Kovalenko, Torshin V.I., Draginskaya E.S.	

COMPARISON OF IMMEDIATE EFFECTS OF VOCAL BREATHING EXERCISES AND PHYSICAL EXERCISES ON HEART RATE VARIABILITY (HRV) IN HEALTHY STUDENTS	245
A.N. Kovalenko, I.V. Kastyro, V.I. Torshin, Y.S. Guhschina, E.S. Doroginskaya, N.A. Kamanina	

AN ALGORITHM FOR DETECTING THE ONSET OF LINGUISTIC SEGMENTS IN CONTINUOUS ELECTROENCEPHALOGRAM SIGNALS	249
C. Tonatiuh Hernández-del-Toro, Carlos A. Reyes-García	

DISCRIMINATION BETWEEN CHILDREN AND ADULT FACES USING BODY AND HEAD RATIO AND GEOMETRIC FEATURES	253
C.A. Reyes-García, E. Morales-Vargas, H. Peregrina-Barreto, C. Manfredi	

SONIFICATION TECHNIQUES APPLIED TO EEG SIGNALS OF NONMOTOR GENERALIZED ONSET EPILEPTIC SEIZURES	257
L. Frassinetti, R. Guerrini, C. Barba, F. Melani, F. Piras, C. Manfredi	

BIOVOICE: A MULTIPURPOSE TOOL FOR VOICE ANALYSIS.....	261
Maria Sole Morelli, Silvia Orlandi, Claudia Manfredi	
INDEX OF AUTHORS	265



MAVEBA
2019
Firenze, Italy

FOREWORD

This book of Proceedings includes the contributions presented at the 11th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications – MAVEBA 2019, held in Firenze from 17 to 19 December, 2019. That is, 20 years since the very first MAVEBA in 1999!

Looking back to those days, I remember well the spirit of adventure that inspired this initiative, both on my side and on that of my colleague Piero Bruscaioni, with whom I also shared many subsequent MAVEBA editions.

MAVEBA started because of our curiosity and continued thanks to the enthusiasm of the participants. And today? Curiosity and enthusiasm are still there, with the awareness of a fascinating and increasingly interdisciplinary world. The large number of contributions collected in this Proceedings is the clear demonstration of this.

The main subjects concern methods for analyzing hoarseness and retrieving features of the human voice related to particular physiological or neurological conditions, with the aim of assessing reliable procedures for objective, quantitative definition of levels of voice disorders, singing voice parameters, newborn cry features, vocal fold and vocal tract modelling. The interdisciplinarity, that has always characterized the MAVEBA workshops, is well highlighted by the themes addressed, listed below.

I wish to give special thanks and greetings to the CoMeT Association, that is present at MAVEBA with a large number of its members. This year is a special one for CoMeT, celebrating the 50th anniversary from its foundation, and I am happy and proud to celebrate it together with the twenty-year anniversary of MAVEBA!

The papers presented at MAVEBA and collected in this volume are divided into nine Sessions, two Special Sessions, professionally coordinated by Dr. Franco Fussi and Dr. Philippe Dejonckere, and a Keynote lecture given by Thanasis Tsanas.

SESSION I – PARKINSON’S DISEASE AND SPEECH

SESSION II – SINGING VOICE

SESSION III - SPECIAL SESSION – EDUCATION AND REHABILITATION OF THE ARTISTIC VOICE

Coordinator: Franco Fussi

SESSION IV - VOICE AND EMOTIONS

SESSION V - VOICE QUALITY

SESSION VI - VOCAL FOLDS DYNAMICS

SESSION VII - COMET SESSION – ACTOR’S AND ACTRESS’ VOICES

Coordinator: Philippe Dejonckere

SESSION VIII – VOCAL FOLDS PARALYSIS/ABNORMALITIES

SESSION IX - KEYNOTE LECTURE

DEVELOPING NEW SPEECH SIGNAL PROCESSING ALGORITHMS FOR BIOMEDICAL AND LIFE SCIENCES APPLICATIONS: PRINCIPLES, FINDINGS, CHALLENGES, AND A VIEW TO THE FUTURE

Thanasis Tsanas

SESSION X - BIOMECHANICS/DEVICES

SESSION XI – SPEECH

SESSION XII - VOICE VS OTHER PHYSIOLOGICAL SIGNALS/DISEASES

I am very grateful to the authors for their contribution and to all participants that stimulated the discussion and helped to propose new research themes and methodologies of analysis in a field that will always be evolving, even and hopefully in the next twenty years.

