# **BioMet®Phon**

## Tool for the Evaluation of Voice Quality and Biometry

## **User's Manual**



Version 2.3 – December 2012 © BioMetro Soft, S.L.

## **BioMet®Phon**

## Tool for the Evaluation of Voice Quality and Biometry

© BioMetro Soft, S.L. C/ Ribera del Loira 42, Ed. 2, 28046, Madrid, Spain Tel.: +34.91.503.0145 Fax: +34.91.403.0099 URL: www.glottex. com e-mail: info@glottex.com

© Copyright and reproduction of the materials in the present User's Manual by any means in any media are restricted by the Copyright holder. This document can not be reproduced, copied or transferred to other parties without the consent in writing by the Copyright owners.

## Index

1	Intr	oduction	4						
2	Wł	nat is Glottex®Technology	4						
	2.1	The engine behind Glottex®Technology							
	2.2	What is the glottal source	5						
	2.3	What is the glottal source useful for							
3	Bio	Met®Phon	8						
	3.1	Management of Patient's Database							
	3.2	Storing New Records in the Active Patient Folder	8						
	3.3	Previous or New Record Analysis from the Active Patient	10						
	3.4	Insertion of a New Patient	14						
4	Ins	Installation							
5	MS	MSDOS Auxiliary Window14							
6	Inte	Intellectual Property15							

## 1 Introduction

BioMetro Soft S.L. is a company created as a 'start-up' from Universidad Politécnica de Madrid, Spain, founded in June 2011. It initiated its activites after getting the first prize in the VII Contest for the Creation of Companies with Technological Basis in December 2010 (actúaUPM). The field of action of BioMetro Soft S.L. is the implantation and commercialization of software applications using the human voice for medical, biometrical and linguistic applications.

### 2 What is Glottex®Technology

Glottex®Technology is a signal processing engine designed to extract the glottal source from voiced speech. It is entirely based on the reconstruction of the inverse vocal tract transfer function to remove its influence from voiced speech. The result is a correlate to the dynamic pressure build-up in the supraglottal side of the vocal folds, which bears behavioral marks on the specific phonation of a particular speaker.

#### 2.1 The engine behind Glottex®Technology

The underlying technology or Glottex®Engine is a specific set of computer programs designed to estimate the vocal tract transfer function removing it from voice. The results are two signals known as the glottal source and the Glottal Flow. The glottal source may be seen as the dynamic pressure build-up in the pharyngeal side of the vocal folds (supraglottal). The Glottal Flow may be seen as the flow of air through the vocal folds at a given time instant during the phonation cycle. The foundations of the technology are to be found in the Fant Model for Voice production given in the figure below:



The speech production system in Fant's source-filter model is composed of an Excitation Model (1 for voiced speech, 2 for unvoiced speech) generating the excitation signal e(n) with power spectral density given in (5), a filter emulating the Vocal Tract (3) imposing resonances and anti-resonances (as in 6) to the excitation to generate the speech before radiation  $s_a(n)$ , and the lip-radiation model (4) explaining the evolution suffered from the sound wave from the

radiation outlet to the microphone to produce the recorded signal  $s_r(n)$ , with power spectral density given by (7).

The Glottex®Engine produces inverse models for (4) and (3) to reconstruct  $s_a(n)$  and e(n) from the recorded signal  $s_r(n)$ . In the case of voiced speech the resulting estimated excitation will be  $\hat{e}(n)$  as an approximation to the original glottal source. Based on this signal different correlates may be produced, as for example the Glottal Flow, the distortion parameters *jitter* and *shimmer*, the harmonics-noise correlate, the biomechanical parameters of the vocal folds (dynamic mass, viscoelasticity) the temporal phonation cycle description, the gap alterations and the tremor in voice, among others.

#### 2.2 What is the glottal source

The glottal source is a correlate of the pressure build-up of the dynamic pressure in the supraglottal side of the vocal folds as depicted in the figure below:



In the uppermost template a segment of phonated speech (voice) is presented (a). The ringing observable in the trace is due to the vocal tract resonances. In (b) the residual obtained when removing the vocal tract resonances is appreciated, where the resulting signal is a train of excitation spikes. The glottal source obtained from the integration of the glottal residual is depicted in (c). The Glottal source is characterized by sudden drops of pressure (marked by asterisks) corresponding to the closure of the vocal folds (closing or contact phase) followed by a recovery to dynamic pressure 0 forming a plateau to expand in a positive raise of pressure during the open phase to drop again during the closing phase. The glottal flow (d) appears as a triangle-like burst of flow (air) during the open phase followed by an almost flat drop during the 5/15

closed or contact phase (the Glottal Flow signal shows fluctuations due to uncompensated integration effects). Seen in detail the glottal source describes a pattern which is known as the Liljenctrants-Fant cycle, represented below.



