Calibration of external lighting and sensing photoglottograph
Anne Bouvet, Annemie van Hirtum, Xavier Pelorson, Shinji Maeda, Kiyoshi Honda, Angelique Amelot

To cite this version:
Anne Bouvet, Annemie van Hirtum, Xavier Pelorson, Shinji Maeda, Kiyoshi Honda, et al.. Calibration of external lighting and sensing photoglottograph. 10th international workshop Models and Analysis of the Vocal Emissions for Biomedical Applications (MAVEBA), Dec 2017, Florence, Italy. hal-01671688
CALIBRATION OF EXTERNAL LIGHTING AND SENSING PHOTOGLOTTOGRAPH

A. Bouvet\(^1\), A. Van Hirtum\(^1\), X. Pelorson\(^1\), S. Maeda\(^2\), K. Honda\(^2\), A. Amelot\(^2\)

\(^1\) CNRS UMR216, Gipsa-lab, Université Grenoble-Alpes, Grenoble, France
\(^2\) Laboratoire de Phonétique et Phonologie (LPP), Université Paris 3, Paris, France
anne.bouvet@grenoble-inp.fr

Abstract: Observation and measurement of vocal folds vibration and glottal opening during speech requires techniques as little invasive as possible for the subject. The LPP has developed the External Photoglottograph (ePGG) system. It consists of illuminating the glottis through the neck skin with an infrared light and recording light variation intensity modulated by glottal movement with a photodiode placed across the larynx. The system is tested on two mechanical larynx replicas. The first one consists of two rigid half cylinders in forced oscillation controlled by a step motor. The second one is flow driven and uses latex tubes filled with water in order to reproduce vocal folds self-oscillation. Time-varying glottal area is measured accurately for both replicas. Experimental results are compared to ePGG recordings in order to assess the correlation between area measurements and ePGG signal. This characterization is used to propose a calibration of the glottal opening as a function of parameters affecting the ePGG signal (distance, angle, skin, tissue, system setting, etc.).

Keywords: ePGG, glottal area, vocal folds auto-oscillation, non-invasive measurement, mechanical replica

I. INTRODUCTION

Observation and measurement of the glottal opening and vocal folds vibration during speech requires the measurement of small displacements (of order of a millimeter) at a high sampling rate. Further, the technique must be as little invasive as possible in order not to prevent from normal speech and/or articulation.

External lighting and sensing ElectroPhotoGlottoGraph (ePGG), developed at LPP [4], relies on transillumination technique (Fig. 1), which consists of illuminating the glottis through the neck skin with infrared light and recording the variation of light intensity modulated by the vibration of the vocal folds with a sample rate of 20 kHz. In contrast to common laryngeal illumination techniques no visible light is used to light the glottal area. Instead light in the near infrared (IR) spectral range is used since wavelengths in this range (700-1000 nm) are reported to transilluminate large sections (many centimeters) of human tissue [1-3]. Compared with other photoglottographic techniques, this device has the advantage that both the light source (IR) and the light sensor (S) are positioned on the speaker’s neck as illustrated in Fig. 1. Non-invasive, it allows to perform measurements without coercion for the speaker and can be used without medical assistance.

The objective of this paper is therefore to assess ePGG measurements during auto-oscillation of a mechanical glottal replica for which the varying glottal area can be accurately quantified during oscillation. The influence of all parameters potentially affecting the outcome of an ePGG measurement is studied in a controlled and repeatable way in order to propose a calibration procedure or/and provide objective guidelines for ePGG usage.

Fig. 1: Glottal transillumination principle.

II. METHODS

In order to realize this objective, experiments are performed using the following experimental setups.

Firstly, to characterize the ePGG outcome, the system (source and sensor) is mounted on an optic bench, so that the emitter and receiver system can be characterized (positioning, settings) and environmental measurement disturbances can be identified. Next, a uniform Plexiglas tube (diameter 25 mm) is used to represent the human trachea and pharynx. The tube is covered with layers of lamb leather to simulate the absorption of light by the human skin. The impact of leather thickness and IR source positioning (angle, distance to sensor) are assessed.
Firstly, the effect of source-sensor distance $d$ (Fig. 1) on the ePGG signal is sought. The ePGG system is positioned on the mechanical airway replica with constant area ($A_u = 491 \text{ mm}^2$). The source-sensor distance $d$, is systematically varied in the range $2 \text{ mm} \leq d \leq 200 \text{ mm}$ and the orientation angle is $27^\circ$.