Claudia Manfredi

ACKNOWLEDGEMENTS

I greatly acknowledge the ScaramuzziTeam Congress Agency for its great professionalism, Dr.Eng. Alice Cavaliere, who manages and constantly updates the website, and Dr. Eng. Lorenzo Frassinetti, PhD student, who collaborated in reviewing the Proceedings and solving the daily difficulties with patience and professionalism.

THE TOMATIS ELECTRONIC EAR EFFECTS ON SIMPLE VOCALIZATIONS

M. Prenassi¹, W. Coppola², G. Ramponi³, T. Agostini², S. Marceglia^{1,3}

¹ IRCCS Ca' Granda ospedale Maggiore policlinico Milano, Milan, Italy

² Dipartimento Scienze della Vita, Università degli studi di Trieste, Trieste, Italy

³ Dipartimento di Ingegneria e Architettura, Università degli studi di Trieste, Trieste, Italy

m.prenassi@gmail.com, walter.coppola@alice.it, ramponi@units.it, agostini@units.it, smarceglia@units.it

Abstract: The Tomatis electronic ear is a device that could modify the natural audio feedback between the emitted voice and the ears of a talking or singing individual. Our aim was to test if the device causes quantifiable vocal variations having the subjects repeat sustained vowel sounds (i.e. /a/, /i/, /u/) with different frequency filters applied by the device. The subjects are 19 native adult Italian speakers (8 females) testing 4 different filtering methodologies: unfiltered feedback (control), low pass filter at 4 kHz, high pass filter at 4 kHz and a high pass filter at 8 kHz. All subjects quantifiably modified their vocalization in response to the varying methodologies for at least one letter of each filtering method: 81.29% of the sessions of all subjects were significantly different in fundamental frequency from the control ($p < 0.05$, Kruskal-Wallis test). Among subjects, the variation trend was significant only for the fundamental frequency of the letter /u/ of a particular subgroups categorized by mean fundamental frequency. This initial work shows that the vocal variations caused by the Tomatis device are quantifiable but subject specific, laying the groundwork to test new parameters to find common trends of configurations.

Keywords: Tomatis, Electronic Ear, audio feedback, audio stimulation

I. INTRODUCTION

Alfred Tomatis was a French scientist, founder of Audio-Psycho-Phonology, an auditory rehabilitation methodology that stimulates the ear modifying the auditory input. This stimulation is delivered through a device called Electronic Ear. This device is based on a series of amplifiers, filters, and electronic controls, which receives the sound, emitted by a source, processes it and sends it back to the subject through a special headset. Tomatis's theory of listening is the product of a series of rigorous neurophysiological studies, based on the phylogenetic and ontogenetic analysis of the development of the nervous system [1][2][3]. It was fundamental to highlight the common origin and the consequent structuring of the organs

responsible for vocal emission (for example, V cranial pair for the musculature of the mandible and for the muscle of the hammer, VII cranial pair for the upper part of the larynx, for facial muscles and for the muscle of the stirrup), thus evidencing the very close correspondence between listening and voice production.

The conclusions reached by Tomatis are as follows:

“the voice can only contain the frequencies that the ear can hear (the larynx emits only the harmonics that the ear can hear)” and “if one modifies the hearing, the voice is unconsciously and immediately modified”[4][5][6]. In 1957, the theory was experimentally corroborated by a team led by Raoul Husson in the Functional Physiology laboratory at the Sorbonne in Paris [7]. After this experiment, fewer than a dozen offshoot and the related training systems have been developed based on this effect, with mild claims of effectiveness [8]. Only a fraction of these studies used the voice of the subject as auditory input. Our aim, in this preliminary work, is to test a new model of the Electronic Ear and the vocal variations that it causes on subjects emitting simple sounds (i.e. single sustained vowels) that are modified and fed back to them through special earpieces. This experiment was chosen to test the effectiveness of the device at a fundamental level, as a first step to map the actual capabilities of the device and of the method.