The lower template shows a series of phonation cycles, one of them being depicted in detail in the upper template (blue). The pattern shown is known as the Liliencrants-Fant cycle. The description in the picture starts during the closing instant (0), where the pressure has experienced a drastic drop below 0. Due to the elastic nature of the gas column in the vocal tract, a recovery is experienced reaching almost a stable value near 0 at tR1. This is known as the recovery phase. From there on till tO2 (opening instant) the dynamic pressure stays close to 0 (resting sub-phase), as the vocal tract is closed by the vocal folds (except for slight escapes of flow represented by the green line). During the open phase from tO2 till the end of the cycle a burst of flow (in green) is expelled. The pressure raises during the first part of the burst injection till tM (abduction sub-phase). At this point the vocal folds initiate an approximation (adduction sub-phase) to close the vocal tract again, which shows a steady and sharp decay in pressure till the end of the cycle (tC). This is the L-F cycle: closing, closed phase (0-tM) composed of recovery phase (0-tR1) and steady closure phase (tR1-tO2), and open phase, composed of abduction sub-phase (tO2-tM) and adduction sub-phase (tM-tC) to be followed by a new closing.

#### 2.3 What is the glottal source useful for

Laryngology: The glottal source expresses two different phonation conditions which have strong correlates with dysphonia or pathology, these are cyclic asymmetry and defective closure. The first condition shows remarkable

dissimilar neighbor phonation cycles, i.e., each vibration pattern is more similar each two cycles than to their neighbor patterns. Asymmetry may be due to anatomical differences. to unilateral partial paralysis, to unilateral neuromuscular path lesions, or to unilateral organic lesions (polyps, cysts, etc.). The second dysphonic condition has to see with deficient or incomplete closure, resulting in air escape during the closed phase. This would result in an inefficient air use, fatigue, excessive turbulence, and irregular L-F pattern with loss of closure sharpness in the negative salience of the glottal source peaks, irregular or too slow recovery phase, unstable resting phase, or distorted abduction and adduction sub-phases. Defective closure has to see with unilateral lesions of the kind commented or with bi-lateral lesions making closure difficult, as nodules. Another source of incomplete closure has to see with a lack of tension in the crico-arytenoidal muscle-cartilage structure, leaving a permanent gap or escape affecting the arytenoidal side of the vocal folds. This results in a permanent gap condition producing airy voice and a loss of efficient air use. All these conditions may be checked by the graphical presentation of the glottal source produced by Glottex®Engine as well as from the numerical distortion parameters as *jitter*, *shimmer*, *harmonic-noise ratio*, *biomechanical* unbalance, cycle time and gap alterations, which are produced by the application. Their use in Phoniatry have to see with the comparison of the parameters against the normal distributions and limits of these obtained from normophonic speakers of both genders. The tool **BioMet®Phon** renders these estimations and contrasts and plots them in semaphores which alert on deviations of the parameters which may show a pathological semantics. These evaluations may be used clinically in comparisons of pre-post surgery or treatment conditions, monitoring success or progress.

<u>Speech Therapy & Rehabilitation</u> The same distortions may also be used in evaluating the rehabilitation progress or success in speech therapy. Besides estimates of the biomechanical parameters related with dynamic vocal fold mass and viscoelasticity may be used by the Speech Therapist to track gesture, modality of phonation, pitch elevation, air use, etc., to educate or restore the phonation function in an objective manner. These parameters may be stored in EXCEL® files for the documentation of validation and contrast of the phonation function.

<u>Voice Forensics</u> Many of the phonation abnormal conditions are true biometrical marks of the speaker appearing in different realizations in a speech segment, which may be used as forensic evidence marks. The parameters estimated by the Glottex®Engine may be used for automatic evaluation tests among Questioned Evidence against the Suspect's Evidence in a line-up Reference Model used as contrast. These contrasts are implemented by Gaussian Mixture Models embedded in the tool **BioMet®Fore** allowing Forensic Voice Evaluation Contrasts in a professional manner. The application generates different graphical material helping in the documentation of real cases before the Court.