In addition, in order to mimic the influence of wall tissue thickness, measurements are performed adding two (thickness 1.4 mm) or three latex layers (thickness 2.1 mm). Measured mean ePGG signals are plotted in Fig. 3. The ePGG signal decreases with $d$ regardless of wall thickness. Linear fitting of measured ePGG signals in the range $d \leq 100 \text{ mm}$ (appropriate for human subjects) and in the range $d \geq 100 \text{ mm}$ (appropriate for mechanical replicas), results in $R^2 \geq 98.9\%$. Consequently, a first order linear approximation can be used to characterize the evolution of ePGG signal with source-sensor distance $d$, while the negative slope depends on wall absorption (thickness) and distance $d$. All remaining experiments are done with 2 layers (thickness 1.4 mm).

Secondly, static geometrical configurations are done to determine the effect of the source orientation angle in the mid-coronal plane (Fig. 1) on the ePGG signal. The ePGG system is again positioned on the uniform mechanical airway replica, i.e. in absence of a glottal constriction (no glottal replica). The source orientation angle is systematically varied from $0^\circ$ up to $40^\circ$ and the source-sensor distance is held constant to $d = 100 \text{ mm}$. Measured mean ePGG signals are plotted in Fig. 4. For orientation angles up to about $15^\circ$, the ePGG signal is minimum and only marginally ($< 0.3 \text{ V}$) affected by the orientation angle due to the source (IR) half beam angle of $22.5 \pm 2.5^\circ$. Further increasing the orientation angle above $15^\circ$ results in a linear ($R^2 = 98.1\%$) increase of the mean ePGG signal. All remaining experiments are done for orientation angle $27^\circ$.

Thirdly, static geometrical configurations are performed to determine the effect of glottal area on the ePGG signal (Fig. 1). The rigid and deformable vocal

**Fig. 2:** ePGG system source/emitter IR and optic sensor/receiver (S) placed at a) motor-driven rigid replica, b) flow-driven auto-oscillating replica.

**Fig. 3:** Mean ePGG signal as a function of source-sensor distance $d$ for the airway replica with 2 and 3 leather layers.

**Fig. 4:** Mean ePGG signal as a function of source-sensor distance $d$ for the airway replica with 2 and 3 leather layers.

**III. EPGG SIGNAL MODEL**

The ePGG system is assessed on mechanical replicas. Since experimental setups are equipped to measure the glottal area, the relationship between ePGG signal and glottal area can be studied on these replicas as a function of parameters potentially affecting the ePGG signal (Fig. 1). In the following, the experimental ePGG signal characterization is presented firstly for static geometrical configurations with constant glottal area and secondly for dynamic geometrical configurations with a time-varying glottal area. During all experiments, the room temperature was $21 \pm 1 ^\circ \text{C}$.
folds mechanical replicas are used to which a uniform mechanical airway wall replica is attached at each end.

![Fig. 4: Mean ePGG signal as a function of source orientation.](image)

Source and sensor are positioned on each airway replica (trachea end and vocal tract end) so that the glottal area of the mechanical replica corresponds to the minimum area of the channel portion between source and sensor. The (minimum) source sensor distance is $d = 150$ mm for the rigid and $d = 257$ mm for the deformable glottal replica. Glottal area $A_g$ is varied in the range 0-55 mm$^2$ (rigid) and 20-100 mm$^2$ (deformable). Measured mean ePGG signals are plotted in Fig. 5. The ePGG signal increases linearly with $A_g$ for both the rigid ($R^2 = 99.2\%$) and the deformable ($R^2 = 98.2\%$) replica. So that ePGG signal and glottal area relate well using a linear approximation. Note that, in general, differences in slope and offset can occur due to 1) positioning of the ePGG system (source-sensor distance $d$, orientation angle) and 2) channel wall properties affecting absorption (thickness, material, etc.). In the next section, time-varying glottal areas are considered.

![Fig. 5: Mean ePGG signal as a function of static glottal area $A_g$ for a rigid and a deformable mechanical glottal replica.](image)

### B. Time-varying glottal area

The correlation between the time-varying ePGG signal and the time-varying glottal area is quantified for the motor driven rigid (oscillation frequency $f_0 \in \{2, 5, 10, 12\}$ Hz, $0 \leq A_g \leq 40$ mm$^2$, $d = 150$ mm) and the flow-driven deformable (fundamental frequency $f_0 \in \{113; 125; 129; 131\}$ Hz for mean upstream glottal pressures $P_u \in \{500, 570, 720, 840\}$ Pa, $20 \leq A_g \leq 100$ mm$^2$, $d = 257$ mm) mechanical replica. Typical examples of correlated time signals for slow (rigid) and fast (deformable) vocal folds displacement are plotted in Fig. 6. Correlation coefficients between ePGG signals and glottal area $A_g(t)$ yield > 90% for the rigid and > 85% for the deformable glottal folds replica.