II. METHODS

In total, 19 native Italian speakers, 8 females and 11 males without speech impairments were recruited. The experimental setup included a microphone (Shure BETA 58A, Beyerdynamic TG V56c), the Tomatis system (Brain-Activator MBL), and an external recording device (M-AUDIO Fast Track Pro, sampling at 44100 samples/s, and a recording computer). The subjects were standing in a pre-marked position with their back and head touching a wall. The microphone was placed in a fixed position.

This preliminary experiment consisted of 4 segments, each composed of three sessions. In the first task of each segment, the subjects had to repeat 20 times the

vocalization of /a/ described as the corresponding vocal sound in Italian; this is called “a” session. After this, they waited 60 seconds and repeated the task with the letters /i/ and /u/ (“i” session and “u” session). The three sessions together define a segment.

In the first segment (NF), the Tomatis system acted as a straightforward audio loop, without any deliberate signal manipulation (no frequency filtering and no delay). NF is used as a control. The following three segments included a frequency filtering of the voice, first a low pass with -3dB cut-off frequency at 4 kHz (LP4K), then a high pass with the same cut-off frequency (-3dB at 4 kHz, named HP4K), and finally a high pass at 8 kHz (also -3dB cut-off, named HP8K). The total 4 segments of 3 sessions (NF, LP4K, HP4K and HP8K, for the sessions /a/, /i/, /u/) resulted in 80 vocalizations per subject and letter. The vocalizations were segmented and analyzed with the software PRAAT (<http://www.fon.hum.uva.nl/praat/>, downloaded June 2019) and a built-in algorithm [9] to extract the fundamental frequency (F0) and the first two formants (F1 and F2 respectively). If the algorithm was not capable of detecting a formant within a predetermined interval IF_i (1) the vocalization was discarded.

$$IF_i = MFV_i \cdot (1 \pm 0.5) \quad (1)$$

MFV_i: mean vowel formant (based on data from [10]);

i = 1: F1;

i = 2: F2.

Table 1. Percentage of sessions, ordered by letter, fundamental frequency and formants, statistically different from the control group (Kruskal-Wallis, $p < 0.05$). MeanV: mean change by letter for all the methods. MeanM: mean change by method for all the letters.

		LP4K %	HP4K %	HP8K %	MeanV %
F0	/a/	68.42	73.68	84.21	75.44
	/i/	78.95	84.21	84.21	82.46
	/u/	94.74	89.47	73.68	85.96
MeanM % F0		80.70	82.46	80.70	81.29
F1	/a/	42.11	42.11	57.89	47.37
	/i/	36.84	47.37	57.89	47.37
	/u/	52.63	68.42	57.89	59.65
MeanM % F1		43.86	52.63	57.89	51.46
F2	/a/	42.11	42.11	47.37	43.86
	/i/	63.16	47.37	68.42	59.65
	/u/	57.89	63.16	52.63	57.89
MeanM % F2		54.39	50.88	56.14	53.80

III. RESULTS

All subjects completed the experiment, resulting in 228 sessions (76 segments), but we had to discard some poor quality samples; overall, we analyzed 18 vocalizations for each session, for a grand total of 4104 vocalizations. Among the segments, regarding F0, 81.29 % of the sessions of all subjects were significantly different from the control NF ($p < 0.05$, Kruskal-Wallis test) as shown in Table 1. The first and second formants were less influenced by the Tomatis loop than F0, with 51.46% of sessions significantly different from NF for F1 and 53.80% of sessions significantly different from F2 (Table I). Still, all subjects responded to some extent to the feedback, modifying their vocalization formants in response to a method for at least one letter.

In Fig. 1 are shown the vocalizations of four different subjects of the letter /i/, even if some sessions are visually different from the control group (Fig. 3, C and D).

