



**HAL**  
open science

## About speech motor control complexity

Pascal Perrier

► **To cite this version:**

Pascal Perrier. About speech motor control complexity. Jonathan Harrington & Marija Tabain. Speech Production: Models, Phonetic Processes, and Techniques, Psychology Press: New-York, USA, pp.13-26, 2006, Macquarie Monographs in Cognitive Science. hal-00371045

**HAL Id: hal-00371045**

**<https://hal.science/hal-00371045v1>**

Submitted on 26 Mar 2009

**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.

## **ABOUT SPEECH MOTOR CONTROL COMPLEXITY**

Pascal Perrier

Institut de la Communication Parlée, UMR CNRS 5009

Institut National Polytechnique de Grenoble & Université Stendhal Grenoble, France

### **ABSTRACT**

A key issue in research about speech motor control is the one of the level of complexity that is required for the internal models: have these models to account accurately for all physical properties of the speech motor system, including the complex tongue-jaw biomechanics? Or would more simple internal representations be sufficient, which model only the static characteristics of the peripheral speech apparatus or which give a rough account of articulatory dynamics? On the basis of experimental and modeling studies of speech movements and human limb movements published in the literature, the adequacy of simplified internal representations for speech motor control is analyzed.

## 1. INTRODUCTION

In the past two decades, the analysis of speech motor control and its modeling were basically inspired by investigations and models of other skilled human movements, such as reaching, pointing or grasping. This research approach has been proved to be efficient and it allowed the emergence of a number of speech control models which served as theoretical backgrounds for numerous studies (see among others Saltzman, 1986; Saltzman & Munhall, 1989; Guenther, 1995; Ostry *et al.*, 1996; Perkell *et al.*, 1997; Laboissière *et al.*, 1996; Perrier *et al.*, 1996; Perkell *et al.*, 2000). Consequently, the issues raised today in the domain of motor control research are very important for future studies of speech motor control, and addressing them is a prerequisite for the improvement of the current models. Among these issues, two questions are of particular interest: (1) How complex are the acquired internal representations of the motor system, built up during the acquisition of the motor task, and used to plan and achieve the movement? (2) What role does low-level, short-latency feedback play in achieving an accurate and stable movement control? In this paper, contributions to these issues are proposed, while taking into account some important peculiarities of the speech production task.

## 2. SPEECH PRODUCTION: A COMPLEX MOTOR TASK.

Compared to many other human motor tasks classically studied in motor control research, speech production has a number of peculiarities that make it particularly complex. Some findings suggesting such a complexity were already frequently discussed in the literature:

- (1) Because of its semiotic nature, the goal of speech production is actually defined in an abstract domain. Hence, its physical characterization is not straightforward, and this has two consequences. First there is no unique physical correlate for a given elementary speech sound and a large variability of patterns can be observed in the neuro-

physiological, articulatory and acoustic domains (see Perkell & Klatt, 1986). Second the issue of the physical space in which the motor task is planned becomes particularly complex, since the distal space can be defined either by articulatory positions, or by spectral properties of the speech signal, or by perceptual characteristics of this signal (see Browman & Goldstein, 1990; Stevens, 1989; Savariaux *et al.*, 1995; Guenther *et al.*, 1998, Tremblay *et al.*, 2003), or a multimodal space associating oro-sensory, auditory and even visual characterizations.

(2)Speech production has a large number of degrees of freedom that confer a many-to-one characteristic on the relationships between motor commands, articulatory positions and acoustic or auditory properties. This characteristic, together with the above mentioned intrinsic variability of the physical correlates of the production of a given sound, has the consequence that a large set of motor equivalence strategies can be used to implement a range of coarticulation strategies or to deal with artificial perturbations (such as a pipe or food in the mouth), or pathological perturbations (such as tongue or mandible surgery) or peripheral perturbations (Perkell *et al.*, 1993; McFarland *et al.*, 1996, Guenther *et al.*, 1999, Savariaux *et al.*, 1999). These multiple strategies obviously contribute to the complexity of speech motor control.

