



HAL
open science

Motor Equivalence in Speech Production

Pascal Perrier, Susanne Fuchs

► **To cite this version:**

Pascal Perrier, Susanne Fuchs. Motor Equivalence in Speech Production. Melissa A. Redford. The Handbook of Speech Production, Wiley Blackwell, pp.225-247, 2015, 978-0-470-65993-9. 10.1002/9781118584156.ch11 . hal-01205656

HAL Id: hal-01205656

<https://hal.science/hal-01205656v1>

Submitted on 29 Sep 2015

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.



HAL Authorization

Motor Equivalence in Speech Production

Pascal Perrier & Susanne Fuchs

Abstract

The first section provides a description of the concepts of “motor equivalence” and “degrees of freedom”. It is illustrated with a few examples of motor tasks in general and of speech production tasks in particular. In the second section, the methodology used to investigate experimentally motor equivalence phenomena in speech production is presented. It is mainly based on paradigms that perturb the perception-action loop during on-going speech, either by limiting the degrees of freedom of the speech motor system, or by changing the physical conditions of speech production or by modifying the feedback information. Examples are provided for each of these approaches. Implications of these studies for a better understanding of speech production and its interactions with speech perception are presented in the last section. Implications are mainly related to characterization of the mechanisms underlying interarticulatory coordination and to the analysis of the speech production goals.

I. What is motor equivalence in speech production?

Motor equivalence is a very important property of motor control in animals and in humans, which could be roughly defined as the capacity to achieve the same motor task differently. Motor equivalence is currently observed in our everyday life: we can grasp an object on the ground by bending the back or by crouching down; we can paint a wall while holding the brush in the right or in the left hand; we can turn on the light by pushing the switch with the finger or with the elbow... These simple examples illustrate well the power of motor equivalence in motor control: motor equivalence offers freedom to the motor system and this freedom can be used to deal with various kinds of additional constraints. Lifting a heavy object from the ground is more efficient if it is done using the legs rather than the back; painting alternately with the right and the left hand is less exhausting than painting continuously with the same hand; turning on the light with the elbow enables us to hold an object in the hands at the same time. This capacity of the motor system to adapt certain strategies depending on external constraints is called the “*plasticity*” of the motor system. The plasticity of the motor system allows the Central Nervous System to fulfill the intended motor task properly, while integrating constraints. These constraints can be, for instance, the search for efficiency, as in lifting an object using our legs rather than using our back, or the need for the parallel execution of other motor tasks. Plasticity is also the basis for motor rehabilitation when the peripheral motor system is damaged and certain body parts take over for others.

Motor equivalence exists in speech production. A typical and well-known example of motor equivalence is the ability to communicate efficiently with listeners either while speaking with a free-moving jaw (the normal way of speaking) or while speaking with a pencil wedged between the upper and lower teeth restraining the jaw in a fixed position. Motor equivalence in speech has been clearly documented in a number of experimental studies. For example, Hughes & Abbs (1976) studied inter-articulatory variability of /æ, i, ε/ vowels during repetitive productions of Consonant-Vowel syllables in native speakers of American English. They observed that for the same distance between the lower and upper lips, different individual positions of the lower lip, upper lip and jaw were reached. Variations in the position of one of these articulators were counterbalanced by coordinated variations of the two

other articulators, in order to keep the distance between the lips at a constant value. This is a motor equivalence strategy, one that uses the property that various inter-articulatory configurations can lead to the same interlabial distance. Interlabial distance was assumed to be the motor goal of the task. Similarly, Maeda (1990) observed the variability of the jaw and the tongue dorsum positions during the production of /i/ and /a/ in various phonetic contexts for native speakers of French. He reported that jaw height and tongue dorsum front-back position cooperated to ensure the achievement of the required vocal tract shape for each vowel. The vocal tract shape was assumed to be the motor goal. The common observation for both vowels was that insufficient jaw height was counterbalanced by a more anterior positioning of the tongue, and, reciprocally, a high jaw position was associated with a more posterior tongue positioning. For vowel /i/, when jaw was low, the more anterior positioning was associated with an elevation of the tongue with respect to the jaw. This was presumably done in order to keep the constriction small and anterior enough. For vowel /a/, when jaw was low, the more anterior tongue positioning was presumably achieved in order to prevent for the reduction of the constriction area in the pharynx.

Motor equivalence has also been found at the level of the vocal tract shape to achieve an acoustic goal (e.g. Perkell et al., 1993). Perkell and colleagues analyzed a large number of repetitions of the rounded vowel /u/ in native speakers of American English. They observed a significant negative correlation between tongue raising, which determines the size of the constriction inside the vocal tract, and lip rounding, which determines the size of the constriction at the lips. A narrow constriction at the lips was associated with a more open constriction in the vocal tract, and vice-versa. These two constrictions are the major factors determining the frequency of the second maximum of amplitude (second formant or F2) in the spectrum of the vowel /u/. A low value of this frequency is the most important characteristics of /u/ from an auditory point of view. This acoustic property was therefore considered to be the motor goal. The narrower both constrictions are, the lower the second formant. The negative correlation observed by Perkell et al. (1993) is consistent with the idea of motor equivalence. If one of these constrictions is too large (a property that tends to increase F2), the other constriction is adjusted accordingly.

Kelso & Tuller (1983) summarized motor equivalence as follows: “Within limits, people (and animals) can achieve the same « goal » through a variety of kinematic trajectories, with different muscle groups and in the face of ever-changing postural and biomechanical requirements.” (p. 217). This description is interesting because it contains the two most important key words that refer to the main issues that are addressed in motor equivalence studies in speech production.

The first key-word is “**goal**”. The relative invariance of the goal across different motor strategies is the basic requirement for the existence of motor equivalence. The definition of the goal determines whether motor equivalence exists or not for a given motor task. One of the peculiarities of speech production, as compared to other human motor tasks, is that the goal is not physical but cognitive in nature. Basically, in speech production the goal is the meaning of the message that is transmitted to the listeners. By defining the goal in this way, we can see a huge variety of possibilities for motor equivalence in speech production. In this context, motor equivalence not only includes equivalences in terms of “muscle recruitments” or “kinematic trajectories”, but also in terms of lexical and semantic features. Studies of motor equivalence in speech production have not considered all these possibilities. Rather, they have strongly reduced the scope of investigation by considering goals at a phonological level. The tasks investigated have concentrated on the production of a phoneme or of sequences of a few phonemes and so on the motor goals associated with the production of a phoneme, or of a short sequence of phonemes. The three examples of experimental studies cited above, illustrate well the complexity of even this restricted problem of phoneme production. In these studies, goals have been defined either in the articulatory domain or in the acoustics. This point will be further developed in section III.A.

The second key-word is “**variety**”. Motor equivalence enables motor tasks to be achieved in a variety of ways. In speech, articulatory and acoustic variability is frequently observed across phonetic contexts and/or speaking styles, and in this respect it is called “coarticulation”. Understanding what the degrees of freedom are for the speech motor system and how these degrees of freedom are taken into account by the Central Nervous System is crucial for explaining and predicting coarticulation phenomena. Thus, studies of motor equivalence in speech production are important for studies of inter-articulatory coordination underlying coarticulation phenomena. This point will be further developed in section III.B.

