Foreword

For Signal and Acoustic Modeling for Speech and Communication Disorders

This fine collection of articles from 10 renowned research groups across the world brings together work related to the analysis and detection of relevant conditions and behavior of humans via use of their speech.

Recent advances in the technology of speech signal processing have the potential to assist in the detection and treatment of a vast range of human disorders, most notably in areas of speech and communication. Until recently, speech analysis has been largely dedicated to coding and recognition applications, for example, the vast network of cellular telephony and the many products and services that allow users to communicate with computers via voice. Such services almost always assume normal speech communications, that is, the user employs a normal voice. However, in the medical field, there is a great need to diagnose various abnormal conditions in humans, and their vocal output has been a largely untapped resource. Surely, various medical tests such as X-rays, temperature, and blood pressure readings, etc., serve as a primary window to human conditions. Nonetheless, one's speech is a very simple indicator of the state of one's health. Till recently, analysts relied only on their intuition as to how to interpret a patient's speech. With recent advances, we can expect to give useful computer analysis tools to such personnel. The set of speech difficulties that one can hope to assist cover the range of congenital defects, developmental disorders, degenerative diseases, and disabling neuromotor conditions.

This anthology brings together researchers from leading universities and laboratories across five continents. The contributors to this volume examine acoustic modeling across a wide range of communication disorders. Causes include Parkinson's, Autism Spectrum Disorder, cleft palate, intellectual disabilities, neuromotor disorders (e.g., dysarthria), stroke, tumors, brain injury, and reactions to medication. The works here use advanced innovations to show how recent speech technology can benefit healthcare.

The methods described are non-invasive and cost-effective, as they can detect, monitor, and gauge speech, language, and communication disorders, without any effect on patients and with minimal equipment (i.e., microphone and computers).

This book has three sections. The first section discusses how specific developments in speech signal processing can be highly instrumental in diagnosing and treating speech and communication disorders, as well as in detecting early signs of medical problems that may impair neuromotor functioning and compromise speech intelligibility.

The first chapter analyzes whispered speech – a common type of vocal output that has been little examined in the past. People with vocal cord problems often

resort to whispering, and others employ such when trying to speak very softly. Study of whispering may help detect heart-related problems.

The second chapter examines hypernasality – a speech quality owing to defective velopharyngeal closure; speech analysis can estimate the degree of congenital abnormality of the lip and palate, and assist in assessing neuromuscular and hearing problems.

The third chapter in the first section of the book examines recent advances in automatic acoustic analysis toward diagnosis of speech disorders. One area of recent work deals with analysis of prosody (intonation, i.e., the durations, pitch and intensity of speech), for which behavior of normal and disordered speech often differs, for example, in speaking rate. The method discussed in this chapter has been successfully tested in several European languages. This chapter also delves into chaos theory, exploiting such aspects of speech to distinguish normal and pathological voices.

The second major section of the book examines how speech relates to cognitive, affective, and developmental disorders. The first chapter in this section compares the speech of children and adults with intellectual disability, to the speech of normal speakers. While language impairments as detected by professionals are a primary way to discover language problems, there is great variability in the acoustic nature of such abnormalities, leading to much difficulty of proper diagnosis, that is, it is easy for even professionals to miss subtle acoustic signs that automatic analysis may find.

Speakers use, in complex ways, both prosody and spectro-temporal features to communicate all sorts of information via speech, most notably, they communicate ideas to humans listening. This occurs via the syntactic, semantic, and pragmatic content of their words. However, the speech signal contains far more information, encoded in highly complex ways, for example, cues to who is speaking, what emotional and medical state they are in, and what language they are speaking.

This section's first chapter shows a speech classification system that differentiates reliably between speakers with an intellectual disability and those with normal cognitive functioning. The next two chapters deal with Autism Spectrum Conditions (ASC). One provides a survey of autism and speech, language, and emotion. People with ASC vary their spoken content in ways often different from normal speech. The second chapter on autism examines clinical applications of speech technology.

The third major section of the book presents studies in the assessment and quantification of speech intelligibility in patients with anatomical defects, disabling conditions, and degenerative diseases. The studies in this section cover both congenital and acquired speech disorders. The first chapter in this section

does automatic assessment of consonant omission and speech intelligibility in cleft palate speech, for 530 participants speaking Mandarin.

Next is a study of distinctive auditory-based cues and rhythm metrics to assess the severity of patients with dysarthria, from either acquired or congenital neuromotor communication disorders that affect their speech intelligibility. The diagnosis method in this chapter uses a model simulating the outer, middle, and inner parts of the ear. Relevant auditory-based cues are combined with Mel-Frequency Cepstral Coefficients (MFCCs) and rhythm metrics. MFCC has been used for decades as a standard way to analyze speech in automatic speech recognition (ASR.)

The Nemours and Torgo databases of dysarthric speech are used. Standard methods of ASR, such as Gaussian mixture models (GMMs), support vector machines (SVMs), and multinomial logistic regression (MLR) are tested for this dysarthric speech assessment.

The third section of the book concludes with an analysis of how wireless sensor networks and embedded systems can be used to measure the motor defects of patients suffering from Parkinson's disease. This chapter shows how data fusion between a sensor network and an embedded system can help quantify the disease. This is particularly relevant today, as healthcare systems increasingly use ubiquitous and pervasive computing to monitor patients at a distance.

The editors of this volume have done a fine job in reviewing recent international work in this field. They have presented here the most current and advanced research findings in signal and acoustic modeling of speech pathologies. The target audience for this book is speech scientists, clinicians, and pathologists.

Prof. Douglas O'Shaughnessy
(Ph.D. MIT USA, IEEE Fellow and ASA Fellow),
McGill University,
Canada