*INTERSPEECH 2016*
September 8–12, 2016, San Francisco, USA

# Bayesian modeling in speech motor control: a principled structure for the integration of various constraints

*Jean-François Patri*[1,2,3,4], *Pascal Perrier*[1,2], Julien Diard[3,4]

[1]Univ. Grenoble Alpes, Gipsa-lab, F-38000 Grenoble, France
[2]CNRS, Gipsa-lab, F-38000 Grenoble, France
[3]Univ. Grenoble Alpes, LPNC, F-38000 Grenoble, France
[4]CNRS, LPNC, F-38000 Grenoble, France
Jean-Francois.Patri@gipsa-lab.fr

## Abstract

Speaking involves sequences of linguistic units that can be produced under different sets of control strategies. For instance, a given phoneme can be achieved with different acoustic properties, and a sequence of phonemes can be performed at different speech rates and with different prosodies. How does the Central Nervous System select a specific control strategy among all the available ones? In a previously published article we proposed a Bayesian model that addressed this question with respect to the multiplicity of acoustic realizations of a sequence of phonemes. One of the strengths of Bayesian modeling is that it is well adapted to the combination of multiple constraints. In the present paper we illustrate this feature by defining an extension of our previous model that includes force constraints related to the level of effort for the production of phoneme sequences, as it could be the case in clear versus casual speech. The integration of this additional constraint is used to model the control of articulation clarity. Pertinence of the results is illustrated by controlling a biomechanical model of the vocal tract for speech production.

**Index Terms**: Speech motor control, Bayesian modeling, Hypo/Hyperspeech, Dynamical constraints.

## 1. Introduction

A number of various experimental studies have suggested that speech is a skilled serial-order motor task that is adapted to achieve time series of goals within a timing that does not allow any on-line cortical processing of feedback signals. This is illustrated for the auditory feedback by the absence of impact on speech production of a feedback delayed by up to 75 ms [1]. In addition, the speech motor system is highly redundant with many available degrees of freedom, which makes the inference of motor commands from physical signals an “ill-posed” inverse problem [2]. These degrees of freedom are used in different ways to deal with sequence planning and variations in the conditions of articulation. For instance, clear anticipatory behaviors have been observed in articulatory and acoustic patterns associated with speech sequences [3].

To deal with this complexity, motor control models classically consider a feedforward control scheme for ongoing speech production under normal conditions. This is associated with a feedback controller enabling, with a certain delay, a correction of the motor commands in case of inadequate feedforward commands or in presence of external perturbations [4]. In this context, speech planning, which aims at solving the “ill-posed” inverse problem by finding the motor command patterns adapted to the production of a speech sequence, has been classically modeled within an optimal motor control framework. This approach has been shown to generate results in close agreement with experimental data, in particular in terms of adaptation to perturbations [5] or in terms of anticipatory behavior [6, 7].

However, criticisms of this approach include key issues in cognitive science, such as the nature of the neuro-physiological mechanisms likely to be associated with cost computation and cost minimization, or the inability to account for the well-known token-to-token speech variability [8]. This last drawback is inherent to the feedforward optimal control scheme, since this scheme basically cancels possible variations along the degrees of freedom directions, by specifying a unique optimal solution to the control problem.

In this context we have published an alternative approach by formulating feedforward optimal control in a Bayesian modeling framework [9], which has been implemented for the control of a 2D biomechanical model of the tongue [10]. The methodology is based on Bayesian Programming [11, 12, 13], which proposes a structure for the construction of Bayesian models. Bayesian modeling, in a nutshell, solves inference problems by computing probability distributions for the solutions instead of proposing specific values; this allows to solve ill-posed problems while conserving token-to-token variability in a principled way [14]. It also preserves the basic principles underlying the search for optimality without being explicitly driven by the minimization of a cost, relying instead on probabilistic evaluation of all possibilities.

