

VISUAL FEEDBACK OF ACOUSTIC VOICE FEATURES IN VOICE TRAINING

C. William Thorpe

School of Communication Sciences & Disorders,
University of Sydney

ABSTRACT: Speech technology has long been utilised to within computer-assisted tools for voice training. Such tools, providing real-time visual feedback of specific characteristics of the voice, complement traditional training methods. However, in order for such technology to be successful, specific aspects of voice training must be considered including the nature of voice perception, motor learning processes, and what information can be reliably extracted from the voice signal. This paper examines the factors that must be considered in developing voice training technology, and provides an overview of the techniques that have been applied to extract information from the voice signal and present it visually for real-time training purposes.

INTRODUCTION

A potentially exciting application of speech technology is the provision of visual feedback in voice training. Traditionally, voice training in areas such as singing, acting, and speech therapy, is based on a model of practice coupled with teacher-provided feedback. The teacher instructs the student as to what kind of performance to practice, and provides feedback about how well his or her performance matches the desired target ideal. Typically, the feedback concerns aspects of either the acoustic quality of the vocalisation, or physiological aspects such as breathing. However, even with highly competent teachers, any feedback given by the teacher has limitations, which may reduce the learning efficiency. Firstly, the teacher must be able to perceive a particular quality of performance in order to provide comment. Some aspects of physiology are of course invisible to the observer, even by manual sensation, and even some acoustic aspects may be difficult to perceive precisely except in highly trained individuals (eg subtle differences between vowels within categorical spaces). Secondly, the teacher must communicate the feedback to the student, either in words, by touch, modelling a "correct" performance, or perhaps by means of metaphor. Thirdly, the student must be able to understand the feedback and use it to alter the subsequent performance in order to effect the desired improvement. Finally, the teacher may be in a classroom situation where attention to individual students is by necessity limited.

Providing ancillary feedback of particular aspects of a student's performance offers the possibility to overcome some of the limitations of traditional teacher-provided feedback. Firstly, such feedback is not limited by (auditory) perceptual processes, but may highlight aspects of the performance that are not easily perceived. Secondly, the feedback can be communicated directly in a visual format that directly mirrors the performance, and furthermore, this feedback can be contemporaneous with the student's vocalisation. This implies that it may be easier to integrate the knowledge gained into improvements in performance, compared to feedback that is delayed.

The purpose of this paper is to provide an overview of the various issues involved in providing visual feedback for the purposes of voice training. In Section 1, issues relating to learning processes themselves are discussed, followed by an overview of how the characteristics of voice perception need to be taken into account during feedback. Section 3 describes the types of acoustic analyses that are available to extract relevant information from the voice, while Section 4 describes ways in which this information can be visually presented to the student. Finally, some of the pedagogical considerations of incorporating computer technology in voice training are discussed in Section 5. The focus here (and the examples included) is on the application to singing, but the factors considered are also broadly relevant to other vocal training tasks including accent training, speech therapy, etc.

1. LEARNING AND FEEDBACK

Vocal training can be considered as a form of motor training in which it is desired to obtain extremely accurate control of all the muscle groups involved in vocalisation. The training process involves a series of trial productions, with feedback about each trial being used to improve subsequent productions. This model of learning involves both internal and external feedback mechanisms. The

Proceedings of the 9th Australian International Conference on Speech Science & Technology
Melbourne, December 2 to 5, 2002. © Australian Speech Science & Technology Association Inc.

bhewitt
C. William Thorpe