(3)Compared to other skilled human motor activities, speech movements in normal conditions can be very short, since vowels have a mean duration of approximately 80 ms and consonants have mean durations around 40ms (O'Shaughnessy, 1981). These characteristics seem to exclude any potential on-line contribution of long-latency oro-sensory feedback that would be processed by the cortex, and to limit the role of auditory feedback to a suprasegmental level and to an *a posteriori* monitoring used to correct segmental aspects of speech after it was produced (Perkell *et al.*, 1997). The absence of on-line use of auditory feedback to control speech at a segmental level is

well supported by experimental work showing that speakers can produce intelligible speech even after hearing loss (Lane & Wozniak, 1991; Manzella *et al.*, 1994).

Compatibly, work on stutterers and normal speakers shows that delaying auditory feedback in the range of 50 - 200ms affects prosodic (speaking rate, fluency, rhythm, intonation and stress) rather than segmental features (Hargrave *et al.*, 1994; Stager & Ludlow, 1993). At the same time, speech gestures have to be accurate enough in order to ensure that the associated acoustic signal can be correctly perceived by a listener. How accuracy can be obtained without the use of long-latency feedback that would be processed by the cortex, is a key issue for speech production research, but not for other human motor tasks except for eye-saccades.

To deal with the complexity and the multimodality of speech task representations, with the numerous motor or auditory equivalence strategies, and with the accuracy requirements in the absence of long-latency feedback going through the cortex, the large majority of speech motor control models published in the literature assume the existence of *internal representations* of the speech apparatus, called *internal models* (Jordan, 1990; Kawato *et al.*, 1990; Laboissière *et al.*, 1991; Hirayama *et al.*, 1992; Jordan & Rumelhart, 1992; Guenther, 1995; Perkell *et al.*, 1997; Perkell *et al.*, 2000; Perrier *et al.*, 2004).

### **3. INTERNAL MODELS AND SPEECH PRODUCTION CONTROL**

#### **3.1. A useful concept to deal with the control of complex motor tasks**

The *internal model* concept was proposed in the research domain to deal with the non-biunivocity between motor commands and position of the final effector, and with the delays associated with the processing of long-latency feedback. The basic hypothesis is that copies of the motor system, or of subsets of it, could be learned in the brain during the acquisition of

the motor skill (in our case in the speech learning phase). Once they are learned, these models could be used to estimate predictions of the consequences of motor command changes on the trajectory of the final effector. According to different existing models of human motor control published in the literature, the role of *internal models* in the execution of the motor task could take different forms.

A first hypothesis suggests that internal models' predictions could be the basis of task planning strategies aiming at ensuring that the final effector moves along a specific desired trajectory in the task space. This is the so-called *desired trajectory hypothesis* (Kawato, 1999). Since human motor systems have usually an excess of degrees of freedom, many different motor command sequences are likely to allow the achievement of the specified trajectory. The basic idea of the *desired trajectory hypothesis* is that the Central Nervous System would use internal models prior to the execution of the movement, in order to select an optimal one from all possible motor command sequences that would both generate the required trajectory and minimize a motor criterion, classically related to the concept of effort. This kind of internal model is classical called *inverse internal model*, since it permits to go from the desired output to the motor commands.

For target oriented movements, an alternative hypothesis suggests that the Central Nervous System would use internal models as *direct* models during the planning to optimize the positioning accuracy of the final effector at the target in presence of neural noise (Harris & Wolpert, 1998). From this perspective, internal models would also be used prior to the execution of movement, but the trajectory of the final effector toward the target would be a consequence of the planning strategy rather than the specification of the task itself. Another, more recent use of direct internal models was proposed by Todorov & Jordan (2002) in the

framework of their *optimal control feedback* model. According to this model the motor control strategy would not consist in specifying a desired trajectory in the task space and in selecting the appropriate commands for the motor system to follow this trajectory. It would instead make use of feedback information during the execution of movement to selectively modify motor commands in an optimal way, when deviations in the task space occur that would endanger the achievement of the task goal. The proposed use of feedback would require long-delay loops that are known to generate instabilities in closed-loops servo-mechanisms and that would consequently induce inadequate corrections from the controller. To overcome these problems, Torodov & Jordan (2002) have proposed using the outputs of the internal models as afferent signals.