Methodologically, motor equivalence in speech has often been investigated via experimental studies. Sometimes these experimental studies have been associated with modeling work, and experimental data have been compared to the results of simulations. The basic idea underlying the design of experimental studies is to generate variability in the realization of the same goal. In normal speech, variability can be generated by varying phonetic context or speaking style (speaking rate, level of clarity, prosodic patterns...). It can be also induced by introducing perturbations to the production-perception loop. The latter approach allows an investigation of the motor task. It has been used in the majority of studies on motor equivalence. The next section is concerned with the presentation of the general experimental methodology of these studies.

II. Methodology for studying motor equivalence in speech

The majority of experimental studies of motor equivalence phenomena in general, and of speech production in particular, have used perturbation paradigms. The basic idea underlying these paradigms is to provide significant changes to the conditions in which the motor task is achieved. These applied changes disturb the way in which motor strategies are used to fulfill the respective motor task under normal conditions. In the absence of motor equivalence, perturbations would definitely prevent the subjects from fulfilling the motor task. When motor equivalence strategies exist for a given task, perturbations are an efficient way of studying the link between coordinated articulators or the perception-action loop. In addition, perturbations allow investigations of the degrees of freedom for a given task and how the Central Nervous System is interacting with the external world to preserve the achievement of the motor goal. Perturbation studies differ with respect to the level they apply to. They can be classified in three ways:

- Perturbations that limit the degrees of freedom of the motor system
- Perturbations that change the physical conditions of speech production
- Perturbations that change feedback information

More details are given below for these three kinds of perturbations.

A. Perturbations that limit the number of degrees of freedom

Perturbations that limit degrees of freedom prevent the use of one or more degrees of freedom in the fulfillment of the speech production task. The perturbation can be static, i.e., constant for the duration of the motor task, or variable in time, i.e., applied at a certain number of time points and for a certain duration. By looking at the way the Central Nervous System uses the remaining degrees of freedom to achieve the task, it is possible to gather interesting information about degrees of freedom in general, about the nature of the goal that is preserved in spite of the perturbation, and about the process underlying the development of motor equivalent strategies in the restricted space of the remaining degrees of freedom.

A number of phonetic studies on different languages have shown that jaw position can vary during the production of vowels. This observation, together with the fact that speaking is possible with an object maintained between the teeth, has led to the conclusion that jaw opening is a degree of freedom that can be manipulated in vowel production. The first motor equivalence studies in speech were applied to the jaw. One of the most popular static perturbations of jaw opening has been a bite-block (Lindblom, Lubker & Gay, 1979; Gay, Lindblom & Lubker, 1981). Bite-blocks are small rigid blocks, a few millimeters in depth and width, that are inserted between the lower and upper teeth of the subject. This perturbation constrains the jaw to stay at a constant position during the production of speech. Depending on the height of the block, the jaw is fixed either at a high position (bite block height=3-5mm) or at a low position (bite block height=20-25mm). Another perturbation of the jaw consists in applying resistive loading to the articulator during the closing movement of a bilabial stop (Folkins & Abbs, 1975; Kelso et al., 1984). The loading prevents the jaw continuing its upward movement toward the position usually reached under normal conditions in the bi-labial stop. However, as long as the loading is not too strong, this perturbation does not prevent the achievement of the labial closure that is required for bilabial stops. Since the labial closure is the result of the combined influences of the jaw, the upper lip and the lower lip positions, the limited amplitude of the jaw can be compensated for by increasing the amplitude of the movements of the lower and upper lips.

A similar unexpected perturbing force can be applied to the lower lip, rather than to the jaw, in bilabial stops (Gracco & Abbs, 1985) or in bilabial fricatives (Gomi et al., 2002). Again the bilabial closure or constriction can be achieved by a downward shift of the upper lip.

Acoustic models of speech production (Fant, 1960) have shown that the position and size of the constriction within the vocal tract, and the shape of the lips (spread or rounded) are the two main factors influencing the spectral characteristics (the formants) that are relevant for the perception of vowels. These models have also shown that for the vowels /u/ and /ʊ/ two different constriction patterns are possible in the vocal tract. A constriction in the palato-velar region is possible in association with a small lip area, or a constriction in the velo-pharyngeal region can be used in combination with a larger lip area. For these vowels, lip opening is one of the degrees of freedom. Perturbation of the lip opening has been manipulated by Savariaux et al. (1995, 1999). The perturbation consists of using a 25mm-diameter tube to prevent the achievement of the small lip area required for rounded lips. Based on the acoustic theory of speech production, it is theoretically possible to produce a vowel in the presence of this perturbation by constricting the velo-pharyngeal region of the vocal tract.

B. Perturbations that change the physical conditions of speech production

A second kind of perturbation consists in modifying the physical conditions of speech production. The idea underlying these paradigms is quite different from the preceding one. It is not a matter of restricting the degrees of freedom in the speech production system, but of changing the way the articulators interact with the vocal tract boundaries or of changing the dynamical constraints, i.e. the forces applied to the articulators. Under these perturbed conditions, the Central Nervous System has to face the fact that the usual motor control strategies do not produce the expected effects, either in the articulatory or in the acoustic domain. The ultimate goal of the task, i.e., the production of a phoneme or of a short sequence of phonemes, does not change, but new strategies have to be found to reach this goal. In that sense it is a motor equivalence problem. However, these perturbations involve a level of adaptation other than the perturbations described in section II.A, since a new mapping between articulators and acoustics or between applied forces and articulatory dynamics has to be taken into account. Available motor equivalence strategies for normal speech may not be used to the same extent. The Central Nervous System has to explore new possibilities for motor equivalence and to elaborate new motor control strategies in order to reach the same goals as under normal conditions. The purpose of this kind of perturbation studies is to observe how the Central Nervous System explores motor plasticity in the articulatory and acoustic domains to preserve the achievement

speech goals. These studies provide information on the nature of the goal, the plasticity of the motor system, and on the way new strategies can be developed.

A first example for this type of perturbation is to modify the morphology of the vocal tract. When the morphology of the vocal tract changes, articulatory movements do not shape vocal tract cavities in the same way as under normal conditions. Consequently, the relationship between articulatory movements and acoustics also changes. Perturbations of the vocal tract morphology are classically introduced with artificial palates of different thicknesses (McFarland et al., 1996) or different shapes (Brunner et al., 2011) or with a dental prosthesis in which the upper incisors are longer than under normal conditions (Jones & Munhall, 2003). These perturbations are static. Unexpected and time variable perturbations have also been provided using an inflatable artificial palate (Honda et al., 2002).

A more complex perturbation to the speech production apparatus consists in applying a time varying force field to the jaw. The intensity of the force field varies in time as a function of the velocity of the jaw movement: the faster the movement, the stronger the force field. No force is applied both at the beginning and at the end of the movement, since the velocity is zero at these positions. Rather, the force field modifies the mechanical conditions for the displacement between these two extreme positions (Tremblay et al, 2003). Thanks to the existence of motor equivalence in the levels of muscle activations underlying a given jaw movement, the Central Nervous System can develop new motor strategies to resist against the influence of the external force field, and ensure the achievement of the goal.

C. Perturbations that change feedback information

The third kind of perturbation does not affect the speech production system itself. It modifies the way the Central Nervous System can assess whether motor goals are reached by altering oro-sensory (tactile and/or somato-sensory) or auditory feedback. Motor equivalence strategies with altered feedback are different from strategies that are used under normal conditions, since the entire production-perception system is taken into account, and goals are defined in sensory rather than in physical terms. Altered feedback perturbation provides information on the motor control processes underlying the use of motor equivalence strategies, on the role of feedback in the selection of motor equivalence strategies, and on the nature of the speech production goals in the sensory domain.

