



HAL
open science

Pinocchio, a biomimetic mechatronic testbed for producing in vitro articulated speech

Nathalie Henrich Bernardoni, Julien Royer, Mounib Tlaïdi, Xavier Laval, Sylvain Arnaud, Lucie Bailly

► To cite this version:

Nathalie Henrich Bernardoni, Julien Royer, Mounib Tlaïdi, Xavier Laval, Sylvain Arnaud, et al.. Pinocchio, a biomimetic mechatronic testbed for producing in vitro articulated speech. 13th International Seminar on Speech Production, May 2024, Autrans, France. hal-04579018

HAL Id: hal-04579018

<https://hal.science/hal-04579018>

Submitted on 17 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Pinocchio, a biomimetic mechatronic testbed for producing *in vitro* articulated speech

Nathalie Henrich Bernardoni¹, Julien Royer¹, Mounib Tlaidi¹,
Xavier Laval¹, Sylvain Arnaud¹, Lucie Bailly²

¹Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

²Univ. Grenoble Alpes, CNRS, Grenoble INP, 3SR, 38000 Grenoble, France

nathalie.henrich@gipsa-lab.fr, lucie.bailly@3sr-grenoble.fr

Introduction.

Speech production results from a fine neuro-controlled coordination between breathing, phonatory and articulatory movements. In physical terms, these movements induce fluid-structure-acoustic interactions within the vocal tract. In the framework of source-filter theory (Fant 1960), the aerodynamic phase of speech is often overlooked, yet being of much importance for voiced and unvoiced speech sound production (Catford 1977). In the case of voiced speech sound, it is taken into account by myoelastic-aerodynamic theory of phonation (Jan G. Švec et al. 2021), in which glottal constriction and vocal folds vibration generate aeroacoustic sources. Our aim is to advance our understanding of speech sound production by developing a biomimetic *in vitro* testbed. Over the last thirty years, vocal-fold testbeds have evolved in complexity and biomimicry. However, most testbeds explore the physics of phonation on geometrically-fixed replicas capable of self-sustained oscillations in fluid-structure interaction, but unable to produce intonative variations. Most of the time, the testbeds are not coupled to vocal tract, and whenever they are, the resonant cavities are static 3D-printed tracts. Reproducing *in vitro* the dynamic movement of speech articulators, such as jaw, tongue, velum and larynx, together with phonation, remains a challenge. We present here the first steps in the design of a biomimetic mechatronic testbed that would integrate all phonatory and articulatory aspects important for voice and speech production.

Methods.

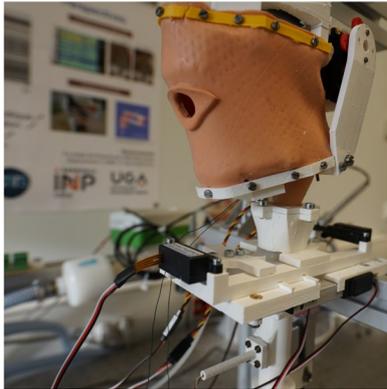
Design efforts focused on three aspects : how to breath, how to phonate, how to articulate.

The first aspect to take into account is the ability to control air supply through flow and pressure parameters. The system controls a valve opening, which is linearly related to airflow rate at the device inlet. Subglottal pressure can also be programmed using a servo loop. To stabilise the upstream jet, a settling chamber is used. A long tube leading from the chamber to the testbed avoids acoustic coupling with the subglottal tract (Lehoux, Hampala, and Jan G Švec 2021).

Choice was made to design a flexible laryngeal envelope combined with a 1:1 scale vocal-folds replica, so as to allow articulatory movements within the larynx and vocal fold stretching. In a first approach, the folds were designed to be inserted interchangeably while maintaining a seal, with the aim of gradually improving their biomimetic characteristics (Luizard et al. 2023). In a second approach, the folds were moulded together with the laryngeal envelope. Several aspects were taken into account and tailored : folds geometry, material tensile stiffness, folds anisotropy. Folds geometry was derived from common M5 models (Murray and Thomson 2012; Luizard et al. 2023). All tested folds were made of either commonly-used silicone elastomers (Ecoflex™series with increasing degrees of shore hardness) or gelatin-based hydrogels. Body and cover were homogenous (moulded together) or heterogenous (moulded apart with different stiffness, and without or with addition of a fibrous structure). The self-oscillation capabilities of each laryngeal replica were assessed.