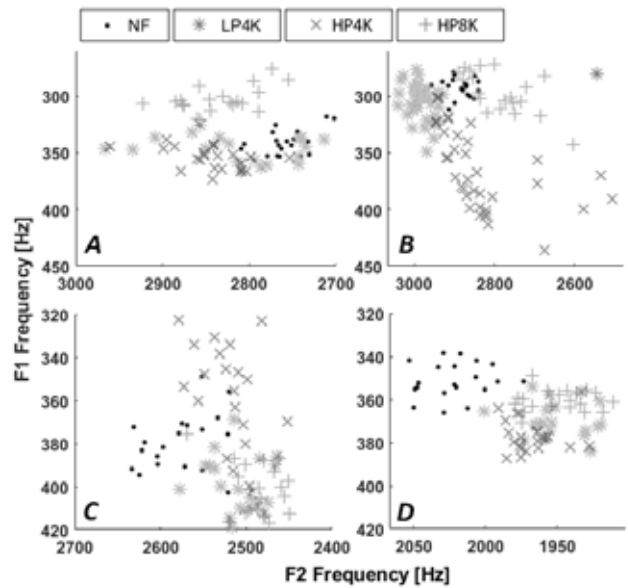


Fig. 1. First and second formants for the letter /i/ of four subjects, divided by methodology. A: subject 1, high F0 group, B: subject 7, high F0 group; C: subject 13, low F0 group, D: subject 19, low F0 group.

To assess the presence of underlying trends, further analysis were conducted.

Means and standard deviations were calculated for each session. Subjects were divided into two groups: high-F0 and low-F0 (the threshold was the total average of F0). The groups were composed of 8 female and 2 male subjects for the high-F0 category (in total 10 subjects) and the remaining 9 male subjects for the

low-F0 category. This classification was performed because group-specific filtering effects were observed in the preliminary analysis.

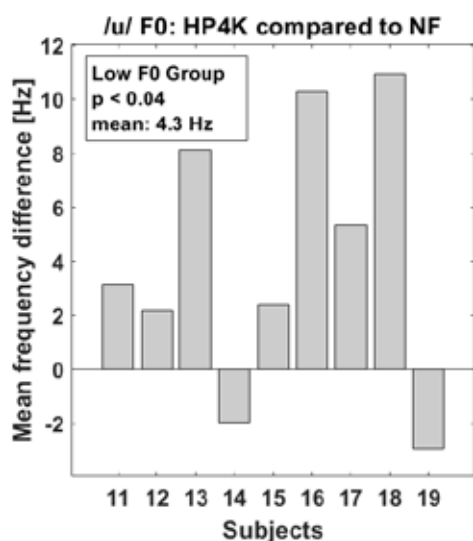


Fig. 2: Difference for the F0 of letter /u/ between the HP4K method and control (NF), regarding the low F0 group.

In the low-F0 group there was a significant increase by 4.3 Hz ($p < 0.04$, Wilcoxon signed rank test, Fig. 2) in the fundamental frequency (F0) of the letter /u/ of the HP4K session. For the same vowel, between HP4K and LP4K, F2 increased by 68.23 Hz ($p < 0.006$, Wilcoxon signed rank test) in the high-F0 group, and increased for the letter /a/ by 33.63 Hz ($p < 0.02$, Wilcoxon signed rank test) for the low-F0 group, as shown in Fig. 3. The standard deviation analysis had a statistical significance only in the F2 formant for vowels /a/ and /i/ in the high-F0 group: it decreased from HP8K to HP4K and increased from HP4K or HP8K to LP4K.

In the low-F0 group, the standard deviation had statistical significance only for the letter /i/: it increased in F1 for HP8K compared to controls, and for LP4K compared to controls. It decreased in F2 for HP8K compared to HP4K. Fig. 4 shows the average values of F0 for all the subjects in the high-F0 and low-F0 groups; all the letters showed a distinct increase between control NF and the methods, except for /a/ in the lower F0 group. However, these results are statistically different only for the letter /u/ in the lower F0 group between NF and HP4K ($p < 0.05$, Wilcoxon signed rank test).