Altered oro-sensory feedback is frequently produced by applying an anesthetic that reduces the amplitude of oro-sensory feedback, though anesthesia has the drawback of effects that are hard to control in their extent. Anesthesia has been applied to the temporomandibular joint, which provides information on jaw position (Kelso & Tuller, 1983), and to the oral mucosa, which provides tactile feedback information (Kelso & Tuller, 1983). Recently, Ito & Ostry (2010) perturbed the kinesthetic information provided by the cutaneous receptors in the skin of the face. These receptors provide information on the stretching of the skin, which is influenced by the positioning of the jaw and by the spreading or protrusion of the lips in natural speech. Perturbing this feedback information modifies the perception of the positions of these speech articulators.

Altered auditory feedback can also be used to perturb the production-perception system. Further details are given in Chapter 14 of this book. A common perturbation consists in restricting the availability of auditory information by presenting very loud white noise (classically 80 to 90 dB) to the subjects via headphones (Kelso & Tuller, 1983; Brunner et al. 2011). This perturbation should be used with caution since it is known to induce significant changes in speech articulation, according to the well-known “Lombard effect” (Summers et al, 1988). Another perturbation, more complex to implement, is altering the spectral characteristics of the speech signal. This perturbation can affect the formant values of vowels in whispered speech (Houde & Jordan, 1998) or in normal speech (Purcell & Munhall, 2006a), the spectral Center of Gravity (COG) of fricatives (Shiller et al., 2009), or the fundamental frequency (Jones & Munhall, 2002). In all these experiments, natural auditory feedback due to acoustic wave propagation in the air and to bone conduction of acoustic vibrations has to be masked. This is why the perturbed auditory feedback is presented to the subjects via headphones and at a relatively high acoustic level (again 80 to 90 dB).

All the perturbation paradigms described above are used to study motor equivalence in speech production with the following two aims: First, to investigate either the nature of the speech motor goals, or the motor control processes underlying the use of the degrees of freedom to organize interarticulatory coordination. Second, to study the role of feedback in speech motor control. Hence, the data recorded during these experiments are acoustic signals, articulatory movements, vocal tract geometries, and muscular activations (using electromyography). Recently, brain activations have also been recorded to look at the cortical correlates of the adaptation mechanisms (Tourville et al., 2006). Acoustic signals are in general analyzed in the spectral domain to investigate aspects of the spectral

envelope. This envelope provides information on the spectral maxima (e.g. formants) and the distribution of the energy in the frequency domain. The measured acoustic properties in the spectral domain often reflect auditory cues to the perception of phonemes. This is why we will use the terms acoustic and auditory interchangeably in this chapter.

III. What can be inferred from motor equivalence studies?

Motor equivalence phenomena are interesting in speech production research because they are an efficient way of addressing fundamental issues: the nature of speech goals and the interarticulatory coordination underlying speech variability due to coarticulation and variations in speaking style. Major contributions to these topics will be developed below.

A. The nature of the goals

Since the end of the seventies, numerous studies have investigated the nature of speech goals. Speech production consists in moving the orofacial articulators in order to provide appropriate changes in the vocal tract shape. The vocal tract shape determines the resonance modes of the vocal tract which are excited by the acoustics associated with vocal fold vibrations and turbulent air flow. These resonance modes are the major factors that determine the spectral peaks of the acoustic speech signal, and these spectral peaks are the main cues used by the auditory speech perception system. In addition, it is known that speech can be perceived, at least partly, in the absence of an acoustic signal or with a highly degraded signal, when listeners can see the face of the speaker and the movements of the orofacial articulators. This multimodal, i.e., motor, articulatory and acoustic, nature of the speech production process has led to a debate about the nature of the goals in speech production. The debate was particularly strong in the eighties in phonetics and phonology, and has now been taken up in the neuroscience domain. In the phonetics-phonology literature the question was: Are goals specified in terms of articulatory movements (Fujimura, 1986), in terms of vocal tract shapes (Fowler, 1986) or in terms of spectral characteristics (Stevens, 1972)?

A major difficulty in the investigation of speech goals is the fact that in normal speech production the characteristics that are measurable for one speaker, in terms of muscle activations, articulatory positions, vocal tract shapes and spectral properties, are quite strongly related for a given sequence of

phonemes. To disentangle these links, experimental paradigms generating motor equivalencies are efficient because they generate variability in one domain, but not necessarily in another. In all the experimental studies on motor equivalence, the methodology is based on the following assumption: When variability occurs in speech production, goals correspond to the least variable properties across conditions; properties that are significantly more variable are the degrees of freedom of the task.

There are numerous studies of motor equivalence that investigate the nature of goals in speech production. Some of these have been cited in Section II to illustrate different experimental paradigms. There exist many others that could be acknowledged here. However, we will focus in this section on only a few that illustrate well, from our point of view, the challenges and approaches of these studies.

1. Bite-block experiment: Motor goals for vowels are the acoustically most significant characteristics of the vocal-tract geometry

Gay et al. (1981) studied five native speakers of Swedish during the production of the long vowels /i, a, u, o/ in a series of Vowel-Vowel-Vowel sequences. Subjects were asked to produce these sequences either under normal conditions or with a bite-block. For the closed vowels /i, u, o/ the bite block was high (22.5mm) and for the open vowel /a/ it was low (2.5mm). The task was to produce the vowels with the same quality across the normal and the bite block conditions. Lateral X-ray views of the vocal tract were recorded. These data allowed measurements of individual articulatory positions and the shape of the vocal tract in the mid-sagittal plane of the head. The acoustic signal was recorded and analyzed. According to simulations with an articulatory model of the vocal tract, changes in jaw position due to the insertion of the the bite block should have induced a significant change in the formant patterns if there was no compensatory change in tongue position. The results of the study were consistent for all subjects: differences in formant patterns between the normal and the bite block condition were much smaller than differences that would have occurred in the absence of compensatory strategy; this was true even in the first trial. The smallest variation in the articulatory domain occurred in the region of maximum constriction in the vocal tract. Cross-dimensional deviations increased with an increase in distance away from the point of maximum constriction. The authors concluded that the goal of vowel production is defined in terms of formant patterns and that this goal was “*coded neurophysiologically in terms of area-function related information and is specified*

with respect to the acoustically most significant area-function features, the points of constriction along the length of the tract.” (p. 809).

Gay et al.'s (1981) study would prove very important in discarding the hypothesis that speech production goals could be specified in terms of absolute position of the articulators in the vocal tract. The results showed the importance of the acoustic domain in specifying goals. That said, perturbations using bite-block do not allow for testing whether vowel goals are specified in terms of constrictions in the vocal tract or in terms of formant patterns in the acoustic domain. A bite block perturbs the usual articulatory positions associated with the production of the vowel, but it does not prevent speakers from attaining the usual geometric shape of the vocal tract associated with the usual formant pattern of this vowel. To test whether vowel goals are specified in terms of vocal tract constriction or formant pattern, Savariaux et al. (1995) conducted a study using a perturbation of the usual geometric shape of the vocal tract. This study is described next.