The approach was illustrated by reformulating GEPPETO, an existing optimal control model for speech production planning developed in the lab [15], into the Bayesian modeling framework. We have shown that these two models are nested, with optimal control as a special case of the Bayesian model: indeed, the Bayesian model is simplified into the optimal control model, when the inferred control commands are strictly limited to those with maximum posterior probability. Variability in the Bayesian model is formally generated by assuming that control is performed by sampling motor commands randomly according to the distribution that solves the inference problem.

An interesting aspect of the proposed formalism is its coherence for dealing with multiple constraints. This is illustrated in the present work by describing an extension of the Bayesian model that takes into account an important additional constraint associated with the trade-off between accuracy and effort at fast speaking rate, which has also been integrated in GEPPETO [16]. While our recently published model specified the con-

http://dx.doi.org/10.21437/Interspeech.2016-441

straints only in terms of auditory goals to be reached at the different phonemes of a speech sequence, the proposed extension integrates the specification of the level of effort (weak, medium, strong) involved in the production of the sequence. Depending on this effort level, different probability distributions are associated to the achievement of the task. The inferred command patterns will be assessed on the biomechanical model of the tongue with a special focus on how increasing the speaking rate affects the articulatory and acoustic accuracy in phoneme production.

The paper is divided into four sections. Section 2 summarizes the main hypotheses involved in the formulation of the Bayesian model, with an emphasis on the additional constraint that is to be included. From these ingredients the extension of the Bayesian model is introduced in Section 3. Section 4 presents the results of the inference and the assessment of the inferred command patterns with the biomechanical tongue model. Finally, the strengths of the proposed modeling framework are discussed with respect to its capacity to deal with different kinds of physical constraint applied to speech production.

# 2. Methodology

## 2.1. Basic control of the biomechanical model

This section summarizes all hypotheses considered in the model. The first half of these hypotheses are exactly the same as those used to define our recently published model [9]. The second half concerns the introduction of the additional total-level-of-force constraint.

As in GEPPETO, we are concerned with the selection of control variables for the production of sequences of phonemes. We only consider phonemes {/i/, /e/, /ɛ/, /a/, /oe/, /ɔ/, /k/} that do not require jaw or lip movements since the version of the biomechanical model that is used for simulations only includes an account of the tongue [17, 10]. Control variables correspond to the six control parameters, called $\lambda$ [18], used to pilot the biomechanical model. The first three formants of the acoustic signal are assumed to define the auditory space in which motor goals associated with phonemes are specified as convex target regions. Finally, the knowledge of the mapping between control variables and formant values is assumed to be stored in an internal model in the CNS [5]. In our study it is implemented by a radial basis function (RBF) network learned from more than $2.10^4$ simulations covering the whole motor space [19]. The internal model is considered to be "static" as it associates commands and outputs at targets.

This first set of hypotheses was used to specify in our previous model the selection of control variables based on the auditory characterization of phonemes only. However, different configurations of $\lambda$ control variables may result in the same tongue shape (and therefore the same acoustic signal), corresponding to an identical equilibrium configuration for different generated forces. The total level of force influences the capacity of the tongue to satisfy the speech requirements associated with increasing speaking rates. GEPPETO characterizes every articulatory configuration at the targets with a corresponding level of effort. These levels of effort are associated with the levels of total muscle force that are categorized in three levels, "Weak", "Medium" and "Strong" [16]. Since muscle force capacity is highly muscle dependent [20, 21] and since phonemes are differentiated by the patterns of recruited muscles [22, 23, 24, 25], for a given total level of effort, the actual muscle force involved depends on the considered phoneme. Thus, effort is not simply linked with the absolute amplitude of muscle force, but with the amplitude of the force relatively to the maximal capacity of the involved muscles to produce force. This is a way to not penalize phonemes, such as /i/ or /u/, that requires the activation of intrinsically strong muscles, such as the Genioglossus or the Syloglossus, as compared to phonemes, such as /a/ that is associated with a weaker muscle, the Hyoglossus. Here again, the knowledge of the control-to-force mapping is assumed to be stored in the CNS in a "static" internal model, which results from a learning process that generalizes the relation between motor commands and generated forces from a number of examples. It is implemented by a second RBF network.