IV. DISCUSSION

As shown in the results section a clearly audible and statistically relevant response is evoked by the filtered Tomatis audio loop, even if a unified trend response

among subjects is not clearly delineated. The first results we present in this paper indicate some ability of the Tomatis system to modify these covariances and as a consequence to act on general voice quality. Indeed, experiments in which the relationship among F0 and the F1 and F2 formants are synthetically modified are reported in the open literature (e.g., [11]). It was shown that the perceived quality of a voice depends on the covariance of the formants, which should correspond

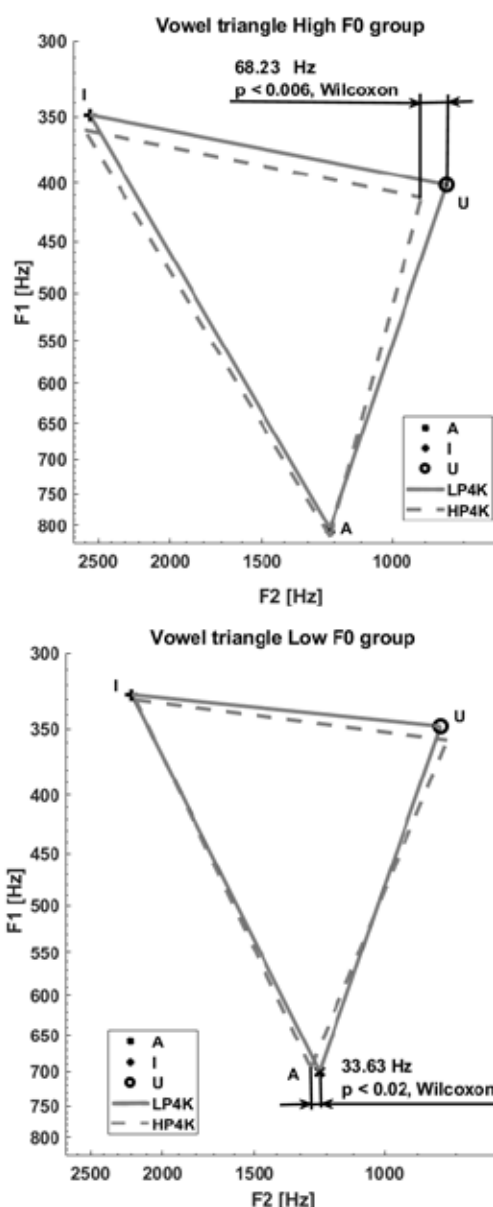


Fig. 3: First and second formant for the all-subject average vowel triangle of the letter /a/, /i/, /u/ for the LP4K and HP4K methods. At the top the high F0 group and at the bottom the low F0 group methods with the significant F2 drifts for the letter /u/ (top) and /a/ (bottom).