2. Lip-tube experiment: Vowel goals are auditory under vocal tract related constraints

Savariaux et al. (1995) perturbed the production of the French /u/, pronounced in isolation, by introducing a 25mm diameter tube between the lips of 11 native speakers of French. The tube, henceforth called a *lip tube*, induced a large increase in lip area, without altering the jaw position much. The subjects were asked to produce an /u/ with the lip tube. Immediately after the insertion of the lip tube the subjects received a training session. In this session the subjects had the lip tube in place and were allowed to produce /u/ 19 times in order to compensate for the perturbation, if they felt that compensation was needed. At the end of the training session, the subjects were asked to produce /u/ once again with the lip tube in place using the strategy they considered to be the best among the 19 preceding training trials. Compensatory strategies were analyzed in the articulatory domain with lateral X-ray views.

The production of /u/ in French is normally achieved with very rounded and protruded lips and a high and back tongue position, generating a vocal tract constriction in the palato-velar region. However, the acoustic theory of vowel production (Fant, 1960) predicts that the same (F1, F2) pattern can also be produced with open lips and with a vocal tract constriction in the velo-pharyngeal region.

An analysis of the acoustic signal in terms of fundamental frequency and formant patterns (F1, F2, F3) was performed, and perceptual tests were run for 10 of the 11 subjects to evaluate the perceptual quality of the vowel /u/ produced under perturbed conditions (Savariaux et al., 1999). Results were as follows. First, in contrast to the observations of Gay et al. (1981), none of the subjects could compensate for the lip tube perturbation in the first trial. Compensation was not achieved immediately, but after a number of trials in the training session. Second, during the training sessions all the subjects shifted the formant pattern (F1, F2) of their /u/ in direction of the pattern produced without the perturbation. Third the extent of this compensation was highly variable across subjects. For four subjects, the perception tests revealed full compensation, for three subjects partial compensation, and for three subjects no compensation at all. The articulatory data indicated that only one subject radically moved the vocal tract constriction from the palatal to the velo-pharyngeal region. The other subjects moved their tongue backwards, but to a lesser extent. Those who improved the perceptual quality also changed their fundamental frequency.

These observations support the hypothesis that the goal of vowel /u/ is specified in the auditory domain in terms of formant frequencies with some additional influence of fundamental frequency. However, variability in the extent of compensation across subjects shows that speakers use also orosensory feedback to compensate for a perturbation, that the role of this feedback is speaker-specific and that it may rely on speaker-specific vocal tract properties

3. Dental prosthesis: Goals for sibilants are auditory

Jones & Munhall (2003) investigated the contribution of auditory feedback in adapting to a dental prosthesis during the production of the fricative /s/ in Canadian English. The dental prosthesis lengthened the upper incisors between 5 and 6 millimeters, without affecting the bite of the six speakers in the study. In contrast to the bite block and the lip tube, this perturbation did not prevent the subjects from producing their usual constriction in the vocal tract, nor did it hinder proprioceptive or tactile feedback in the vocal tract.

In the production of /s/, the noise source arises from a jet of air, generated by the vocal tract constriction, hitting the surface of the front teeth (Shadle, 1989). The resulting noise excites

resonances mainly in the small front cavity located between the constriction and the lips, which maximizes spectral energy in the high frequencies (above 5kHz). According to Jones & Munhall (2003), lengthening the upper incisor teeth induces an effect that is comparable to the enlargement of the front cavity. In the absence of a correction, the larger cavity results in lowering the frequency of maximum spectral energy, such that /s/ will sound more like /ʃ/.

Jones & Munhall's (2003) experimental session consisted of two sub-sessions. Each of these sub-sessions consisted of 15 blocks of 10 repetitions of /tas/ under 4 different conditions: (C1) normal condition; (C2) without the dental prosthesis in the mouth and with masked auditory feedback consisting of white noise; (C3) with the prosthesis in the mouth and with masked auditory feedback; (C4) with the prosthesis in the mouth and with normal auditory feedback. The ordering of the 15 blocks was as follows: C1, C2, C3, C4, C3, four alternations (C4-C3), C2, C1. The acoustic production of /s/ was evaluated by measuring the Center of Gravity of the spectrum and the ratio between the slope of the spectral envelope below 2.5kHz and the slope between 2.5kHz and 8kHz. Perception tests were also run with 16 listeners.

Results were as follows. In the first block of condition C3 (no auditory feedback), the acoustic production of /s/ was altered by the dental prosthesis, and the spectral impact conformed to the theoretical prediction: /s/ sounded more like /ʃ/. When auditory feedback was available, a trend for improvement was found in the slope ratio. The perception tests also revealed that utterances were of significantly higher quality when auditory feedback was available during the production than when it was masked. In addition, the perception tests showed that the sounds produced in the presence of auditory feedback improved over several repetitions with training. This was not the case when auditory feedback was masked.

The observation that compensatory strategies are developed when a dental prosthesis is in place, together with the findings from the perception test, which show the role of auditory feedback in the development of such strategies, supports the hypothesis of an auditory goal for the sibilant /s/.

4. Velocity dependent force fields: Articulatory components are part of the goals

Tremblay et al. (2003) applied a velocity dependent force field to the jaw during speech and non-speech mandibular movements. The force was applied in the mid-sagittal plane of the head along an axis parallel to the occlusal plane and in the direction of jaw protrusion (methodology described above in Section II.B). Three different conditions were tested with native speakers of English: productions of the utterance /siat/ in slow and clear speech; articulation of the same utterance slowly and clearly without vocalization (silent speech); non-speech jaw movement that matched the amplitude and duration of the two speech conditions. For each condition the session started with 20 repetitions of the task without perturbation, it continued with 20 repetitions with perturbations of the jaw, and finally the perturbation was removed and the task was again repeated 20 times. The movement of the jaw in the mid-sagittal plane was recorded together with the acoustic signal. The first two formant frequencies (F1, F2) were measured to characterize the produced speech sequence in the auditory domain.

Results were as follows. In the first trials following the introduction of the force field, a noticeable modification of the motion path of the jaw was observed for all subjects and for the three conditions. After training, an adaptation to the perturbation was observed for the two speech conditions (silent and vocalized speech), but not for the non-speech condition. After a few trials, the motion path of the jaw during speech became similar to the one produced without the perturbation, and differences in the (F1, F2) patterns were minimal. In addition, perceptual tests indicated that listeners could not systematically distinguish between speech stimuli with and without perturbations. These last two results suggest that the compensation to force field after training was not guided by any goal in the auditory domain. No similar compensatory effects were observed after training in the non-speech condition.

Since speech and non-speech movements used the jaw in very similar ranges of displacement and duration, the differences observed between speech and non-speech conditions cannot be attributed to any peripheral phenomenon, such as muscle mechanics or jaw dynamics. Rather, the differences must have their origins in the motor control strategies underlying the speech and non-speech movements. According to Tremblay et al. (2003), the different strategies reflect differences in the specification of the goals. For the speech task, time-varying somatosensory feedback during movement is part of the specification of the goal, while this is not the case for non-speech movements.