<u>Neural Pathology Detection & Monitoring</u> This is a very promising line currently under research and development which is already implemented and normalized against a population of normophonic speakers of both genders. Currently three cyclicality expressing parameters are derived with tremor detection capabilities useful in monitoring Parkinson's Disease progress, for instance. These parameters are indirectly estimated from the viscoelastic estimates, which are known to be directly related to neuromuscular activity of the branches of the laryngeal nerve acting on the transversal and oblique arytenoid and cricothyroid muscles responsible for the vocal fold adduction and abduction (cricothyroid, transverse and oblique arytenoid and posterior cricoarytenoid). Therefore from assymetric and viscoelastic studies the neurological condition of the patient affecting vocal fold tension may be inferred and monitored.

<u>Singing Voice</u> The use of viscoelastic estimates in monitoring vocal fold performance allows the Speech Therapist or the Singing Trainer or Educator to track the education, rehabilitation and care of the professional singer. These functions may be directly obtained from the tool **BioMet®Sing**.

<u>Linguistics</u> The detection of voiced and unvoiced segments in running speech separating the glottal function (prosody and modality controlled) from the vocal tract influence (articulation controlled) is very efficient in producing acoustic-phonetic labeling, as well as prosodic and rhythmic information which can be used for advanced linguistic studies. This information is provided by a sophysticated tool as **PangurBan®ACAD**.

#### 3 BioMet®Phon

It is a powerful application specifically conceived for the Analysis of Voice Quality by Voice Care Practitioners and Speech Therapists. It is based on a General User Interface (version 2.3) running Glottex®Engine (version 7.3) as the one shown in Figure 1. The application incorporates the following features:

#### 3.1 Management of Patient's Database

Patients Records are organized by Patient's Folders using a Patient's Name Convention. Other folders and records may be also processed assuming they are stored as '\*.wav' audio files. The main functions available under the Patient's Database Management Interface are:

**Browse Patient's Control** button, which allows opening a specific Patient's Folder.

**Change Patient** button, which allows declaring a New Patient, a New Folder, or simply changing the current Active Patient. As this procedure opens in a new User Interface it is described later.

#### 3.2 Storing New Records in the Active Patient Folder

The Active Patient is the one referenced by the Active Patient Folder Path in the upper leftmost text window. Once this path is open the user may proceed acquiring a new record to be included in the Active Patient's Folder by means of the **Record** button. Prior to activating this button the user is expected to select one of the four Record Type buttons immediately below:

- Vowel /a/, for maintained utterances of vowel /a/ lasting several seconds
- Split vowels, for utterances of trains of vowels, like /ah/, /eh/, /ih/, /oh/, /uh/ in stacatto
- Sentence, for utterances of fixed short sentences as /I owe you a letter/ or any similar
- Reading, for utterances of read text lasting some minutes, as the "Rainbow Passage"



Figure 1 Main Graphic User Interface of **BioMetr®Phon** showing the buttons and active windows in analyzing a voice frame.

The activation of the **Record** button will open a Recording Session by:

Activating the recording window on the Signal Window immediately above the **Record** button with a time length given by:

- once the duration established in Record Length for Vowel /a/ sessions
- twice the duration established in Record Length for **Split Vowels** sessions
- four times the duration established in Record Length for **Sentence** sessions
- ten times the duration established in Record Length for **Reading** sessions

During the recording the resulting trace will be shown in the Signal Window as it is being produced second by second, and the VU-meter will show a scale with the maximum level recorded in each second interval in colors.

An automatic name will be selected for the record with the following convention:

<L-yyyymmdd-hhnnss-session>.wav

where:

L is a label with the same name than the Active Patient's Folder yyyy is a four character label with the current year mm is a two character label with the current month dd is a two character label with the current day hh is a two character label with the current hour nn is a two character label with the current minute ss is a two character label with the current second session is one of the following according with the Record Type selected:

> vowela for **Vowel /a/** type splitvowels for **Split Vowels** sentence for **Sentence** reading for **Reading**

After the recording, a segment of the duration given by the window **Analysis Frame (ms)** in the **BioMetr®PhonCONFIG** Interface (activated by **Change Patient**) on the central part of the recording will become automatically the Active Segment for analysis and lit in green. The active segment may be reproduced with the **Play Seg** button. The limits of the Active Segment will be published in the editable windows immediately below the **Analyze** button. It is possible to change the Active Segment by clicking on the Signal Window area. A new 200 ms frame will be highlighted.

