Available online at www.sciencedirect.com

SciVerse ScienceDirect

Speech Communication 55 (2013) 667–690

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

A review of lumped-element models of voiced speech

Byron D. Erath [a,*], Matías Zañartu [b], Kelley C. Stewart [c], Michael W. Plesniak [c], David E. Sommer [d], Sean D. Peterson [d]

[a] Department of Mechanical and Aeronautical Engineering, Clarkson University, Potsdam, NY 13699, United States
[b] Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile
[c] Department of Mechanical and Aerospace Engineering, The George Washington University, Washington DC 20052, United States
[d] Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada

Received 11 September 2012; received in revised form 6 February 2013; accepted 8 February 2013
Available online 21 February 2013

Abstract

Voiced speech is a highly complex process involving coupled interactions between the vocal fold structure, aerodynamics, and acoustic field. Reduced-order lumped-element models of the vocal fold structure, coupled with various aerodynamic and acoustic models, have proven useful in a wide array of speech investigations. These simplified models of speech, in which the vocal folds are approximated as arrays of lumped masses connected to one another via springs and dampers to simulate the viscoelastic tissue properties, have been used to study phenomena ranging from sustained vowels and pitch glides to polyps and vocal fold paralysis. Over the past several decades a variety of structural, aerodynamic, and acoustic models have been developed and deployed into the lumped-element modeling framework. This paper aims to provide an overview of advances in lumped-element models and their constituents, with particular emphasis on their physical foundations and limitations. Examples of the application of lumped-element models to speech studies will also be addressed, as well as an outlook on the direction and future of these models.

Keywords: Lumped-mass models; Glottal aerodynamics; Acoustics; Vocal fold models; Vocal tract; Subglottal system; Acoustic interaction

Contents

* Corresponding author. Tel.: +1 3152686584.
E-mail addresses: berath@clarkson.edu (B.D. Erath), matias.zanartu@usm.cl (M. Zañartu), kstewart@gwu.edu (K.C. Stewart), plesniak@gwu.edu (M.W. Plesniak), peterson@mme.uwaterloo.ca (S.D. Peterson).

0167-6393/$ - see front matter
http://dx.doi.org/10.1016/j.specom.2013.02.002

1. Introduction

The widely accepted myoelastic-aerodynamic theory of vibration states that vocal fold (VF) vibrations are produced by a coupling between the aerodynamic forces and the tissue parameters (van den Berg, 1958), which generates a rich spectrum of acoustical sound that can feed back and influence both the VF dynamics and the fluid flow. Interest in the field of voiced speech production can be traced back nearly a century (Wegel, 1930) and arises primarily from a desire to understand the basic physics of phonation, as well as to provide diagnosis and treatment to individuals that suffer from any number of vocal pathologies. Due to the complexity of voiced speech, simplified models are often employed for both scientific and clinical investigations. As with all modeling practices, scientific speech investigations seek to reduce the complex problem down to the essential physical foundation to gain traction in understanding the phenomena. This is, of course, a fine balancing act between reducing the model enough to be tractable, but not so far as to be trivial.

The most common modeling framework in voiced speech investigations is the so-called lumped-element approach, wherein the VF structure is modeled as a collection of discrete coupled mass-spring-damper systems subjected to some external aerodynamic and/or acoustical loading function. Generally speaking, in a damped vibrating system such as the VFs, the energy imparted to the system must be as great as the energy dissipated by it in order to sustain self-oscillation. This occurs physiologically, as the aerodynamic loading is physically coupled to the VFs and supplies energy to the system during each oscillation cycle. Lumped-element models have proven capable of emulating the physiological VF kinematics and acoustic output. In the lumped-element framework, the dynamical response of the VFs is governed by a system of coupled nonhomogeneous autonomous ordinary differential equations. The number of equations needed to describe the system is directly related to the number of degrees of freedom employed in the model. The system of equations can usually be solved with relative ease using an explicit, forward time-marching scheme, with appropriate initial conditions.

Lumped-element VF model investigations were initially utilized because of the low computational cost that enables quick, efficient calculations. In spite of ever-increasing computer speeds, they still remain an attractive approach for scientific speech investigations due to their clinical significance and their ability to perform broad parametric investigations relatively quickly (Erath et al., 2011b; Steinecke and Herzel, 1995). In addition, as lumped-element models become more refined, the prospect of developing short run-time models that can mimic patient-specific phonatory conditions becomes more realistic.

The objective of this paper is to review advancements in lumped-element models of speech, focusing specifically on