## 2.2. Bayesian model

### 2.2.1. Model definition

**Variables** Variables correspond to the formal representation of all the relevant quantities selected for the description of the system. In the Bayesian framework, these quantities correspond to probabilistic variables. These variables are extracted from the control scheme presented in Section 2:

- $\boldsymbol{\Phi}$ is a discrete variable representing phonemes. It corresponds to the set of phonemes specified in Section 2.
- $\boldsymbol{S}$ is a continuous vector variable in the auditory space described as the 3-dimensional formant space: $S = (F_1, F_2, F_3)$.
- $\boldsymbol{M}$ is a 6-dimensional continuous vector variable representing the motor commands controlling the tongue: $M = (\lambda_1, \cdots, \lambda_6)$.
- $\boldsymbol{\nu}$ is a continuous scalar variable representing the amount of total muscle force generated.
- $\boldsymbol{N}$ is a discrete variable representing the level of effort generated at the target articulation of a phoneme. Its elements are the 3 levels of effort considered in GEPPETO.

**Decomposition** Our aim is to completely specify the joint probability distribution $P(M\ S\ \Phi\ \nu\ N)$. To this aim, the joint probability distribution is first decomposed following the standard chain rule of probabilities. Then, assumptions related to dependencies linking different variables are exploited, in order to simplify the obtained decomposition. The resulting decomposition is:

$$\begin{aligned} P(M\ S\ \Phi\ \nu\ N) \quad &= \quad P(M)\, P(S \mid M)\, P(\Phi \mid S) \\ &\qquad P(\nu \mid M)\, P(N \mid \Phi\ \nu). \qquad (1) \end{aligned}$$

The assumptions leading to the terms in the first line of Eq. (1) are the same as for our previous model [9]. Terms appearing in the second line of Eq. (1) correspond to the additional constraint included in the model. The simplifications leading to these terms are the assumptions that the total force level $\nu$ is uniquely specified by the motor variable $M$, and that the level of effort $N$ is completely characterized by the knowledge of $\Phi$ and $\nu$. It is worth noting that the global structure of the obtained decomposition results in the combination of the structure involved in our previous model with the structure relating muscle forces and related level of effort. Figure 1 illustrates the incorporation of these two structures into the current model.

**Parametric forms** Once defined the decomposition for the joint distribution $P(M\ S\ \Phi\ \nu\ N)$ (Eq. (1)) has been defined, it is necessary to specify the form taken by each of its factors.

$\boldsymbol{P(M)}$, the probability of occurrence of a motor pattern $M$, is specified by a uniform distribution, since, in the absence

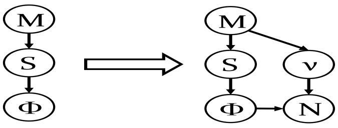

Figure 1: *Incorporation of the additional effort-constraint into our previous Bayesian model [9].*

of further knowledge, there is no reason to assume that some motor control patterns should be favored as compared to others.

$\boldsymbol{P(S \mid M)}$ and $\boldsymbol{P(\nu \mid M)}$ represent probabilistic accounts of the knowledge of the motor-to-auditory and motor-to-force mappings respectively (see Section 2). Since these mappings uniquely specify the auditory and force outputs corresponding to motor variable $M$, they are described by Dirac delta functions located on the outputs of the corresponding RBF networks. These outputs are denoted by $s^*(M)$ and $\nu^*(M)$ respectively.

$\boldsymbol{P(\Phi \mid S)}$ corresponds to the probability of associating phoneme $\Phi$ with the auditory input $S$. It is specified in the same way as in our previous model [9], by defining a sub-model that implements the inference of this categorization based on Gaussian descriptions of the $P(S \mid \Phi)$ distributions, describing the dispersion of auditory variable $S$ for each phoneme.