The recording may be stopped at any time with the **Stop** button.

The recorded trace may be listened with the **Play All** button.

The reproduction may be stopped at any time with the **Stop** button.

#### 3.3 **Previous or New Record Analysis from the Active Patient**

The Active Patient is the one referenced by the Active Patient's Folder Path in the upper leftmost text window. Once this path is open the user may proceed selecting a former record from the Active Patient by means of the **Browse Patient's Folder** button and double clicking in the desired file name, or to analyze a new recording just acquired by clicking on the **Analysis** button. The Active Segment will be used in the analysis. The analysis will take some few seconds after which period the following results will be observed:

A **Spectral Contents Window** as the one in the next figure will be published:



Figure 2 Spectral Contents Window. Upper template: in blue the Fourier Spectral Density of the voice signal; in red the LPC Transfer Function of the Vocal Tract. Lower template: in blue the Fourier Spectral Density of the Glottal source; in red the average LPC profile of the same signal background.

This window shows the Fourier Analysis of the Active Segment in blue and the LPC spectral envelope of the vocal tract in red marking the positions in Hz and the relative heights of the voice formants in dB. The average behavior of the glottal source will be marked in green. It is very important to notice that the estimations of the vocal tract spectral behavior is free of the glottal source influence, which makes it a very precise estimation compensated in spectral tilt and free of the Glottal Formant.

A frame with the glottal source cycles contained in the Active Segment is published in the **Glottal Source Clipping** window in the lower right part of the User Interface. One of this cycles (closest to the middle of the Active Segment) is published in detail in the **Glottal Source and Mucosal Wave Correlates** window in the upper right part of the User Interface. Estimations of the Recovery Time (tR1 and tR2), the Opening Time (tO1 and tO2) and the glottal source Maximum Instant (tM) are given in the template, as well as estimations of the following Glottal Gap Alterations:

ContGap for the Contact Gap Alteration, i.e., the maximum residual opening defect found during the Closed Phase.

AdduGap for the Adduction Gap Alteration, i.e., the maximum residual counter-closing or retract of the vocal folds during the Closing Phase.

PermGap for the Permanent Gap Alteration, i.e., the maximum residual opening found during the complete phonation cycle due to defective closure in the arytenoidal extreme of the vocal folds.

An estimation of the Reliability of the algorithmic extraction process and the Average Pitch are published in the editable windows immediately below the Recording Type panel.



Figure 3 Summary of the analysis: In the left hand side from top to bottom the description of the tool, the patient's data and the statistics of the analysis for the parameters selected are to be found. In the right hand side from top to bottom the graphical results for a glottal cycle and the set of glottal cycles analyzed are to be found.

Values for five Voice Quality Parameters (*jitter, shimmer, body mass, body stiffness* and *body mass unbalance*) are published in the table immediately below. These parameters are found to be largely correlated with certain types of dysphonias. The deviation values of a distribution of normal speakers (normophonic) are given in black, the actual estimation for the specific speaker being analyzed given in blue.

Time (s)

A set of semaphores is lit immediately to the right of the table, with the following semantics:

- green: the estimate is within distribution limits
- · red: the estimate is above the distribution upper limit

• blue: the estimate is under the distribution lower limit

A **Summary** of the above information and figures including the patient's available data is published in an independent figure as in Figure 3.

This figure shows the specific Active Patient's information in the upper left most part, the values of the Voice Quality Parameters and their distributions in the lower middle, and a copy of the figures published in the User Interface in its right most part.

The published figures in '\*.pdf' format and an **Excel**® file with the complete Voice Quality Parameters (61 different including dispersion, cepstral, biomechanical and cyclical) are stored in a folder named **results** within the Active Patient's Folder if the **Save Results** button under **Analyze** is activated.

