# 3D biomechanical modelling for speech production

Jean-Michel GERARD
Institut de la Communication Parlée – UMR CNRS 5009
INP Grenoble & Université Stendhal Grenoble

Speech production is the result of complex interactions between articulators and motor control. Yet, although data processing tools and measurement methods have made great improvements, they still have limitations in helping to elaborate theories that explain mechanisms involved in the speech production process. In 1969, Perkell was the first to design and use a biomechanical tongue model as a tool to study speech production, and nowadays, the use of such models to study speech production is still common (Laboissière et al. (1996), Payan & Perrier (1998), Wilhelms-Tricarico (2000), Dang & Honda (2003), Gérard et al. (2003)).

Setting up a model is quite a complex task, and given the great number of possible approaches, several questions have to be answered prior to building the model. What is to be studied? What do we have to model and which level of complexity do we need to achieve it? The answer to those questions will determine the choice between two general frameworks. On the one hand, geometrical models, using a statistical description of deformations, are usually used for animation of talking heads and can be useful in studying some speech production phenomena. On the other hand, the biomechanical models, more complicated to design and to use, give us the opportunity to study the interaction between mechanics and motor planning, that is, tongue deformations along with their causes. Each has advantages and drawbacks, and the key point is to make an adequate compromise between the level of complexity and the problem to be studied. Complex models will generally imply a complicated design with huge computation times and algorithmic issues in terms of convergence processes during simulations, even if mathematic techniques can reduce those issues.

In biomechanical modeling, several calculation techniques can be used: the Finite Element Method (FEM), the Finite Difference Method (FDM) or the Mass-Spring system. The simplicity of the Mass-spring system is counterbalanced by its lack of accuracy in obtained results and instabilities. The variability of phenomena in speech production makes the study of those phenomena interesting in several respects, requiring an adaptation of the model to various speakers. Currently, matching techniques can easily adapt a generic mesh to the geometry of any speaker starting from MRI or CT-scan data. Thus the FEM, which uses a mesh, is generally used, rather than the FDM, which uses surfaces. The choice between a 2D or 3D description of the tongue is guided by the kind of study to be conducted. For example, a 3D representation will allow the study of deformations in a transverse plane, but will make the computation more complex. Temporal variations of tongue deformations, and therefore variations of the acoustic signal, are important for a good understanding of the meaning contained in the signal. Time can be partly controlled by motor planning but is also constrained by the mechanics of the system (inertia, damping). The mechanical properties of tongue tissues feature viscoelasticity, anisotropy and a highly nonlinear behavior. The lack of data published about tongue mechanical properties led us to make our own measurements. One method could be to use simple mechanical properties as a first step, and in case of inadequate accuracy, make the description more complex, step-by-step.

In the talk, each step of a biomechanical model design will be discussed and some answers to all those issues will be given, illustrated by our work on a 3D tongue model currently in development at ICP.