$\boldsymbol{P(N \mid \nu\ \Phi)}$ characterizes the link between the level of effort $N$ and the total level of force $\nu$ for a given phoneme $\Phi$. It can be seen, for each phoneme, as the categorization of the generated forces $\nu$ into one of the three level of efforts $N$. As for $P(\Phi \mid S)$, this term is specified by a sub-model that performs this categorization based on the probability distributions $P(\nu \mid N\ \Phi)$ of forces $\nu$ generated with respect to each level of effort $N$ and for each phoneme $\Phi$. $P(\nu \mid N\ \Phi)$ are assumed to be Gaussian, with location parameters (means) specified by the effort constraint implemented in GEPPETO, and variances scaled by parameter $\kappa_\nu$, which controls the strength of the constraint.

#### *2.2.2. Inference of motor commands $M$ producing phoneme $\Phi$ with a desired level of effort N*

Having specified the joint distribution $P(M\ S\ \Phi\ \nu\ N)$, we now formulate the question that is to be solved by the model. As the problem is to infer the control commands $M$ producing a desired phoneme $\Phi$ under a desired level of effort $N$, the approach consists in using the model to infer the probability distribution over $M$, conditioned on the specified values of $\Phi$ and $N$. The corresponding distribution, $P(M \mid \Phi\ N)$, is obtained by standard Bayesian inference and is given by:

$$P(M \mid \Phi\ N) \quad = \quad \frac{\sum_{S,\nu} P(M\ S\ \Phi\ \nu\ N)}{\sum_{M,S,\nu} P(M\ S\ \Phi\ \nu\ N)}. \qquad (2)$$

Since in Eq. (2) the denominator is constant for fixed $N$ and $\Phi$, we focus on its numerator and make the denominator implicit by using a proportionality symbol "$\propto$". Using the decomposition given by Eq. (1) and including the constant $P(M)$ term in the proportionality symbol, we obtain:

$$\begin{aligned} &P(M \mid \Phi\ N) \\ &\quad \propto \sum_{S,\nu} P(S \mid M)\ P(\Phi \mid S)\ P(\nu \mid M)\ P(N \mid \Phi\ \nu) \\ &\quad \propto P(\Phi \mid s^*(M))\ P(N \mid \Phi\ \nu^*(M)), \qquad (3) \end{aligned}$$

where the summation over $S$ and $\nu$ is reduced to the terms $\nu = \nu^*(M)$ and $S = s^*(M)$ for which $P(\nu \mid M)$ and $P(S \mid M)$ are not equal to zero. Eq. (3) completely defines the distribution from which motor variables $M$ can be selected in order to produce a specific phoneme with a desired level of effort.

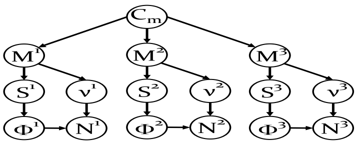

Figure 2: *Diagram of the sequence-planning model including the effort constraint.*

#### *2.2.3. Inference of motor commands producing a sequence of phonemes with a desired level of effort*

Until now we have only considered the selection of motor commands for the production of isolated phonemes. The interest of taking into account the level of effort corresponding to the planned motor variable is made clear in the generation of sequences of phonemes. We previously proposed a Bayesian model planning motor commands for the production of sequences of phonemes, which only considered the auditory constraint [9]. Based on the above methodology for the planning of a single phoneme, it is straightforward to include the additional effort constraint into the sequence planning model. Figure 4 represents the diagram corresponding to this extended sequence-planning model from which we can infer motor variables for the production of sequences of phonemes with different levels of effort. Variables are duplicated and indexed by their order in the intended sequence. The additional variable $C_m$ implements a motor constraint that imposes the proximity of planned motor commands resulting in anticipatory coarticulation effects [9].

