

A DYNAMIC GLOTTAL MODEL THROUGH HIGH-SPEED IMAGING

Jiangping Kong*

Peking University

ABSTRACT

This paper is a study for an improved dynamic glottal model through high-speed imaging (HSI).¹ As is well known, speech production comprises three parts, namely speech source, speech resonance and lip radiation. Among these three parts, speech source is the most important one because it is the basis of speech. In research on speech production, acoustical models of speech source have been well established. But the physiological speech source, that is to say, the activity of the glottis is seldom researched, because the vibration of the vocal folds is difficult to observe and sample. A study on the glottal model was established many years ago (Kong 2007), and in that model, the static glottis was modeled by four quarters of ellipses in three modes namely normal mode, leakage mode and open mode. The dynamic glottal control function was modeled by an approximation of multiplication of sine and exponential. The problem of the dynamic glottal model is that the control parameters can't be well explained, though the glottis can be simulated. In this study, more high-speed images were sampled, the image processing was greatly improved and the dynamic glottal control function was modeled with parameters which were significant to speech perception.

SUBJECT KEYWORDS

High-speed imaging Vibration of vocal folds Dynamic glottal model

* **Jiangping Kong** 孔江平 is a professor of Peking University in the Center for Chinese Linguistics, Department of Chinese Language and Literature, and Joint Research Center for Language and Human Complexity. Address: Dept. of Chinese Language and Literature, Peking University, Beijing 100871, China; [jpkong@pku.edu.cn]

基於高速數位成像的動態聲門模型

孔江平*

北京大學

提要

本文利用高速數位成像技術對動態聲門模型進行了研究。眾所周知，言語產生包括嗓音聲源、聲道共鳴和唇輻射三個方面，其中嗓音聲源尤其重要，因為嗓音聲源是言語產生的基礎。在言語產生的研究中，聲學模型已經有了很深入的研究，但由於聲帶振動難於觀察和採集樣本，嗓音的生理模型研究的很少。多年前作者建立了一個動態聲門模型（Kong 2007），在此模型中，靜態聲門是用四個四分之一橢圓來建模的，並有正常、漏氣和敞開四種模式。模型的動態聲門控制函數是通過正弦和拋物線的乘積來建模。雖然這種方式有效，但合成嗓音的參數解釋性較差。在本項研究中，採集了更多更高品質的聲帶振動高速數位成像樣本和大大改進了數位影像處理的技術，最終模型的動態聲門控制函數所用的參數對嗓音聲源的感知具有很好的解釋性。

主題詞

高速數位成像 聲帶振動 動態聲門模型

* 孔江平為北京大學教授，執教研究于中文系，中國語言學研究中心，語言與人類複雜系統聯合研究中心；地址：中國北京，郵編 100871，北京大學中文系；[jpkong@pku.edu.cn]