First implemented articulatory movements were within the larynx itself, in the framework of Laryngeal Articulator Model (Esling et al. 2019). The vocal folds can be adducted or abducted, imitating the phases of breathing and phonation. They can be stretched and compressed into the anterior-posterior direction, and compressed or uncompressed into the medial-lateral one. Secondly, a geometrically-realistic and articulatory-driven vocal tract was designed, inspired from Arai's vocal-tract models (Arai 2016). A FE model of the tongue was used to design the tongue mould (Hermant, Perrier, and Payan 2017). Two primary articulatory movements were implemented : mandibular motion and shaping of the oral cavity with the movement of tongue body and apex.

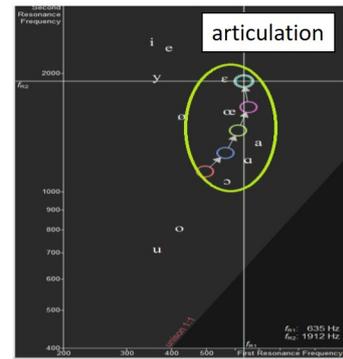
A. Pinocchio vocal-tract, jaw and cheeks



Tongue



C. Behavior assessment



B. Pinocchio larynx



vocal folds

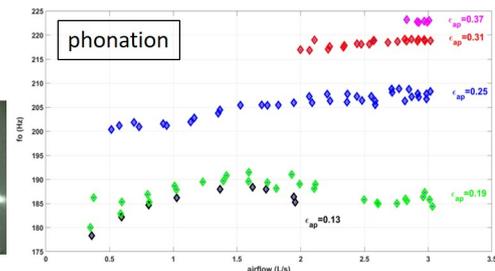
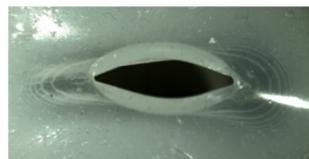


Figure 1: Presentation of the pinocchio testbed.

Results and discussion.

Figure 1 presents the complete system, called *pinocchio*, and its phonatory and articulatory capabilities.

Phonation The vocal-folds replica were able to self-oscillate with and without a vocal-tract, over a wide range of flow rates and for different degrees of tensile stretch. First main factor of variability in vibratory behaviour was the folds geometry. While long (20mm at rest) and thick (4mm) folds did not allow a self-oscillation frequency in the expected range of human speech (Luizard et al. 2023), shortening the length at rest (down to maximum 12mm) and slimming the fold (down to 1.5mm) achieved the goal of a self-oscillation frequency up to the range of female speech (f_o from 180 to 220 Hz) as shown in Figure 1C. The latest version of the vocal-fold replica allowed for increasing f_o with increasing pre-strain in the anterior-posterior direction (ϵ_{ap}).

Articulation We explored the vocalic space produced by movements of jaw, tongue body and apex movements (see Figure 1C top panel). F1 varied moderately between 500 and 700 Hz, mainly in relation to mouth opening induced by jaw movement. F2 varied between 1000 and 2000 Hz with tongue and jaw movements. At this early stage of his life, Pinocchio was able to produce vowels in the surrounding of [a].

References.

- Arai, Takayuki (2016). "Vocal-tract models and their applications in education for intuitive understanding of speech production". In: *Acoustical Science and Technology* 37.4, pp. 148–156.
- Catford, John Cunnison (1977). *Fundamental Problems in Phonetics*. Indiana University Press.
- Esling, John H, Scott R Moisiak, Allison Benner, and Lise Crevier-Buchman (2019). *Voice quality: The laryngeal articulator model*. Vol. 162. Cambridge University Press.
- Fant, G. (1960). *Acoustic Theory Of Speech Production*. The Hague: Mouton.
- Hermant, Nicolas, Pascal Perrier, and Yohan Payan (2017). "Human tongue biomechanical modeling". In: *Biomechanics of Living Organs*. Elsevier, pp. 395–411.
- Lehoux, Hugo, Vít Hampala, and Jan G Švec (2021). "Subglottal pressure oscillations in anechoic and resonant conditions and their influence on excised larynx phonations". In: *Scientific reports* 11.1, pp. 1–14.
- Luizard, P, L Bailly, H Yousefi-Mashouf, R Girault, L Orgéas, and N Henrich Bernardoni (2023). "Flow-induced oscillations of vocal-fold replicas with tuned extensibility and material properties". In: *Scientific reports* 13.22658, pp. 1–19.
- Murray, Preston R. and Scott L. Thomson (2012). "Vibratory responses of synthetic, self-oscillating vocal fold models". In: *The Journal of the Acoustical Society of America* 132.5, pp. 3428–3438. (Visited on 09/21/2022).
- Švec, Jan G., Harm K. Schutte, C. Julian Chen, and Ingo R. Titze (2021). "Integrative Insights into the Myoelastic-Aerodynamic Theory and Acoustics of Phonation. Scientific Tribute to Donald G. Miller". In: *Journal of Voice*.