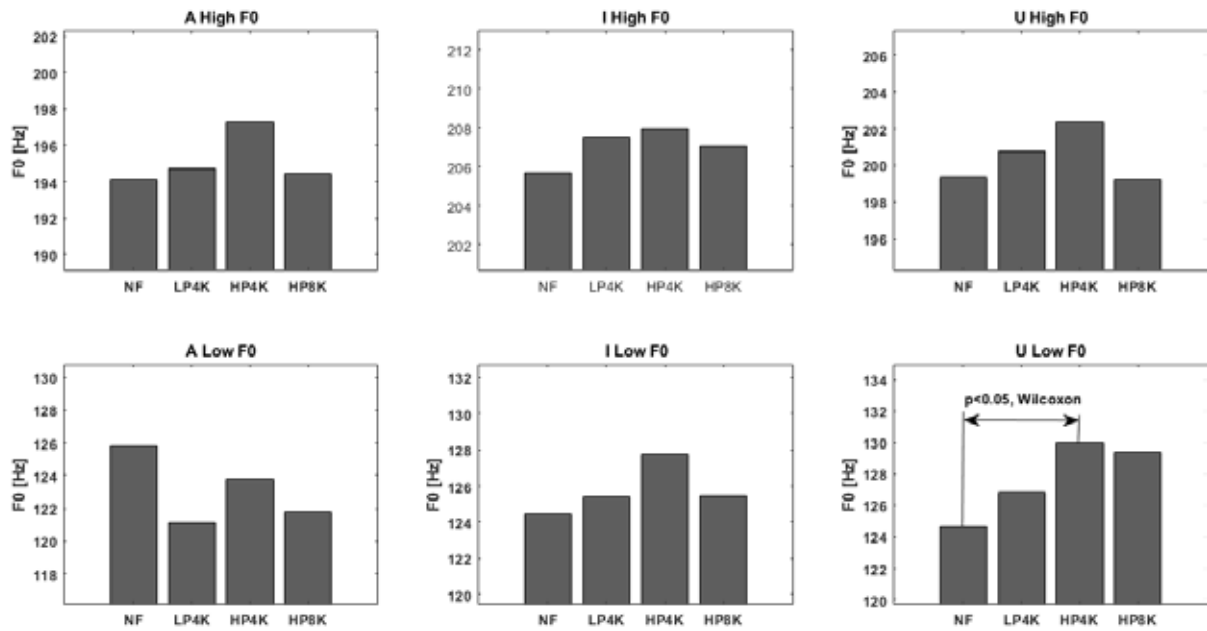


Fig. 4: Average values of F0 for all the subjects of the high-F0 and low-F0 groups. In U, Low F0, the differences between NF and HP4K are statistically significant ($p < 0.005$, Wilcoxon signed rank test)

to an internalized representation of human voice. Even the perception of emotions of the speaker depends on the formants' properties [12]. Even if not statistically relevant, it is also worth of notice the precise pattern followed by the all-subject average of F0 shown in Fig. 4. The HP4K method is able to elicit bigger changes in the fundamental frequency compared to the controls.

V. CONCLUSION

In this preliminary experiment, we showed that the vocal changes elicited in the subjects by the Electronic Ear are quantifiable. The specific effects depend on the type of vocalization and on the class of the subject (high-F0 or low-F0). A clear trend in the fundamental frequency was detected only for the /i/ and /u/ vowels. The standard deviation analysis suggests that a central (4 kHz cut-off frequency) high pass voice filter tends to increase the sound variation. This initial work may be useful to understand the capabilities of the Electronic Ear and the Tomatis methodology.

REFERENCES

- [1] P. Sollier, *Listening for Wellness. An Introduction to the Tomatis Method*. Walnut Creek, CA: Mozart Center Press, 2005.
- [2] W. Coppola, *Itinerari sonori. Corpo, linguaggio ed espressione nell'opera audio-psico-fonologica di Alfred Tomatis*. Milan: Mimesis, 2015
- [3] C. Campo, *L'orecchio e i suoni fonti di energia. Il metodo Tomatis*, Riza Scienze, 74, 1993.
- [4] A. Tomatis, *L'oreille et le langage*, Paris: Édition du Seuil, 1963.
- [5] A. Tomatis, *Vers l'écoute humaine, Tome I, Qu'est-ce que l'écoute humaine? Tome II, Qu'est-ce que l'oreille humaine?*, Paris: ESF, 1974.
- [6] A. Tomatis, *L'oreille et la voix*, Paris: Laffont, 1987.
- [7] R. Husson, "Physiologie de la Phonation", *Practica Oto-Rhino-Laryngologica*, 1957, vol. 50.2, pp. 123-139.
- [8] B.M. Thompson and S.R. Andrews, "An historical commentary on the physiological effects of music: Tomatis, Mozart and neuropsychology", *Integrative Physiological and Behavioral Science*, 2000, vol. 35, pp. 174.
- [9] D. G. Childers, *Modern spectrum analysis*. 1st ed., IEEE Computer Society Press, 1978, pp. 252-255.
- [10] P. Ladefoged, and K. Johnson, *A course in phonetics*, 6th ed., USA: Wadsworth, 2011.
- [11] P.F. Assmann, and T.M. Nearey. "Relationship between fundamental and formant frequencies in voice preference," *The Journal of the Acoustical Society of America*, 2007, 122.2 EL35-EL43.
- [12] W. Xixin, L. Songxiang, C. Yuewen, L.Xu, Y. Jianwei, D. Dongyang, M. Xi, H. Shoukang, W. Zhiyong, L. Xunying, M. Helen, *ICASSP 2019 - IEEE International Conference on Acoustics, speech and Signal Processing (ICASSP)*, 2019.