5. Auditory feedback perturbation: Articulation is modified to reach auditory goals

Purcell & Munhall (2006b) studied the effect of perturbing auditory feedback during the production of the vowel / ϵ / in “head” in ten native speakers of English. The perturbation consisted of either a positive or a negative shift of the first formant F1. After a first phase where the subjects produced the word “head” without any perturbation, F1 was gradually shifted within a range of +/-200Hz in 50 successive steps of 4Hz. At the end of this gradual shift, the subjects were asked to repeat the word 15 times again. The formant patterns were measured continuously during three phases of production: no perturbation, positive and negative shifts in F1. On average, the subjects provided changes to F1 in the opposite direction of the shift. However, compensation was never complete. In addition, subjects did not change their production immediately after the shift started. A minimal shift was necessary, which was on average 76Hz. This value is well above the psychoacoustic threshold (around 20Hz for vowel / ϵ /, according to Kewley-Port, 2001), where spectral differences start to be perceived.

In summary, the Purcell & Munhall (2006b) experiment confirms that vowel goals are in the auditory domain and that speakers change their articulation and corresponding vocal tract shapes if they do not generate the expected auditory goal. The threshold value and the incomplete compensation at the end of the shift phase suggest that compensation is related to a phonetic goal (i.e. a region in the acoustic space) and not to a psychoacoustic threshold.

6. Time varying inflated palate: Speech goals are articulatory and auditory

Honda et al. (2002) designed a time-varying perturbation of the palate based on a balloon glued on to a thin artificial palate. The balloon could be inflated or deflated within a maximum of 60ms by external air pressure through a lead tube. The maximum thickness of the palate (inflated balloon) was 5mm, and the minimum thickness (deflated balloon) 1mm. The perturbation was applied to two native speakers of Japanese during utterances containing repetitions of the syllables / \int a/ or /t \int a/, and was done in two different ways. In one session, the artificial palate was mainly deflated and became randomly inflated in 20% of the utterances; in another session, the pattern was similar with the inflated palate as the basic condition and deflation as the random perturbation. The perturbation was applied just before the first syllable of each utterance and was maintained throughout the utterance. Each session was run twice, first with auditory feedback available, and then with auditory feedback masked

by pink noise. Articulatory data were recorded in the mid-sagittal plane and perception tests were run. The listeners' task was to identify the consonant in the syllable. The articulatory data recorded in the random inflation and deflation conditions were compared with productions in the steady-state deflated palate condition.

The results were as follows. The random deflation of the palate generated very few errors. The impact of the perturbation was much stronger when the palate was randomly inflated: we will therefore concentrate on this perturbation. For both subjects, the first syllable was incorrectly produced and incorrectly perceived under the perturbed condition, but the second syllable was significantly improved on these measures compared to the first. The articulatory data revealed that compensatory adjustments of the tongue position started within 75ms and 150ms after the contact onset between tongue and palate in the first syllable. In the authors' opinion, this latency is too short to enable the processing of auditory feedback at a cortical level. In addition, in the first two syllables no difference could be noted in the perception tests between the condition with auditory feedback available and the condition with masked auditory feedback. Hence, we interpret these results as evidence that the compensatory response to the inflation of the palate is driven initially by tactile feedback rather than by auditory feedback. Significant differences between the two conditions, with and without auditory feedback, started in the third syllable and they were maintained up to the end of the utterance. With auditory feedback available, the quality of the consonant improved rapidly. Thus, from the middle of the utterance onwards no significant difference could be found in the identifications of the sounds pronounced with and without perturbation. With masked auditory feedback, no additional improvement in production was noticed after the second syllable.

In summary, Honda et al.'s (2002) results suggest that important characteristics of the goal are specified in terms of the proximity of the tongue and the palate, which is monitored by tactile feedback. The results also suggest that auditory feedback is necessary if the goal is to be reached with the required accuracy. This supports the idea of speech goals specified both in articulatory and acoustic terms, with a predominant requirement in the acoustics.

7. Summary

Since the pioneering work of Gay et al (1981), numerous studies have investigated motor equivalence phenomena in speech production in order to understand the nature of speech production goals. As the

selection of studies we have described exemplifies, our interpretation of this rich literature is that speech goals have components both in the acoustic/auditory and in the articulatory/motor domains. The studies by Gay et al (1981), Savariaux et al. (1995), Jones & Munhall (2003), Purcell & Munhall (2006b) clearly show that in response to various perturbations subjects change their usual articulatory strategies and their usual vocal tract shapes to generate the expected acoustic/auditory properties. To our knowledge, there exists no example of speech perturbations in the articulatory/motor domain that do not cause subjects to use compensatory strategies if the perturbations prevent them from achieving the acoustic component of the speech goal. In addition, compensatory strategies systematically tend to make acoustic properties closer to those produced under normal speaking conditions. Similarly, there is no example of experiments where modified acoustics are accepted by subjects, and subjects keep their articulatory/motor strategies constant. Hence, speech goals must be primarily acoustic in nature.

However, goals are also defined in the articulatory/motor domain. Tremblay et al. (2003) showed that if speech acoustics are not altered, then preferred articulatory patterns exist. Honda et al. (2002) found evidence in support of the predominant role of tactile feedback in the first stage of a compensatory response to randomly applied palatal perturbations. They also showed that accurate compensation requires auditory feedback in a second stage. Savariaux et al.'s (1995) results suggest that, in the absence of any possibility of reaching the acoustic/auditory goal, subjects use their preferred vocal tract shape to produce /u/. Our interpretation of these perturbation studies is that speech goals have both articulatory/motor and acoustic/auditory components, but that there is a hierarchy between these two components: both components of the goals can influence the emergence of compensation strategies, but the articulatory/motor component is influential only if its achievement is compatible with the achievement of the acoustic/auditory component.

Our interpretation of the results from perturbation studies is compatible with a recent experimental study by Feng et al. (2011). Their eight subjects, all native speakers of English, were asked to produce CVC sequences under four conditions, where V was either /e/ or /æ/. The conditions were a normal speaking condition, a condition with a perturbation of the auditory feedback corresponding to a positive shift of the first formant F1, a condition with a vertical force field applied to the jaw either

upwards or downwards, and a condition with both the auditory feedback perturbation and the force field perturbation.

The results were as follows. Subjects systematically compensated for both auditory feedback and force field perturbations when they were applied separately. When the perturbations were applied together, subjects systematically compensated for the auditory feedback perturbation, but they compensated for the force field only in the cases where the compensation was compatible with the achievement of the acoustic goal. This experiment clearly supports our conclusion: speech goals are auditory and articulatory, but auditory first.

A more recent experimental study by Lametti et al. (2012) has tempered the strength of our conclusion. Here again the authors combined a perturbation of the auditory feedback (shift of the first formant) with a perturbation of the somatosensory feedback (alteration of the jaw path with a robotic device). Seventy-five native speakers of English were recorded during the production of the words “had” and “head”. They were distributed among five sub-groups according to five different experimental conditions: (1) speech production with jaw perturbation only; (2) speech production with auditory perturbation only; (3) speech production with both auditory and jaw perturbations simultaneously; (4) speech production with an auditory perturbation in the first half of the trials and with a jaw perturbation in the second half of the trials ; (5) speech production with a jaw perturbation in the first half of the trials and with an auditory perturbation in the second half of the trials. In the group 3 (both perturbation at the same time), part of the subjects compensated more for the auditory perturbation than for the somatosensory perturbation (in agreement with Feng et al, 2011), but another part compensated more for the somatosensory perturbation than for the auditory perturbation (in opposition with Feng et al., 2011). Similar results were obtained with the subjects of groups 4 and 5, for which the perturbations were applied sequentially and in different orders. These results suggest that the predominance of a modality could be strongly speaker-dependent: for some speakers acoustic could be first, as we hypothesized; but for some other speakers, articulation could be first...