Thus, from the perspective of motor control models such as the *desired trajectory hypothesis*, or Harris & Wolpert's (1996) proposal, or the *optimal control feedback* model, internal models are very powerful tools for dealing both with the many-to-one nature of the relations between motor commands and vocal tract configurations or spectral characteristics of the acoustic signal, as well as with the long-latency of feedback processed by the cortex, and with the possible multimodality of the speech task representations. And there is considerable experimental evidence to suggest that internal models are set up and used in the control of speech production control.

### **3.2. Experimental findings support the existence of internal models for speech production control.**

Most of the data supporting the internal model hypothesis for speech production were obtained in perturbation experiments where a noticeable change was artificially introduced either in the speech production apparatus or in speech perception feedback.

Savariaux *et al.* (1995, 1999) perturbed the production of the French [u] by introducing a 2 cm diameter tube between the lips of their subjects. None of the subjects could compensate for the perturbation without a training phase where they could listen to their acoustic speech production. However, immediately after the insertion of the tube, the large majority of the subjects showed articulatory changes that tended to partly compensate for the perturbation by reducing the acoustical and perceptual gaps between their normal and their perturbed [u]'s. In addition, none of them provided changes in a direction that would have been incompatible with a compensation. These observations support the idea that prior to their first production with the lip tube, speakers would use an internal representation to indicate in which direction articulatory changes should be provided to compensate for the perturbation.

Perkell *et al.* (2000) reported results about the spectral characteristics of speech produced by cochlear-implanted patients according to whether the cochlear implant was turned on or off. It was shown that the degradation of speech without the auditory feedback was very slow, and that it was perceptually noticeable only in alveolar fricatives that require accurate tongue positioning. This supports the hypothesis of the existence of internal representations of the relations between motor commands and spectral characteristics of the signal which would allow the subjects to predict their speech signals in spite of the absence of auditory feedback. Note that Perkell *et al.* (2000) suggested that auditory feedback would be necessary to maintain these representations.

Houde & Jordan (1998) set up an original experimental protocol where subjects were asked to produce whispered steady state vowels while listening via headphones to acoustic feedback in that had been distorted by slightly shifting the formants.. Subjects were asked to



compensate for the auditory perturbation so that they no longer perceived any perturbation. All subjects could compensate after training, and an after-effect was also systematically observed once the auditory feedback perturbation was removed. These findings support the idea that due to the distortion of the auditory feedback, subjects learned new relations between their articulatory commands and the associated spectral characteristics of the speech signal, i.e. a new internal model, and that they relied after training on this new representation to control their speech production.

The three experimental paradigms described above could in fact be explained by static internal models describing the link between steady-state articulatory configurations of the vocal tract and spectral patterns. However, a recent experiment carried out by Tremblay *et al.* (2003) suggests that a more complex internal representation could be required for speech motor control. Subjects were asked to produce vowel-to-vowel transitions while a velocity-dependent perturbing force was applied to their jaw. Thus, perturbations modified the dynamical state of the speech apparatus and not only the static relations between articulation and perception. After training, the subjects were able to compensate for this perturbation in such an accurate way that the jaw moved along the same articulatory path as under normal conditions. In addition, an after-effect was observed once the perturbation was removed. This suggests that the subjects were able to learn a new internal model taking into account the modified dynamical properties of their jaw, and to use it in their speech production. Interestingly, the learning was not observed, when the task only consisted of moving the jaw from a closed to an open position without speaking.