## 3. Results

### 3.1. Inferred control patterns

The decision policy underlying the selection of a set of control variables for the production of phoneme $\Phi$ with effort level $N$ consists in random sampling based on $P(M \mid \Phi\ N)$. This sampling is approximated by the Metropolis-Hasting algorithm that performs a Markov Chain Monte Carlo *(MCMC)* random walk that converges to the desired distribution. Simulations were performed with Matlab's "mhsample" function with 20 chains of $2.10^4$ samples each and a burning period of $10^3$ samples. Tests revealed that further increasing the number of samples had no influence on the global shape of the obtained distributions.

In order to validate the performance of the model it is necessary to evaluate whether control samples $M$ obtained from this sampling policy effectively result in intended auditory and force values that satisfy the constraints. Figure 2 gives an example of the results. It represents the histograms of the first three intended formant values corresponding to the samples $M$ obtained for phoneme /i/ from $P(M \mid \Phi\ N)$ at the three effort levels. These intended formant values where computed from the RBF network used for the internal models described in Section 2. It can be seen that the formants correctly distribute inside the target regions for all the categories of effort. This remains true for all other phonemes. Similarly, Figure 3 represents the histograms of the total intended muscle force obtained from the same set of samples $M$ used in Figure 2. Again, these levels

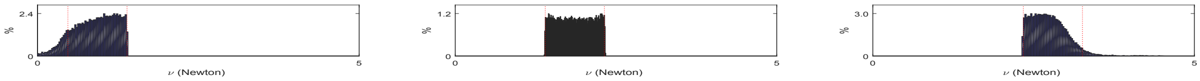

Figure 3: *Histograms of the total level of force corresponding to samples $M$ obtained by the the probability distribution $P(M|\Phi N)$ for phoneme /i/ and effort levels N="Weak"* ***(Left)****, N="Medium"* ***(Middle)*** *and N="Strong"* ***(Right)****. The vertical dotted lines indicates the borders of the regions characterizing the effort constraint of GEPPETO.*

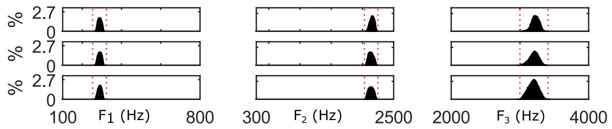

Figure 4: *Histograms of auditory consequences of samples $M$ sampled from the probability distribution $P(M \mid \Phi\ N)$ for phoneme /i/ and effort levels N="Weak"* ***(Bottom)****, N="Medium"* ***(Middle)*** *and N="Strong"* ***(Top)****. The vertical dotted lines indicate borders of the auditory regions characterizing the auditory constraint for this phoneme.*

of intended total force were obtained from the RBF network. It can be seen that the forces correctly distribute according to the corresponding regions that define each level of effort. In summary, the Bayesian model correctly infers control samples $M$ that jointly satisfy both the auditory and the effort constraints characterizing the speech planning task.

### 3.2. Sequence planning at various levels of effort

We consider for illustration the sequence /aiɛ/ planned with the N="Weak" and N="Strong" levels of effort. The set of control variables having the highest inferred probabilities are selected and the resulting tongue trajectories are generated by the biomechanical model presented in Section 2. Two speaking rates are implemented for each set of control variables by specifying a slow and fast transition rate between the motor commands of the first and second phoneme in the sequence, together with a long and a short duration of the second phoneme. Results are illustrated as formant trajectories in auditory space shown in Figure5. It can be seen that for a fast speaking rate, motor commands planned with the N="Weak" level of effort result in a formant trajectory that misses the auditory target region for the middle phoneme /i/ (Top left image). This is not the case for control variables planned with the N="Strong" levels of effort (Top right image). In a slow speaking rate however, these same two sets of control variables result in formant trajectories that both reach auditory target regions (Bottom images).