Further studies will be necessary to assess the possible predominance of the acoustic component of the speech goal more precisely, and to evaluate whether it applies in the same way to all sounds in all languages. Motor equivalence will continue to be an excellent paradigm for this purpose. However, a refinement of the data analysis could be necessary. In this regard, the concept of “Uncontrolled

Manifold” proposed by Scholz & Schöner (1999) could be useful. For a given motor task, some motor commands do not vary much across repetitions. For this reason, they are assumed to be related to the fundamental goal of the task. Some other motor commands, less closely related to the achievement of the goal, can vary more. All the possible values of these less strictly controlled motor commands define a subspace of the motor command space. Scholz & Schöner (1999) assumed that this subspace has a continuous structure that they called “Uncontrolled Manifold” (UCM). Each UCM is attached to a specific motor goal, and UCMs change across motor goals. For a given motor task, UCMs correspond to the subspace defined by the degrees of freedom of the task, as introduced in Section I. The subspace that is orthogonal to the UCM attached to a motor goal, corresponds to motor commands whose variations induce a change of the motor goal value. Scholz & Schöner (1999) suggested that, in order to assess whether a motor goal is constant across repetitions of the task, it is useful to look at the structure of the variance within the UCM that is related to the motor goal. If the variance is mainly within the UCM, it can be concluded that the motor goal is preserved across the repetitions of the task. If the variance is orthogonal to the UCM, it can be concluded that the motor goal varies across repetitions. The computation of the UCM requires the existence of a model of the motor system in the brain, which enables a general description between motor commands and motor goals. Depending on the nature of the motor commands, this can be more or less complex. In speech production studies, if motor commands can be assumed to be the measurable positions of the articulators, the proposed estimation of the variance could be calculated quite easily from experimental data. If motor commands are assumed to be muscle-related commands, the implementation could be complex and could require the use of physiological models of the orofacial articulators to infer motor commands from experimental data. A critical analysis of the use of UCM in speech motor control has been presented in Saltzman et al. (2006).

B. Motor control strategies for interarticulatory coordination

The concept of motor equivalence has also been used to study motor control strategies that enable coordination between articulators. Motor equivalence phenomena are interesting in this regard because they give insight into how degrees of freedom are selected in various conditions to reach a motor goal. There is an ongoing debate between the concept of “coordinative structures” and the “prediction-error-correction” principle.

The concept of “*coordinative structures*” (Bernstein, 1977) considers that, for a given motor task, the degrees of freedom of the motor system are not controlled individually, but within more global functional units, called “coordinative structures”. Coordinative structures are not inherent to the speech production system. They are functional units and are used to achieve a specific motor task (Saltzman, 1986, 1991).

The “*prediction-error-correction*” principle assumes that motor equivalence strategies are developed on the basis of a model that takes into account the conditions under which the motor system moves and predicts the output of the system from the motor commands. The predicted output is compared with characteristics of the desired goal, and an error is computed. If the error is significant, as would be the case in a perturbation where a degree of freedom is inhibited, correction is provided using the predictive model. This principle is based on three important hypotheses: (1) a predictive model exists in the Central Nervous System; (2) the Central Nervous System can compute an error; (3) the predictive model is able to adjust the motor commands to reduce the error. The concept of “*internal model*” introduced by Kawato et al. (1987) in motor control research provided support for this principle on the basis of neurophysiological data (see for details Wolpert et al., 1998).

In an experiment presented by Kelso et al. (1984), a native speaker of English was asked to repeat the utterances “a /bæb/ again” and “a /bæz/ again” several times under normal and perturbed conditions. The perturbation was a sudden loading applied to the jaw during the closing gesture of the second consonant. This perturbation was applied randomly across repetitions of the utterances, in such a way that the subject could not predict when it would occur. When it occurred, the perturbation was systematically triggered at the same point of the upward trajectory of the jaw. The perturbation prevented the jaw from reaching the vertical position used in the absence of perturbation. In a first run of the experiment, the movements of the jaw and lips were recorded. In a second run, muscle activations were recorded with electromyography. The activation of the Orbicularis Oris Superior generates a lowering of the upper lip, that of Orbicularis Oris Inferior generates an elevation of the lower lip, and the posterior part of the Genioglossus enables the elevation of the tongue required for the production of /z/. In both runs, the acoustic signal was recorded.

Results were as follows. The subject was able to systematically compensate for the perturbation, and compensation started rapidly after the onset of the perturbation, i.e., with an average latency of 20ms.

Compensation was phoneme specific. A downward shift of the upper lip was observed for the bilabial stop /b/. No change in upper lip movement was noticed for the alveolar fricative /z/, instead there was a stronger activation of the Genioglossus. Overall, the results were compatible with the framework of coordinative structures: the very rapid compensation within a time interval shorter than the time required for feedback processed at a cortical level, supports the hypothesis of a functional coupling. The phoneme specific responses are evidence in support of the hypothesis that the functional coupling is oriented toward the achievement of a specific goal that is defined in terms of closure either at the lips, for /b/, or in the alveolar region, for /z/.

Another motor equivalence study that supports the concept of coordinative structures was conducted by Kelso & Tuller (1983). These authors ran a bite block experiment similar to the one of Gay et al. (1981) but with speakers of English and, more importantly, with a strong reduction in oro-sensory and auditory feedback. Oro-sensory feedback was reduced with anesthesia and auditory feedback was masked by a loud white noise (see section II.C for details). The results showed that subjects were able to compensate for bite block perturbations very rapidly, and with a similar accuracy as reported in Gay et al. (1981). Kelso & Tuller (1983) argued that there is no way to compute an error between the current state of the motor system and the desired goal for the Central Nervous System in the absence of feedback. Thus, the capacity to compensate in the absence of feedback contradicts the “prediction-error-correction” principle. Kelso & Tuller concluded that the very rapid compensatory responses support a functional coupling of the articulators.

From this selection of studies, it can be concluded that evidence in support of coordinative structures lies mainly in the great rapidity of the response to perturbation and in the efficiency of compensation in the absence of feedback. These two pieces of evidence were quite convincing at the time of the experiments. But this was before the concept of an *internal model* was introduced to the field of speech motor control (Bailly et al., 1991).

The internal model hypothesis assumes that two models of the motor system exist in the brain: a forward internal model that provides an approximation of the output of the motor system; and an inverse model that can find the motor commands appropriate to the achievement of the task on the basis of a computation of costs (for a tutorial in the context of speech production see Perrier, 2012). Once a forward model is assumed, the rapidity of response to perturbation and the ability to

compensate in the absence of feedback no longer provide evidence only in favor of coordinative structures. Once developed, a forward model provides the Central Nervous System with a reliable approximation of biological feedback, this is the so-called internal feedback. Internal feedback can be processed very rapidly (within 20ms) by the Central Nervous System, and the Central Nervous System can then rely on this, rather than on biological feedback, in ongoing control. Thus, an internal model can explain why compensations are possible in the absence of feedback.