### 3.3. What kind of internal models for speech production control?

As shown in the preceding section, different sources of evidence show that the control of speech production could incorporate and make use of internal models. These models could be purely static, as suggested in the first three examples, or could require a more complex description of the physics as shown by Tremblay *et al.* (2003). This raises the question of the level of complexity of the internal models of speech production. This is a key issue, and not only for speech production research, but for motor control research in general as illustrated by the controversy between Kawato and colleagues and Ostry and colleagues in their respective attempts to explain arm trajectories in pointing tasks (Gomi & Kawato, 1996; Gribble *et al.*, 1998). Does the motor control system need internal models that describe the motor system in its whole complexity, including dynamical and biomechanical aspects (Kawato *et al.*'s hypothesis), or could a rough description be sufficient (Ostry *et al.*'s assumption)?

In the case of speech motor control this question becomes particularly acute, since the physical properties of speech apparatus are particularly complex:

- (1) The tongue is a hydrostat, i.e. a soft body with highly complex, typically non linear, physical properties and deformation capabilities (see Gerard *et al.*, in this issue, for more details). This articulator is carried by the mandible, a hard body with dynamical properties that differ strongly from the ones of the tongue. Hence, the mechanical coupling between these two articulators cannot be described by simple mechanical laws.
- (2) The tongue moves in a narrow space delimited by the pharyngeal walls, the palate, the teeth, and the cheeks. Consequently, the tongue is often in contact with external structures. The amount of contacts as well as their location vary during speech

production, and this induces over time changes in the physical constraints on tongue movements (damping, limitation of movement amplitude in a given direction...).

Thus, the mechanical and dynamical properties of the peripheral speech apparatus are not only, very complex. They also vary during the movements, since they depend on the positions of the articulators in relation to each other and on the strain of the soft tissues. Consequently, building up accurate internal models of this apparatus means, first, accounting for a strong mechanical and dynamical complexity, and, second, making the models sufficiently flexible in such a way that they can adapt based on information about the internal structure of the soft tissues and about their relations to solid structures. Recent publications have shown evidence for the very good capacity of the Central Nervous System to learn adaptable internal models of complex and time variable dynamics to control accurately the arm or the limb (Flanagan & Lolley, 2001; Dingwell *et al.*, 2002; Waincott *et al.*, 2005). However, in the case of speech articulators, things seem to be even more complex, since, in this case, to be accurate, internal models would have to adapt very quickly and on the basis of short latency proprioceptive afferent signals. Hence, it seems reasonable to assume that internal models of speech production are inaccurate functional descriptions of the speech motor system. From this perspective, different levels of complexity, from a pure static model to a sophisticated dynamical one, can be considered. In this context, an important challenge consists in collecting evidence supporting one hypothesis or the other. In the following sections, results will be presented that support the hypothesis that simplified internal representations could be efficiently used for speech motor control.

#### **4. SIMPLIFIED INTERNAL MODELS COULD BE SUITED TO THE CONTROL OF SPEECH PRODUCTION**

##### **4.1. Complex trajectories do not necessarily require complex internal models**

The hypothesis that planning of skilled movement sequences would be based on cost minimization principles using dynamic variables such as jerk (Hogan, 1984) or torque change (Kawato *et al.* 1990) is often proposed to explain specific kinematic characteristics of human movements such as the curved path observed for the finger in pointing tasks (Gomi & Kawato, 1996) or for the tongue in speech tasks (Löfqvist & Gracco, 2002). However, our recent work carried out with a 2D biomechanical model of the tongue does not support this point of view (Perrier *et al.*, 2003).

We simulated tongue movements in Vowel - Velar Stop – Vowel sequences with a 2D biomechanical model of the tongue, which includes the main muscles responsible for shaping and moving the tongue in the midsagittal plane (posterior and anterior parts of the genioglossus, styloglossus, hyoglossus, inferior and superior longitudinalis and verticalis). Tissue elastic properties are accounted for in finite-element (FE) modeling. Muscles are modeled as general force generators which (1) act on anatomically specified sets of nodes of the FE structure, and (2) modify the stiffness of specific elements of the model to account for muscle insertions into tongue tissues. The model was controlled via a target-based model using Feldman's Equilibrium Point Hypothesis (Feldman, 1986). In our simulations, each sound was specified within a given speech sequence by a set of motor commands that determine a static equilibrium position of the tongue. The transition from one sound to the next was produced by shifting the motor commands at a constant rate of shift. Hence, the temporal paths of the motor commands were strictly dependent on the time of occurrence of discrete sets of motor commands specifying the successive equilibrium positions associated

