Phonovibrography:
Visualising voice production

DR JÖRG LOHSCHELLER

Voice disorders are caused when vibrations of the vocal folds are disturbed. To research this phenomenon, it is vital that these aberrant vibrations can be analysed. Here, **Dr Jörg Lohscheller** explains the groundbreaking technology that his team has employed to produce the most accurate measurements so far

In order to provide some background to your investigations, what limitations are associated with the current stroboscopic examination of voice production?

Stroboscopy was introduced to provide a visualisation of vocal fold vibrations, and is still the visualisation technique used in clinical practice. Stroboscopy uses low-frequency repetitive light flashes (slower than the fundamental frequency of the vocal folds) that illuminate the vibrating vocal folds at specific points in time. In doing so, a video can be produced that represents a virtual slow-motion recording of vocal fold vibrations.

Problematically, the vibrations of the vocal fold tissue are the fastest process of tissue motion within the human body – their frequency ranges between 80 and several hundreds of oscillations per second. In order to record the vibrations correctly, image sensors are required to provide a sampling rate that is several times higher than the fundamental frequency.

In terms of stroboscopic videos, their interpretation is only valid for highly periodic vocal fold vibrations. The more irregular the vibrations become, the more erroneous is the interpretation of these videos. In addition, it is not possible to derive quantitative parameters that describe all details of vocal fold vibrations from stroboscopic videos.

How will your study differ from previous investigations on vocal fold vibrations?

There are several promising approaches dealing with the quantitative analysis of high-speed recordings using sampling rates of 4,000 images per second and more. Kymography analyses vocal fold vibrations at a specific position. However, it does not provide information about the rest of the vocal folds. Further approaches mainly analyse the time-varying space between the vocal folds. But no differentiation between the left and right vocal fold is possible. By contrast, we have developed the phonovibrography approach which is the only methodology that incorporates the entire visible two-dimensional (2D) vibration pattern of both vocal folds.

Could you give an insight into how the phonovibrography approach works?

With this approach, the motion information of the left and right vocal fold edges, which is extracted from a high-speed video, is transferred into a single static 2D image, called a phonovibrogram (PVG). It can then be visually assessed, although the PVG is not solely for compact visual assessment: the data behind the PVG image can be further used for a computerised, and thus objective, analysis of vocal fold vibrations.

What makes endoscopic high-speed video recording the preferred method for visualising irregular vocal fold vibrations?

Only a high-speed video system using a 2D image sensor is capable of capturing all of the information on vocal fold vibrations along the entire vocal fold length. In physiological voices, the deflection amplitudes as well as the cycle duration of the vocal folds are stable. In pathological voices, the deflection amplitudes and the cycle times become more and more irregular, which leads to a hoarse voice. Depending on the pathology, one or both vocal folds can be affected or the irregularity can just be pronounced at a specific region of a vocal fold. In functional voice disorders, the specific effect of the pathology can only be identified by inspecting the vocal fold during vibration. Thus, high-speed imaging is indispensable for a precise medical examination.

How successful has the implementation of computerised analysis been during your examination of voice disorders?

We have shown that it is possible to automatically describe and classify healthy and pathological vibrations using PVG analysis in a number of studies. In several consecutive works, we have developed a computerised approach that processes a PVG in such a way that the vibration type of vocal folds can be described by a very limited number of clinically meaningful parameters. We showed that these parameters are an objective pendant to the current subjective European Laryngological Society (ELS) classification guideline specifying the type of vocal fold vibration.

Through our computerised PVG analysis, we provide a complete and precise description of the vibration pattern comprising quantitative measures of glottal closure, symmetry, regularity and synchronicity. On the basis of these measurements, objective normative maps of healthy vocal fold vibrations can be derived for the first time. The PVG analysis further successfully enables vocal fold paresis and healthy vocal fold vibrations to be classified with an accuracy of more than 90 per cent. In a further study, it was shown that after surgical removal of vocal fold polyps, the improvement of vocal fold vibration regularity could be quantified using the PVG approach. This is an essential step towards establishing PVG as an objective measure to assess the outcome of conventional voice therapy and surgical interventions.