Numerous experiments involving motor equivalence strategies in response to perturbations have provided support for the internal model hypothesis, especially when they have shown the existence of after-effects. After-effects correspond to the fact that compensation strategies developed during the production of speech with perturbation are still used for a time after the perturbation has been removed. After-effects suggest that new mechanisms have been learned and stored in the brain (Shadmehr & Mussa-Ivaldi, 1994). Whether these mechanisms are linked with motor memory or a generalized model of the motor system is still a debated issue (Conditt et al., 1997, Purcell & Munhall, 2006b). However, they are often interpreted as evidence that new internal models are acquired when speech production conditions change (Houde & Jordan, 1998). Internal models can be in the form of coordinative structures, but they can also be more complex and deliver internal feedback. After-effects were found in force field perturbations (Tremblay et al, 2003) and in auditory feedback perturbations (Houde & Jordan, 1998, Purcell & Munhall, 2006b, Cai et al., 2010). There is still uncertainty over the exact nature of the learning that takes place, whether it involves a simple motor memory, a complete generalized model of the speech production system, or something in between. Depending on the nature of the learning, it is possible to assume or to contest that the Central Nervous System is able to compensate accurately and rapidly in response to perturbations using a “prediction-error-correction” principle. Current experimental studies of motor equivalence mechanisms in the presence of perturbations are mainly focused on investigating the nature and complexity of the representations of the speech motor system that are stored in the brain.

IV. Conclusion

Motor equivalence, the capacity to achieve the same motor task in different ways, is an exceptional capacity of the motor system. It allows flexibility in general and adaptability under certain

circumstances. Since the physical conditions in which a certain motor task has to be fulfilled can vary, different solutions may be possible. These solutions depend on both speaker-specific morphology and biomechanics in human speech production, and on the inner and outer conditions that define the effort that must be expended to solve the task. Most studies on motor equivalence in speech production use short term perturbation paradigms. This methodology allows investigation of the goals of speech, a fundamental concept that is crucial to the understanding of speech motor control. The results from perturbation experiments strongly suggest that speech goals are specified in both the articulatory/motor and the acoustic/auditory domains. Many results also suggest the pre-eminence of the acoustic/auditory goal when a perturbation makes reaching both articulatory and acoustic goals impossible. Current work aims to investigate and clarify the proposed pre-eminence of acoustic/auditory goals over articulatory/somatosensory goals.

Motor equivalence also contributes to the generation of hypotheses regarding neural principles that underlie speech motor control, such as the coupling of coordinative structures or the prediction-correction-error principle (internal model). Motor equivalence may to a large extent reflect motor learning. Current works using perturbations of the production-perception loop in speech communication is being used to investigate the nature of the learnt representations of the speech production system in the brain.

V. References

- Bailly, Gérard, Rafael Laboissière & Jean-Luc Schwartz. 1991. Formant trajectories as audible gestures: An alternative for speech synthesis. *Journal of Phonetics*, 19. 9-23.
- Bernstein, Nikolai A. 1967. *The Coordination and Regulation of Movements*. London, UK: Pergamon Press.
- Brunner, Jana, Phil Hoole & Pascal Perrier. 2011. Adaptation strategies in perturbed /s/. *Clinical Linguistics & Phonetics* 25(8). 705-724.

Conditt, Michael A., Francesca Gandolfo & Ferdinando A. Mussa-Ivaldi. 1997. The motor system does not learn the dynamics of the arm by rote memorization of past experience. *Journal of Neurophysiology* 78. 554-560.

Fant, Gunnar. 1960. *Acoustic Theory of Speech Production*. Mouton, The Hague: The Netherlands

Feng, Yongqiang, Vincent L. Gracco & Ludo Max. 2011. Integration of auditory and somatosensory error signals in the neural control of speech movements. *Journal of Neurophysiology* 106(2). 667-679.

Folkins, John W. & James H. Abbs. 1975. Lip and jaw motor control during speech: Responses to resistive loading of the jaw. *Journal of Speech and Hearing Research* 18. 207-220.

Fowler, Carol .A. 1986. An event approach of the study of speech perception from a direct-realist perspective. *Journal of Phonetics* 14. 3-28.

Fujimura, Osamu. 1986. Relative invariance of articulatory movements: An iceberg model. In Perkell, Joseph S. & Dennis H. Klatt (eds.), *Invariance and Variability in Speech Processes*, pp. 226-242. Hillsdale, NJ., USA: Lawrence Erlbaum Associates.

Gay, Thomas, Björn Lindblom & James Lubker. 1981. Production of bite-block vowels: Acoustic equivalence by selective compensation. *Journal of the Acoustical Society of America* 69(3). 802-810.

Gomi, Hiroaki, Masaaki Honda, Takayuki Ito & Emi Z. Murano. 2002. Compensatory articulation during bilabial fricative production by regulating muscle stiffness. *Journal of Phonetics* 30. 261-279.

Gracco, Vincent L. & James H. Abbs. 1985. Dynamic control of the perioral system during speech: Kinematic analyses of autogenic and nonautogenic sensorimotor processes. *Journal of Neurophysiology* 54(2). 418-432.

Honda, Masaaki, Akinori Fujino & Tokihiko Kaburagi. 2002. Compensatory responses of articulators to unexpected perturbation of the palate shape. *Journal of Phonetics* 30. 281-302.

Houde, John F. & Michael I. Jordan. 1998. Sensorimotor adaptation in speech production. *Science* 279. 1213-1216.

Hughes, Marie Olive & James A. Abbs. 1976. Labial-mandibular coordination in the production of speech: Implications for the operation of motor equivalence. *Phonetica* 33(3). 119-221.

- Ito, Takayuki & David J. Ostry. 2010. Somatosensory contribution to motor learning due to facial skin deformation. *Journal of Neurophysiology* 104. 1230-1238.
- Jones, Jeffery A. & Kevin G. Munhall. 2002. The role of auditory feedback during phonation: Studies of Mandarin tone production. *Journal of Phonetics* 30. 303-320.
- Jones, Jeffery A. & Kevin G. Munhall. 2003. Learning to produce speech with an altered vocal tract: The role of auditory feedback. *Journal of the Acoustical Society of America* 113(1). 532-543.
- Kawato, Mitsu, Kazunori Furukawa & Ryoji Suzuki. 1987. A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics* 57. 169-185.
- Kelso, J.A. Scott & Betty Tuller. 1983. "Compensatory articulation" under conditions of reduced afferent information: A dynamic formulation. *Journal of Speech and Hearing Research* 26. 217-224.
- Kelso, J.A. Scott, Betty Tuller, Eric Vatikiotis-Bateson & Carol A. Fowler. 1984. Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance* 10(6). 812-832.
- Kewley-Port, Diane. 2001. Vowel formant discrimination II: Effects of stimulus uncertainty, consonantal context, and training. *Journal of the Acoustical Society of America* 110(4). 2141–2155.
- Lametti, Daniel, R., Sazzad M. Nasir & David J; Ostry. 2012. Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of Neuroscience* 32(27). 9351-9358
- Lindblom, Björn & Johan E.F. Sundberg. 1971. Acoustic consequences of lip, tongue, jaw and larynx movement. *Journal of the Acoustical Society of America* 50(4). 1166-1179.
- Lindblom, Björn, James Lubker & Thomas Gay. 1979. Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics* 7. 147-161.
- Maeda, Shinji. 1990. Compensatory articulation during speech; Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In Hardcastle, William J. & Alain Marchal

(eds.), *Speech Production and Speech Modeling*, pp. 131-149. Dordrecht, The Netherlands: Kluwer Academic Publishers.

McFarland, David H., Shari R. Baum & Caroline Chabot. 1996. Speech compensation to structural modifications of the oral cavity? *Journal of the Acoustical Society of America* 100(2). 1093-1104.