with the phonemes of the sequence. These paths were not determined by any requirement on the articulatory trajectory itself. Under these conditions, the simulated tongue trajectories replicated nicely the well-known curved aspect observed in real speakers (see Perrier *et al.*, 2003, for details). Consequently, there is no need to hypothesize any cost minimization based on dynamical variables, nor on dynamical internal models, to explain the curvature of the articulatory paths: it could be simply due to biomechanical properties of the tongue, i.e. passive tongue elasticity, muscle arrangements within the tongue, and force generation mechanisms. These results do not disprove the hypothesis of complex internal representations, since jerk or torque minimization principles could act in combination with biomechanical factors. However, our simulations suggest that this hypothesis is not a necessity to explain curved articulatory trajectories in speech.

We do not want to discard the possibility that internal representations could include some knowledge about the dynamics of the speech apparatus. It is well-known that time control is a major issue in speech production, since phone durations can carry a significant part of the linguistic information. Since time is strongly dependent on the dynamics of the motor system, it is very unlikely that the dynamics could be excluded from the internal representations that the brain builds up to control the movements. Now, the question is whether coarsely specified dynamical internal models could be sufficient to reach the time and positioning accuracy required for speech production.

#### **4.2. Inaccurate dynamical representations are compatible with movement accuracy**

Inaccurate internal models of the speech production apparatus will induce errors in the estimation of motor commands from the specification of the desired gestures. Consequently, they will induce gesture and acoustic inaccuracy, unless motor commands can be corrected

during the execution of movement. To deal with this problem, Kawato and colleagues (for example Kawato *et al.*; 1987; Gomi & Kawato, 1992; Kawato, 1999) proposed to use a combination of a *feedforward controller*, inferring motor commands from the desired output by using inverse dynamics, and of a *feedback controller*; taken into account the output errors of the motor system to correct *on line* the motor commands. This is definitely a good framework to further think about how to deal with internal model inaccuracy. Still, there are two important problems remaining :

- (1) How to deal with feedback delays?
- (2) Is such a feedback controller compatible with a sparse description of the desired output, in which only target points are specified?

As explained above, Todorov & Jordan (2002) proposed a solution to the first problem, by suggesting using the outputs of the internal model as feedback signals. However, it is not clear how their control model would behave, if the feedback signals taken into account are inaccurate due to internal model inaccuracy. An alternative, very interesting, contribution to this issue was proposed by Hanneton *et al.* (1997), who analyzed hand movements in visuo-manual tracking tasks. Instead of providing a gradual correction of the motor commands based on the continuous comparison between the actual position of the motor system and a desired trajectory , their feedback controller generates from time to time a very strong change to the motor commands in order to push the motor system back to right track.

Hanneton et al.'s subjects were seated in front of a computer screen on which a point moved continuously along a random trajectory. By moving a joystick commanding a simulated second order mechanical system, the subjects had to drive another point on the screen, so that it followed as accurately as possible the first point. Hanneton *et al.* (1997) observed that hand

movements were made of a number of small successive sub-movements associated with high velocity peaks and separated from each other by short periods in which velocity was small. Spectral analysis of these trajectories showed that the dynamical system controlled by the subjects and consisting of the joystick and the simulated mechanical system had non linear dynamical properties.

The authors could reproduce the observed movement patterns with a control model based on the processing of an intermediate *sliding variable* (Slotine & Li, 1991), while assuming linear characteristics for the dynamical system controlled by the subjects.

*“A sliding variable is a specific combination of the instantaneous error and its successive time derivatives. [...] Sliding variables can be viewed either as a prediction about the error or as a criterion to be minimized by an algorithm adapted to the dynamics of the system being controlled. [...] In the dynamic case, [...] control mechanisms are built based on arrangements of feedforward and feedback components: the feedforward part provides an estimate of the inverse dynamics of the controlled system while the feedback part compensates for uncertainties about this estimate.”* (Hanneton *et al.*, 1997, p. 382).