## 4. Conclusion

We have presented an extension of a recently published Bayesian model of speech motor control that plans control variables for the production of phoneme sequences [9]. This extension shows how an additional constraint, related to the level of total muscle force generated in the production of phonemes, can be included in a modular manner into the model. We have shown that the model correctly infers motor commands that jointly satisfy both constraints characterizing the speech task: intended auditory outputs lie inside the desired phonemic regions and total level of muscle force satisfies the desired effort interval. The model was constructed as a reformulation of GEPPETO, an existing optimal control model that rests on the minimization of a cost function. We have shown in a previous paper [9] that the Bayesian model includes the optimal control approach of GEPPETO as the particular case of finding motor commands with the maximum posterior probability. In this sense the Bayesian framework offers a more general approach that has valuable properties. For instance, the Bayesian approach results in a distribution of solutions while the optimal control approach gives a unique point-like value. These distributions allow us to better explore the space of solutions and study the possibility that patterns of variability may be attributed to the distribution of possible equivalent motor solutions.

We see that the proposed Bayesian Programming approach offers an interesting framework for the integration of multiple constraints. The flexibility and coherence of the Bayesian methodology allows us to further include additional constraints considered in speech motor control. In this context, current work is focused in the combination of constraints related to proprioceptive characterization of phonemes along with the present auditory representation.

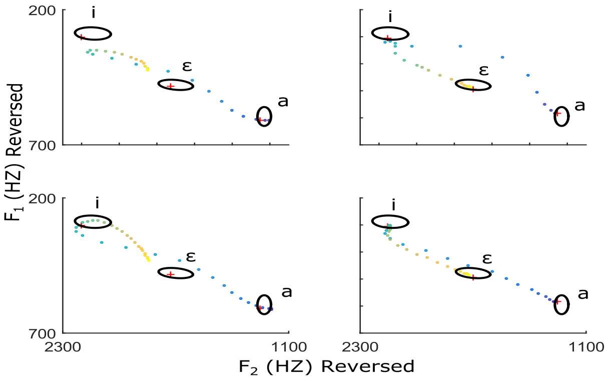

Figure 5: *Formant trajectories corresponding to the motor commands planned for the sequence /aiɛ/ with the N="Weak" (**Left**) and N="Strong" (**Right**) effort levels, produced with fast (**Top**) and slow (**Bottom**) speech rates. For a fast speaking rate, the trajectory planned with the N="Weak" effort level misses the auditory target for the middle phoneme /i/ (Top left image), but not for the N="Strong" effort level. Both commands produce trajectories that reach the auditory targets for slow speech rate (Bottom images). Red crosses indicate the intended auditory output of the motor commands.*

## 5. Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Communitys Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 339152, "Speech Unit(e)s", PI: Jean-Luc Schwartz).

# 6. References

[1] A. Stuart, J. Kalinowski, M. P. Rastatter, and K. Lynch, "Effect of delayed auditory feedback on normal speakers at two speech rates," *The Journal of the Acoustical Society of America*, vol. 111, no. 5, pp. 2237–2241, 2002.

[2] B. S. Atal, J. Chang, M. V. Mathews, and J. W. Tukey, "Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique," *The Journal of the Acoustical Society of America*, vol. 63, no. 5, pp. 1535–1555, 1978.

[3] E. Farnetani and D. Recasens, "Coarticulation and connected speech processes," *The handbook of phonetic sciences*, pp. 371–404, 1997.

[4] D. M. Wolpert, R. C. Miall, and M. Kawato, "Internal models in the cerebellum," *Trends in cognitive sciences*, vol. 2, no. 9, pp. 338–347, 1998.

[5] F. H. Guenther, M. Hampson, and D. Johnson, "A theoretical investigation of reference frames for the planning of speech movements." *Psychological review*, vol. 105, no. 4, p. 611, 1998.

[6] F. H. Guenther, "Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production." *Psychological review*, vol. 102, no. 3, pp. 594–621, 1995.

[7] P. Perrier and L. Ma, "Speech planning for V1CV2 sequences: Influence of the planned sequence," in *Proceedings of the 8th International Seminar on Speech Production*, ISSP-2008, Ed., Universit de Strasbourg, France, 2008, pp. 69–72.

[8] S. Perkell, J. and L. Nelson, W., "Variability in production of the vowels /i/ and /a/," *Journal of the Acoustical Society of America*, vol. 77, pp. 1889–1895, 1985.