![Fig. 6: Illustration of correlated time signals of scaled ePGG (full line, ePGG) and glottal area $A_g$ (dashed line, Ref): a) rigid replica at $f_0 = 10$ Hz, b) deformable replica at $f_0 = 129$ Hz.](image)

Consequently, the ePGG signal and glottal area are correlated at all times during the oscillation. In the following section, it is aimed to model the relationship between ePGG signal and glottal area $A_g(t)$ accounting for the different variables affecting the ePGG signal.

### IV. EPGG SIGNAL MODELING

In Section III, it was shown that the ePGG signal is mainly determined by 1) the source-sensor distance, 2) the minimum area of the channel portion between the source and the sensor and 3) the measurement condition determined by the combination of wall properties (e.g. absorption), environment (e.g. light) and ePGG system settings and positioning (e.g. orientation angle). In the following, an ePGG signal model is proposed accounting for each of these factors. Next, its parameters estimation and initialization is outlined. Finally, the sought relationship between ePGG signal and glottal area $A_g(t)$ is discussed.

Following the transillumination principle shown in Fig. 1, ePGG sensor voltage $U$ is proportional to light intensity $I$ at distance $d_k$ from the light source,

$$U(d_k) \propto I(d_k)$$  \hspace{1cm} (1)

Transmitted light intensity $I(d_k)$ at sensor position $d_k$ is then expressed using light flux $\phi$ as

$$I(d_k) = \int_{A_{min}} \phi(d_k)dA,$$  \hspace{1cm} (2)

where
\[ A_{\min}(d_k) = \min_{d \in [0, d_k]} (A(d)) \]  
(3)

is the minimum area encountered by the transmitted light flux between the source position and the sensor position. Furthermore, in Section III was shown that the dependence on \( d \) and \( A_{\min} \) can be described using a first order linear approximation. Consequently, light flux \( \phi(d) > 0 \) can be approximated by model \( \phi(d) \) defined as

\[ \phi_m(d) = \alpha_d d + \beta_d \]  
(4)

with slope \( \alpha_d < 0 \) and offset \( \beta_d > 0 \) (see Section IV). From (2), \( I(d_k) \) is now modeled as

\[ I_m(d_k) = A_{\min}(d_k) \cdot \phi(d_k), \]
\[ = A_{\min}(d_k) \cdot (\alpha_d d_k + \beta_d). \]  
(5)

Inserting (5) in (1) results in modeling the ePGG voltage \( U(d_k) \) as \( U_m(d_k) \) given by

\[ U_m(d_k) = \gamma((\alpha_d d_k + \beta_d) \cdot A_{\min}(d_k)) + \eta \]
\[ = (\alpha_d d_k + \beta_d) \cdot A_{\min}(d_k) + \eta. \]  
(6)

where \( \eta > 0 \) is the remaining signal measured for \( A_{\min}(d_k) = 0 \) and \( \gamma > 0 \) is the scaling factor of (1). For sake of simplicity, let us denote \( \alpha_v = \gamma \alpha_d < 0 \) and \( \beta_v = \gamma \beta_d > 0 \). It is worth noting that \( A_{\min} = 0 \) corresponds to glottal closure so that no direct light is transmitted. Therefore, \( \eta \) is independent of \( d \) and \( A_{\min} \) so that \( \eta \) reflects solely the measurement condition. Considering now a time-variation of the glottal opening, model (6) can be directly extended as

\[ U_m(d_k, t) = (\alpha_d d_k + \beta_d) \cdot A_{\min}(d_k, t) + \eta. \]  
(7)

Consequently using model (7), extracting area \( A_{\min}(d_k(t)) \) from measured ePGG signals \( U(d_k(t)) \) reduces to a parameter estimation problem which is solved using the iterative gradient descent method with appropriate initialization [12].

V. CONCLUSION

An experimental systematic characterization of the ePGG system has been assessed. Main parameters affecting the ePGG outcome are identified. Based on this result, glottal area modulation is studied experimentally on a motor-driven and on a flow driven mechanical larynx replica. It is seen that the correlation between ePGG signal and area measurement yields more than 85%. Next, a calibration method for the ePGG system is proposed. Further research focuses on a systematic validation of the proposed method.

ACKNOWLEDGMENT


REFERENCES