Thus, these authors proposed that the sub-movements associated with high velocity peaks could be produced in a ballistic-like way by the Central Nervous System when the absolute value of the sliding variable exceeds a constant threshold. Once triggered, these parts of the movements would be independent of any feedback (*open-loop* movements). The role of these rapid open-loop movements would be to bring rapidly the sliding variable value back below the threshold. The other parts of the movement, those having a small velocity, would

correspond to the classical *closed-loop* response of the dynamical system to the motor commands.

This approach is very appealing, because it makes possible to reach gesture accuracy in spite of internal model inaccuracy, and because it seems to be compatible with a target-based model of control, since a sparse description of the output based on discrete target sequences would not allow a continuous feedback correction. Hanneton et al. (1987) have demonstrated that system stability is possible, even if feedback corrections to the inputs are not continuous and strong. However, it should be noted that Hanneton et al.'s (1997) experiment involved long-latency visual feedback, since according to their model, the delay time of this feedback control would be 290 ms. This would be obviously far too long for speech production.

However, the corrective movements proposed by Hanneton et al. (1997) do not seem to require a cortical processing of the feedback, since they are ballistic and fully determined by the initial impulse. Hence, it seems possible that this model could be applied to speech by taking into account low-level short latency feedback, such as oro-sensory afferences. Further work is obviously needed to test this hypothesis for speech production.

## **5. CONCLUSION**

In this paper, we have provided various types of evidence in support of the hypothesis that internal models of the speech production apparatus, from the motor commands to the articulatory patterns or to the spectral characteristics of the acoustic signal, could exist and be used in the planning and/or the execution of speech production. We have also shown that complex articulatory patterns would not necessarily require complex internal models, since a part of the measured complexity could arise from the physical characteristics of the speech production system.



Nevertheless, the necessity for speakers to control temporal events in speech production with accuracy raises the question of the level of complexity that is required for the internal models to take into account dynamics. The sliding variable hypothesis offers a robust framework to think about inaccurate dynamical descriptions of the speech production system, without the need to specify the complete desired articulatory paths. However, a role for short-latency feedback needs to be considered given that in speech production movement durations are shorter than long-latency feedback delays.

Very sensitive mechanoreceptors have been found in the human tongue (Trulsson & Essick, 1997). Neural linkages between the upper lip and the jaw have also been hypothesized (Gracco & Abbs, 1985), in order to explain compensatory movements of the upper lip in case of a sudden perturbation applied to the jaw (note that Gomi *et al.*, 2002, have suggested that, in addition to the neural linkage, a dynamical coupling between lips and jaw could contribute to the compensatory movement). Hence, short latency reflexes exist in the neural circuitry of the speech articulators, and they seem to contribute to the stability of the speech output. Hence, it is worthwhile looking further for general neuro-physiological evidence that short latencies reflexes could be used in combination with a sparse description of the desired articulatory paths and simplified internal representations to produce accurate speech movements.

## 6. REFERENCES

- Browman, C.P. & Goldstein, L.M. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, 18, 299-320.
- Dingwell, J.B., Mah, C.D. & Mussa-Ivaldi, F.A. (2002) Manipulating objects with internal degrees of freedom: evidence for model-based control. *Journal of Neurophysiology*, 88: 222-235
- Feldman, A.G. (1986). Once more on the Equilibrium-Point Hypothesis ( $\lambda$  model) for motor control. *Journal of Motor Behavior*, 18, 17-54.
- Flanagan, J.R. & Lolley, S. (2001). The inertial anisotropy of the arm is accurately predicted during movement planning. *Journal of Neuroscience*, 21(4), 1361-1369.
- Gerard, J.M., Perrier, P. & Payan, Y. (2005). 3D biomechanical tongue modeling to study speech production. *In this issue*
- Gracco, V.L. & Abbs, J.H. (1985). Dynamic control of the perioral system during speech: kinematic analyses of autogenetic and nonautogenetic sensorimotor processes. *Journal of Neurophysiology*, 54, 418-432.
- Gomi, H., & Kawato, M. (1992). Adaptive feedback control models of the vestibulocerebellum and spinocerebellum. *Biological Cybernetics*, 68(2), 105-114.
- Gomi, H. & Kawato, M. (1996). Equilibrium-point control hypothesis examined by measured arm stiffness during multijoint movement. *Science*, 272,117-120.
- Gomi, H., Honda, M., Ito, T. & Murano E.M. (2002). Compensatory articulation during the bilabial fricative production by regulating muscle stiffness. *Journal of Phonetics*, 30, 261-279.
- Gribble, P.L., Ostry, D.J., Sanguineti, V. & Laboissière, R. (1998). Are complex control signals required for human arm movement? *Journal of Neurophysiology*, 79, 1409-1424.