glottex_neu_cap_config	The second s													100 C		
	BioMe								ol	Met®PhonCONF	et®PhonCONFIG					
	®BioMetr							©Biol	Met	tre Soft S.L., Yerrion 2.3 (7.3) July 2012						
Patient's Name:												Patients' Books:				
Patient's MidN:										PatientLog-old.xls PatientLog.xls					^	
Patient's SurN:																
Patient's Age:							_									
Patient's Code:			referenc	e <sup>⊗</sup> N	lale	© Female										
Sampling Freq. (Hz)	):			44	100										-	
Record Length (s):					1.0							Patients' List:				
Analysis Frame (ms	i):			1	100										*	
Visible Parameters:		1	2	3	4 5		8	7 8								
Normo Male Model:		NMM-50-7-1-22k-20120604-101028.mat						-101028.m	nat							
Normo Female Mod	lel:	N	NMF-50-7-1-22k-20120604-102019.mat			-102019.m	nat					-	-			
Origin e-mail:												Diagnose/Comments:				
SMTP Server:															^	
Destin. e-mail:																
EITHER SELECT Bool	k from "Patients' I	Books" O	R Press "E	Browse Pa	tient's Boc	oks"										
I:\Taller\Voz_Patologic	ca\trabajos\ejecut	tables\exe	_glottex\cc	ompiladas\	BioMetroP	honPl	ROF_23	_32\distrib	)gral	abaciones\						
Browse Pi	atients' Books					New P	atient							Exit		

Figure 4 Graphical User Interface for the configuration of the tool and the selection of a given patient's folder and profile.

Use the Change Patient button to change the Active Patient or declare a New Patient. The activation of this button opens a new Graphic User Interface as the one in Figure 4.

The list of active Patients' Books (Excel® files containing different patient lists which the user may have created before) is given in the sliding window **Patients' Books**. To select a patient the user must proceed as follows:

- The user must click on one of the active Patients' Books to open it (for instance on PatientLog.xls, which is the book by defect). This will open the list of available patients' labels in the Patients' List sliding window.
- 2. Select from the Patients' List by clicking on the corresponding label. The path pointing to the new patient will be highlighted and transferred to the path editable window (in the lower part of the figure).
- 3. The user may now leave the patient selection interface by clicking on **Exit**. This will complete the Patient Selection Process.

IMPORTANT OBSERVATION: An associate **PatientLog-old.xls** file is also resident in the same Patient's Database Folder containing a back-up copy. The user is supposed to select only **PatientLog.xls** and not the back-up copy, which is saved for security purposes only.

#### 3.4 Insertion of a New Patient

If the user wants to define a New Patient the operation is as follows:

- 1. Open a Patient's List as commented before.
- 2. Activate the **New Patient** button.
- 3. The Patient's Information demanded in the upper left most editable windows is to be filled. The user may choose to fill the Name, Middle and Surname of the Patient, or fill an arbitrary label on the Patient's Code window. If the Name, Middle and Surname info is supplied the Patient's Code will be filled automatically by the application with the initials of each field. Otherwise the Patient's Code provided by the user will prevail. Optionally the user may provide the age (Patient's Age window)
- 4. Selection of gender (Male/Female button) is compulsory.
- 5. If desired the Sampling Frequency may be changed as well as the Record Length. This last parameter will be used in configuring the duration of the Record Type sessions as commented before. Selection of the Analysis Frame length in ms is also optional.
- 6. An optional text to be included in the Patient's Fields at the time of New Patient Creation may be introduced in the Diagnose/Comments window at this time.
- 7. The operation ends clicking on Exit. This will create a new subfolder for the New Patient and declare it the Active Patient for the session and complete the process of New Patient Declaration. In case the Patient's Code is pointing to an already existing subfolder two messages will indicate this situation asking for user's confirmation by clicking on the Right Mouse button.

#### 4 Installation

Copy the file 'BioMetroPhon\_pkg.exe' in the folder of your preference for installation.

Double click on it and follow the sequence of windows which will be displayed by the Matlab RunTime Library expansion application.

### 5 MSDOS Auxiliary Window

In Windows XP and related versions an auxiliary MSDOS window will appear (similar to the one in Figure 5) to collect and show the messages generated by the application, especially if execution errors are found. In such case, collect a copy of this window with CTRL+ALTGR+PRTYSCR, paste this figure with CTRL+V in a word file and send it to info@glottex.com with a brief explanation of the circumstances accompanying the error.



Figure 5 Console of the application showing an error not captured internally.

## 6 Intellectual Property

Glottex®, Glottex®Engine and Glottex®Technology are registered names. Glottex®Technology is the subject of protection under Patent No. P201131069 (Spain) and PTC/ES2012/000137 (Europe).

The Intellectual Property Rights holder is Universidad Politécnica de Madrid, Madrid, Spain (www.upm.es).