DR JÖRG LOHSCHELLER

Computerised dysphonia diagnostic

By mapping high-speed movies of vocal fold vibrations into two-dimensional diagrams, an international mutlidisciplinary team is set to help better understand human voice production and treat vocal disorders

THOUGH MANY OF us may take it for granted, verbal communication is a vitally important skill in our post-industrial society. In an economy that is based upon the provision of services, approximately 60 per cent of the working population is reliant upon well-functioning communication skills. With this in mind, it becomes clear why voice disorders might well be a pressing scientific concern. Indeed, voice disorders are quite common, with up to 5 per cent of the general population suffering from them; and more than 10 per cent within the group of 'professional voice users'. Without adequate early diagnosis and timely voice therapy, dysphonia can develop and thus lead to occupational downtimes, accompanied by significant socioeconomic costs.

Voice disorders arise from disturbed vibrations of the vocal folds, which are situated within the human larynx. In clinical practice, vocal fold vibrations are predominantly examined using stroboscopy. Typically, endoscopic videos are assessed subjectively by visual inspection. Yet, somewhat surprisingly, there is still no established clinical standard used to describe and diagnose vocal fold vibrations objectively.

It is in this area of science that Professor Jörg Lohscheller from the University of Applied Sciences Trier in Germany and his team have been channelling their research efforts for the last 10 years. As part of this study, Lohscheller's group intends to provide clinicians with the first objective system for the diagnosis of dysphonia by a computerised analysis of vocal fold vibrations.

IMPRESSIVE AIMS

In assessing the vocal folds' movement patterns, the team hopes to achieve a number of goals. Traditionally, vocal fold vibrations are examined using stroboscopy in combination with a visual assessment of the vocal fold vibration patterns. However, there are some inherent deficits concerning this subjective evaluation of the highly complex vocal fold movements.

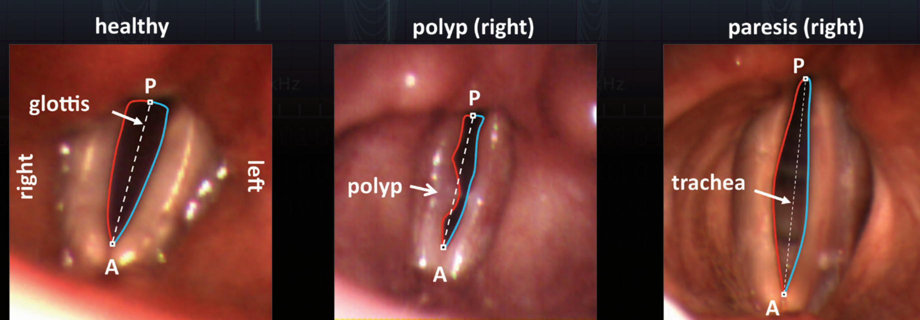

FIGURE 1: Segmentation of the left and right vocal fold edges obtained between the anterior (A) and posterior (P) endings of the visible glottis. Left: Healthy subject, Middle: vocal fold polyp on the right vocal fold, Right: unilateral paresis (right vocal fold)

Assessing the properties of vocal fold vibrations demands a visual rating of the movement characteristics of two vocal folds simultaneously. However, human visual perception is strongly limited in precisely characterising particular features of motion processes. Human beings are instead more capable of capturing and evaluating static geometric shapes. For this reason, considerable inter-rater discrepancies occur in the subjective assessment of vibration properties. Furthermore, since there is a lack of reliable quantitative parameters that describe the vibration pattern of vocal folds precisely, the specific properties of healthy and pathologic voice signals cannot be properly investigated.