- Guenther, F. H. (1995). Speech sound acquisition, coarticulation and rate effects in a neural network model of speech production. *Psychological Review*, *102*, 594–62.
- Guenther, F. H., Hampson, M. & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, *105*, 611–633.
- Guenther, F. H., Espy-Wilson, C.Y., Boyce, S.E., Matthies, M.L., Zandipour, M. & Perkell, J.S. (1999). Articulatory tradeoffs reduce acoustic variability during American English /t/ production. *Journal of the Acoustical Society of America*, *105*, 2854–2865.
- Hanneton, S., Berthoz, A., Droulez, J. & Slotine J.J.E. (1997). Does the brain use sliding variables for the control of movements? *Biological Cybernetics*, *77*, 381-393.
- Hargrave, S., Kalinowski, J., Stuart, A., Armson, J., & Jones, K. (1994). Effect of frequency-altered feedback on stuttering frequency at normal and fast speech rate. *Journal of Speech and Hearing Research*, *37*, 1313-1319.
- Harris, C.M., & Wolpert, D.M. (1998). Signal-dependent noise determines motor planning. *Nature*, *394*, 780-784.
- Hirayama, M., Vatikiotis-Bateson, E., Kawato, M. & Jordan, M.I. (1992). Forward dynamics modeling of speech motor control using physiological data. In R.P. Lippman, J.E. Moody & D.S. Touretsky (Eds.). *Advances in neural information processing systems*, *4*, 191-198. San Mateo, CA: Morgan Kauffmann Publishers.
- Hogan, N. (1984). An organising principle for a class of voluntary movements. *Journal of Neuroscience*, *4*, 1745-2754.
- Houde, J. F. & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, *279*, 1213–1216.
- Jordan, M.I. (1990). Motor Learning and the Degrees of Freedom Problem. In M. Jeannerod (Ed.), *Attention and Performance* (pp. 796-836). Hillsdale, NJ: Erlbaum.

- Jordan, M.I. & Rumelhart, D.E. (1992). Forward Models: Supervised Learning with a Distal Teacher. *Cognitive Science*, 16, 316-354.
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinions in Neurobiology*, 9, 718-727.
- Kawato, M., Maeda, Y., Uno, Y. & Suzuki, R. (1990). Trajectory formation of arm movement by cascade neural network model based on minimum torque-change criterion. *Biological Cybernetics*, 62, 275-288.
- Laboissière R., Schwartz J.L. & Bailly G. (1991). Modelling the Speaker-Listener Interaction in a Quantitative Model for Speech Motor Control: A Framework and Some Preliminary Results. *PERILUS XIV* (pp. 57-62). Stockholm, Sweden: Institute of Linguistics.
- Laboissière, R., Ostry, D.J. & Feldman A.G. (1996). Control of multi-muscle systems: human jaw and hyoid movements. *Biological Cybernetics*, 74, 373–384.
- Lane, H., & Wozniak, J. (1991). Speech deterioration in postlingually deafened adults. *Journal of the Acoustical Society of America*, 89, 859-866.
- Löfqvist, A. & Gracco, V.L. (2002). Control of oral closure in lingual stop consonant production, *Journal of the Acoustical Society of America*, 111, 2811-2827.
- Ostry, D.J., Gribble, P.L. & Gracco, V.L. (1996). Coarticulation of jaw movements in speech production: is context-sensitivity in speech kinematics centrally planned? *Journal of Neuroscience*, 16, 1570–1579.
- Manzella, J., Wozniak, J., Matthies, M., Lane, H., Guiod, P., & Perkell, J. (1994). Speech production before and after deafening. *Journal of the Acoustical Society of America*, 95, 3012(A).
- McFarland, D., Baum, S. R. & Chabot, C. (1996). Speech compensation to structural modifications of the oral cavity. *Journal of the Acoustical Society of America*, 100, 1093–1104.