[9] J.-F. Patri, J. Diard, and P. Perrier, "Optimal speech motor control and token-to-token variability: a Bayesian modeling approach," *Biological Cybernetics*, vol. 109, no. 6, pp. 611–626, 2015. [Online]. Available: http://dx.doi.org/10.1007/s00422-015-0664-4

[10] P. Perrier, Y. Payan, M. Zandipour, and J. Perkell, "Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study," *The Journal of the Acoustical Society of America*, vol. 114, no. 3, pp. 1582–1599, 2003.

[11] O. Lebeltel, P. Bessière, J. Diard, and E. Mazer, "Bayesian robot programming," *Autonomous Robots*, vol. 16, no. 1, pp. 49–79, 2004.

[12] P. Bessière, E. Mazer, J. M. Ahuactzin, and K. Mekhnacha, *Bayesian Programming*. Boca Raton, Florida: CRC Press, 2013.

[13] P. Bessière, C. Laugier, and R. Siegwart, Eds., *Probabilistic Reasoning and Decision Making in Sensory-Motor Systems*, ser. Springer Tracts in Advanced Robotics. Berlin: Springer-Verlag, 2008, vol. 46.

[14] F. Colas, J. Diard, and P. Bessière, "Common Bayesian models for common cognitive issues," *Acta Biotheoretica*, vol. 58, no. 2-3, pp. 191–216, 2010.

[15] P. Perrier, L. Ma, and Y. Payan, "Modeling the production of VCV sequences via the inversion of a biomechanical model of the tongue," in *Proceedings of Interspeech 2005*, Lisbon, Portugal, 2006, pp. 1041–1044.

[16] R. Winkler, L. Ma, and P. Perrier, "A model of optimal speech production planning integrating dynamical constraints to achieve appropriate articulatory timing," in *Proceedings of the 9th International Seminar on Speech Production*, ISSP2011, Ed., vol. Abstracts, Montral Canada, 2011, pp. 235–236.

[17] Y. Payan and P. Perrier, "Synthesis of VV sequences with a 2D biomechanical tongue model controlled by the equilibrium point hypothesis," *Speech communication*, vol. 22, no. 2, pp. 185–205, 1997.

[18] A. G. Feldman, "Once more on the equilibrium-point hypothesis ($\lambda$ model) for motor control," *Journal of motor behavior*, vol. 18, no. 1, pp. 17–54, 1986.

[19] T. Poggio and F. Girosi, "A theory of networks for approximation and learning," Artificial Intelligence Laboratory & Center for Biological Information Processing, MIT, Cambridge, MA, USA, Tech. Rep., 1989.

[20] G. Pruim, H. De Jongh, and J. Ten Bosch, "Forces acting on the mandible during bilateral static bite at different bite force levels," *Journal of biomechanics*, vol. 13, no. 9, pp. 755–763, 1980.

[21] R. A. Brand, D. R. Pedersen, and J. A. Friederich, "The sensitivity of muscle force predictions to changes in physiologic cross-sectional area," *Journal of biomechanics*, vol. 19, no. 8, pp. 589–596, 1986.

[22] K. Honda, "Organization of tongue articulation for vowels," *Journal of Phonetics*, vol. 24, pp. 39–52, 1996.

[23] T. Baer, P. Alfonso, and K. Honda, "Electromyography of the tongue muscles during vowels in /ɔpʊp/ environment," *Ann. Bull. RILP*, vol. 22, pp. 7–19, 1988.

[24] S. Buchaillard, P. Perrier, and Y. Payan, "A biomechanical model of cardinal vowel production: muscle activations and the impact of gravity on tongue positioning," *J Acoust Soc Am.*, vol. 126, no. 4, pp. 2033–51, 2009.

[25] S. Waltl and P. Hoole, "An EMG study of the German vowel system." in *Proceedings of the 8th International Seminar on Speech Production (ISSP)*, R. Sock, S. Fuchs, and Y. Laprie, Eds., 2008, pp. 445–448.