It is in this respect that Lohscheller and his team seek to excel. To date, they have provided a computerised system which is able to visualise and precisely quantify the properties of vocal fold vibrations. This has enabled better analysis of the interrelation between vocal fold dynamics and the emitted voice, which is important to their understanding of the principles of voice production. Using this technology, Lohscheller aims to increase our knowledge about vibratory function. This, in turn, will enable scientists to build models of the vocal folds that can be used to study morphologic changes on vibratory function: "This is crucially important not only for the voice quality in human speech and its impairment, but also for fine-tuning the voice quality in human singing, and for the entire field of mammal vocalisation," explains Dr Christian Herbst, one of Lohscheller's collaborators.

PHONOVIBROGRAPHY AND ITS CLINICAL APPLICATIONS

It is the employment of the phonovibrography approach which really sets Lohscheller's team apart. The approach extracts the vibrating vocal folds out of high-speed videos and transfers the information on movements into static geometric shapes visualised within a two-dimensional diagram called a phonovibrogram (PVG). As a result, the entire properties of the vibration patterns of the vocal folds, even the intra-cycle movements, are fully assessable. Since a PVG encodes vocal fold movements as geometric shapes, vibration abnormalities are detectable without any difficulty. Furthermore, by analysing the geometric PVG structures objectively, a precise and objective characterisation of vocal fold vibration types can be established. Thus, PVG maps out the different reasons for dysphonia, capable of quantifying the extent of the pathology.

Phonovibrography's groundbreaking method of visualising vibrations is by no means its only impressive feature. Indeed, this method could well have further impressive clinical applications too. At present there is no standardised procedure to describe and quantify vocal fold vibrations, making examination results collected from different clinics difficult to compare. Therefore, phonovibrography has an ideal opportunity to be the gold standard in analysing vocal fold vibrations in future. The PVG parameters representing the vibration type can be used to generate normative maps of healthy vocal fold vibrations that allow the precise classification of the individual vibration type of a patient.

Also, due to its efficient implementation, it does not require excessive computer time to generate a PVG from a high-speed video. Thus, PVG visualisation and analysis can be applied directly after the endoscopic examination in the presence of the patient. The PVG image can then be used to explain the result of the analysis directly to the patient. Phonovibrography has further potential to assess the outcome of surgery and therapeutic interventions and, in contrast to a high-speed video, a PVG can easily be documented within a patient's record. Overall, PVGs provide objective information, which is especially useful with conditions that require numeric data – such as measuring blood pressure or categorising the pathology.

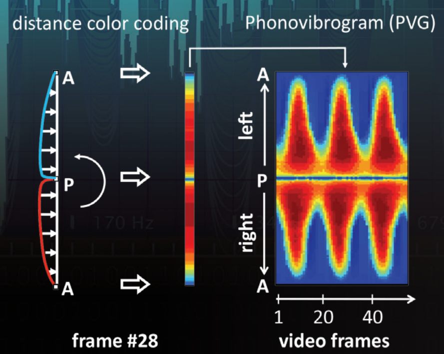

FIGURE 2: Process of PVG construction: For each frame of a high-speed video the distances between the vocal fold edges and the glottal midline are color coded

STRENGTHS AND WEAKNESSES

In terms of the developments that the team has already made to the phonovibrography approach, one major strength is the large amount of processed clinical data and the possibility to clearly differentiate abnormal pathologies from normal ones; thereby demonstrating its applicablity in measuring therapy effects. Lohscheller has visibly shown that their approach will give more detailed insight into larynx function than anyone could have imagined before. Another main strength of the PVG is that it is very easy to understand and is therefore very effective in aiding communication between doctors and patients. It has the potential to be hugely helpful in clinical practice. A PVG incorporates all information about the two-dimensional vibration pattern and, although it constitutes an immense data reduction can act as a precise fingerprint of vocal fold vibrations. Yet arguably the greatest strength of the phonovibrography approach is the ability to derive clinically meaningful parameters that describe the vibration pattern of the vocal folds in the most complete way. Simultaneously, it allows the compression and visual expression of these characteristics in a plain image.