- O'Shaughnessy, D. (1981). A study of French vowels and consonant durations, *Journal of Phonetics*, 9, 385-406.
- Perkell, J.S. & Klatt, D.H. (1986). *Invariance & Variability in speech processes*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Perkell, J. S., Matthies, M. L., Svirsky, M. A. & Jordan, M. I. (1993). Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot "motor equivalence" study. *Journal of the Acoustical Society of America*, 93, 2948–2961.
- Perkell, J., Matthies M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*, 22, 227-250.
- Perkell, J.S., Guenther, F.H., Lane, H., Matthies, M.L., Perrier, P., Vick, J., Wilhelms-Tricarico, R. & Zandipour, M. (2000). A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. *Journal of Phonetics*, 28, 233-272.
- Perrier P., Løevenbruck H. & Payan Y. (1996). Control of tongue movements in speech: The Equilibrium Point hypothesis perspective. *Journal of Phonetics*, 24, 53-75.
- Perrier, P., Payan, Y., Zandipour, M., & Perkell, J. (2003). Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. *Journal of the Acoustical Society of America*, 114, 1582-1599.
- Perrier, P., Payan, Y., & Marret, R. (2004). Modéliser le physique pour comprendre le contrôle : le cas de l'anticipation en production de parole. In R. Sock & B. Vaxelaire (Eds.), *L'anticipation à l'horizon du Présent* (pp. 159-177). Sprimont, Belgium: Pierre Margala (Collection Psychologie et Sciences Humaines).
- Saltzman, E.L. (1986). Task dynamic coordination of the speech articulators: a preliminary model. *Experimental Brain Research Series*, 15, 129-144.

- Saltzman, E.L. & Munhall, K.G. (1989). A dynamical approach to gesture patterning in speech production. *Ecological Psychology, 1*, 1615-1623.
- Savariaux, C., Perrier P. & Orliaguet, J.-P. (1995). Compensation strategies for the perturbation of the rounded vowel [u] using a lip-tube: a study of the control space in speech production. *Journal of the Acoustical Society of America, 98*, 2428-2442.
- Savariaux, C., Perrier, P., Orliaguet, J.P. & Schwartz J.L. (1999). Compensation strategies for the perturbation of French [u] using a lip tube. II. Perceptual analysis. *Journal of the Acoustical Society of America, 106*, 381-393.
- Slotine, J.J.E. & Li, W. (1991). *Applied non-linear control*. Prentice-Hall.
- Stager, S.V. & Ludlow, C.L. (1993). Speech production changes under fluency-evoking conditions in non-stuttering speakers. *Journal of Speech and Hearing Research, 36*, 245-253.
- Stevens, K.N. (1989). On the quantal nature of speech. *Journal of Phonetics, 17*, 3-45.
- Todorov, E., & Jordan, M.I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neurosciences, 5(11)*, 1226-1255.
- Tremblay, S., Shiller, D.M. & Ostry, D.J. (2003). Somatosensory basis of speech production. *Nature, 423*, 866-869.
- Trulsson, M., & Essik, G.K. (1997). Low-threshold mechanoreceptive Aafferents in the human lingual nerve. *Journal of Neurophysiology, 77*, 737-748.
- Wainscott, S.,K., Donchin, O. & Shadmehr, R. (2005) Internal models and contextual cues: encoding serial order and direction of movement. *Journal of Neurophysiology, 93*: 786-800