Perkell, Joseph S., Melanie L. Matthies, Mario A. Svirsky & Michael I. Jordan. 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(5). 2948-2961.

Perrier, Pascal. 2012. Gesture planning integrating knowledge of the motor plant's dynamics: A literature review from motor control and speech motor control. In *Speech Planning and Dynamics*, Susanne Fuchs, Melanie Weirich, Daniel Pape & Pascal Perrier (eds.), pp. 191-238. Frankfurt am Main, Germany: Peter Lang Publishers.

Purcell, David W. & Kevin G. Munhall. 2006a. Compensation following real-time manipulation of formants in isolated vowels. *Journal of the Acoustical Society of America* 119(4). 2288-2297.

Purcell, David W. & Kevin G. Munhall. 2006b. Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *Journal of the Acoustical Society of America* 120(2). 966-977.

Saltzman, Elliot. 1986. Task dynamic coordination of the speech articulators: A preliminary model. *Experimental Brain Research Series* 15. 129-144.

Saltzman, Elliot. 1991. The task dynamic model in speech production. In Peters, Herman F.M., Wouter Hulstijn & C. Woodruff Starkweather (eds.), *Speech Motor Control and Stuttering*, pp. 37-52. Amsterdam, The Netherlands: Elsevier Science Publishers B.V.

Saltzman, Elliot, Masayoshi Kubo & Cheng-Chi Tsao. 2006. Controlled variables, the uncontrolled manifold method and the task-dynamic model of speech production. In Divenyi, Pierre, Steven Greenberg & Georg Meyer (eds.), *Dynamics of Speech Production and Perception*, 21-31. Amsterdam, The Netherlands: IOS Press.

Savariaux, Christophe, Pascal Perrier & Jean-Pierre Orliaguet. 1995. Compensation strategies for a lip-tube perturbation of the rounded vowel [u]. *Journal of the Acoustical Society of America* 98(5). 2428-2442.

- Savariaux, Christophe, Pascal Perrier, Jean-Pierre Oriaguet. & Jean-Luc Schwartz. 1999. Compensation strategies for the perturbation of French [u] using a lip tube. II. Perceptual analysis. *Journal of the Acoustical Society of America* 106(1), 381-393.
- Scholz, John P. & Gregor Schöner. 1999. The uncontrolled manifold concept: Identifying control variables for a functional task. *Experimental Brain Research* 126. 289-306.
- Shadle, Christine H. 1989. Articulatory-acoustic relationships in fricative consonants. In Hardcastle William J. & Alain Marchal (eds.), *Speech Production and Speech Modeling*, 211-240. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Shadmehr, Reza & Ferdinando A. Mussa-Ivaldi. 1994. Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience* 14(5). 3208-3224.
- Shiller, Douglas M., Marc Sato, Vincent L. Gracco & Shari B. Baum. 2009. Perceptual recalibration of speech sounds following speech motor learning. *Journal of the Acoustical Society of America* 125(2). 1103-1113.
- Stevens, Kenneth N.. 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. In David, Edward E., Jr. & Peter B. Denes (eds.), *Human Communication: A Unified View*, 51-66. New York, USA: Mc Graw Hill.
- Summers, Walter van, David B. Pisoni, Robert H. Bemacki, Robert I. Pedlow & Michael I. Stokes. 1988. Effects of noise on speech production: Acoustic and perceptual analyses. *Journal of the Acoustical Society of America* 84(3). 917-928.
- Tourville, Jason A., Kevin J. Reilly & Frank H. Guenther. 2006. Neural mechanisms underlying auditory feedback control of speech. *NeuroImage* 39, 1429-1443.
- Tremblay, Stéphanie, Douglas M. Shiller & David J. Ostry. 2003. Somatosensory basis of speech production. *Nature* 423. 866-869.
- Wolpert, Daniel.M., R. Chris Miall & Mitsuo Kawato. 1998. Internal models in the cerebellum. *Trends in Cognitive Sciences* 2(9). 338-347.

VI. Further Reading

Grimme, Britta, Susanne Fuchs, Pascal Perrier & Gregor Schöner. 2011. Limb versus speech motor control: A conceptual review. *Motor control* 15(1). 5-33, provides a comparison of methods and issues, including motor equivalence, in limb motor control and speech motor control research.

Guenther, Frank H., Michelle Hampson & Dave Johnson. 1998. A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review* 105(4). 611-633, presents an extensive modeling work addressing the issues of speech goals, internal models and motor equivalence.

Latash, Mark L., Mindy F. Levin, John P. Scholz & Gregor Schöner. 2010. Motor control theories and their applications. *Medicina (Kaunas)* 46(6). 382-392, presents a theoretical framework to study motor control issues, including motor equivalence, without internal models and using a combination of theories, namely the Uncontrolled Manifold concept and the Equilibrium-Point-Hypothesis.

Munhall, Kevin. G., Ewen G. MacDonald, S. K. Byrne & I. Johnsrude. 2009. Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *Journal of the Acoustical Society of America* 125(1). 384-390, presents experimental results showing an automaticity of the compensation in the auditory domain when auditory feedback is altered.

Rochet-Capellan, Amélie, Lara Richer & David J. Ostry. 2012. Non-homogeneous transfer reveals specificity in speech motor learning. *Journal of Neurophysiology* 107, 1711–1717, presents the results of experiments using auditory feedback perturbations, which suggest that for speech production the relation between motor commands and spectral patterns is learnt locally in the acoustic space.

Sakata, Jon T. & Michael S. Brainard. 2009. Social context rapidly modulates the influence of auditory feedback on avian vocal motor control. *Journal of Neurophysiology* 102(4). 2485-2497, shows that, in the presence of auditory feedback perturbation birds, when they are singing, develop compensatory strategies, and the extent of the compensation is modulated by the social context of the song production.

Tremblay, Stéphanie, Guillaume Houle & David J. Ostry. 2008. Specificity of speech motor learning, *Journal of Neuroscience* 28(10). 2426-2434, presents the results of force field perturbation experiments that suggest for speech production that the Central Nervous System learns task specific, instead of generalized, representations of the relations between motor commands and movements.

Vahdat, Shahabeddin, Mohammad Darainy, Theodore E. Milner & David J. Ostry. 2011. Functionally specific changes in resting-state sensorimotor networks after motor learning. *Journal of Neuroscience* 31(47). 16907-16915, presents the results of an fMRI study in which a network of brain regions has been found, which is active during motor perturbation experiments and can be associated with motor learning.

VII. Key-words

Speech production, speech motor control, motor equivalence, motor goals, speech coordination, motor perturbations, auditory feedback perturbations

VIII. Authors biographies

Pascal Perrier was awarded a PhD from Grenoble University. He is a Professor at Grenoble INP/Gipsa-Lab. His research interests include speech motor control and orofacial biomechanics. He worked with the ZAS/Berlin, the Speech Group at MIT/Cambridge, and the JAIST in Ishikawa/Japan. He is the editor of the book series “Speech Production and Perception” by Peter-Lang Publisher

Susanne Fuchs was awarded a PhD from QMUC, Edinburgh. Currently, she is working at ZAS in Berlin. She is interested in experimental phonetics and speech motor control, using a variety of methods (Electrophalography, articulography; intraoral pressure measures, Respitrace). She is the editor of the book series “Speech Production and Perception” by Peter-Lang-Publisher