However, there are still some barriers to the clinical adoption of the phonovibrography approach. Currently, in clinical practice rigid endoscopes are used during endoscopic examination. Since the endoscope can be distracting in the oral cavity, it influences the process of voice production itself – causing, for example, unnatural muscular tension within the larynx and a potential gag reflex. This can influence the result of an examination. The process of acquiring and transferring high-speed videos from the camera to the computer is another challenge to solved because at present it is quite time-consuming for a busy clinical practice.

Thus, Lohscheller and his colleagues seek to improve and refine the technique as part of their future work. The solution to the aforementioned issues, they think, will be to use flexible endoscopes which can be inserted through the nose. Flexible endoscopy would affect the process of voice production less and would consequently lead to an improved PVG analysis and heighten the quality of medical examination procedures.

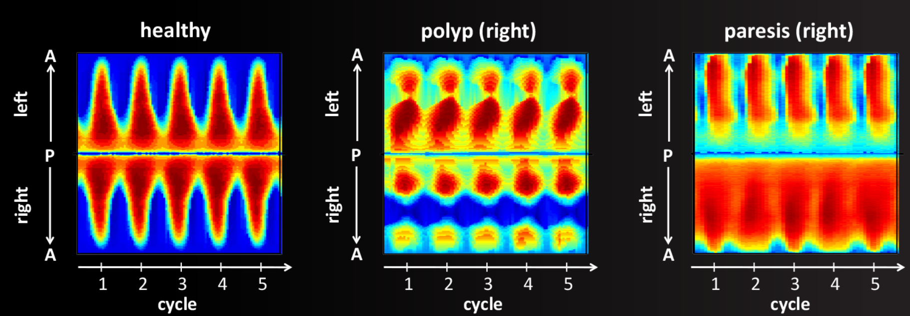

FIGURE 3: Visualisation of vocal fold vibrations using Phonovibrograms (PVG) obtained from the subjects shown in figure 1. Within the PVG the characteristic properties of the vocal fold vibrations are transfered into distinct geometric PVG patterns.

INTELLIGENCE

PHONOVIBROGRAPHY: DEVELOPMENT OF A COMPUTERISED SYSTEM FOR THE DIAGNOSIS OF DYSPHONIA

OBJECTIVES

- To develop a computer-aided analysis system for voice disorders, providing a new way of quantifying, differentiating and classifying these conditions
- To validate this system in close cooperation with national and international partners from medical and scientific disciplines

KEY COLLABORATORS

Christian Herbst, PhD, Department of Cognitive Biology, University of Vienna, Austria

Melda Kunduk, PhD, Department of Communication Sciences and Disorders, Louisiana State University, Baton Rouge, USA

Prof. Dr Bernhard Schick, MD, Department of Otorhinolaryngology, Saarland University Hospital Homburg/Saar, Germany

Dr Maria Schuster, MD, ENT Clinic, Ludwig-Maximilians-University Munich, Germany

Jan Svec, PhD, Department of Biophysics, Faculty of Sciences, Palacky University Olomouc, Czech Republic

FUNDING

Deutsche Forschungsgemeinschaft (DFG) project number LO1413/2-1/2

CONTACT

Professor Dr Jörg Lohscheller
Project Coordinator

University of Applied Sciences Trier
Department of Computer Science
Medical Informatics
Schneidershof
54293 Trier
Germany

T +49 651 8103 578
F +49 651 8103 454
E j.lohscheller@hochschule-trier.de

PROFESSOR DR JÖRG LOHSCHELLER

studied electrical engineering at RWTH Aachen University, Germany. He received his doctoral degree in engineering (with highest distinction) from the University of Erlangen-Nuremberg in 2003 and, since 2009, has been full Professor of Medical Informatics at the University of Applied Sciences Trier, Germany. In addition, Lohscheller has been Adjunct Professor at the Louisiana State University, Baton Rouge, USA since 2008. He has also been a member of the medical faculty of Saarland University since 2011.

Deutsche Forschungsgemeinschaft

HOCHSCHULE TRIER
für Technik, Wirtschaft und Gestaltung
Trier University of Applied Sciences