

On the complex nature of speech kinematics

Susanne Fuchs

Zentrum für Allgemeine Sprachwissenschaft, Berlin, Germany

Institut de la Communication Parlée, UMR CNRS 5009, Institut National Polytechnique de Grenoble & Université Stendhal, Grenoble, France

Pascal Perrier

Institut de la Communication Parlée, UMR CNRS 5009, Institut National Polytechnique de Grenoble & Université Stendhal, Grenoble, France

Studying kinematic behavior in speech production is an indispensable and fruitful methodology in order to describe for instance phonemic contrasts, allophonic variations, prosodic effects in articulatory movements. More intriguingly, it is also interpreted with respect to its underlying control mechanisms. Several interpretations have been borrowed from motor control studies of arm, eye, and limb movements. They do either explain kinematics with respect to a fine tuned control by the Central Nervous System (CNS) or they take into account a combination of influences arising from motor control strategies at the CNS level and from the complex physical properties of the peripheral speech apparatus. We assume that the latter is more realistic and ecological. The aims of this article are: first, to show, via a literature review related to the so called '1/3 power law' in human arm motor control, that this debate is of first importance in human motor control research in general. Second, to study a number of speech specific examples offering a fruitful framework to address this issue. However, it is also suggested that speech motor control differs from general motor control principles in the sense that it uses specific physical properties such as vocal tract limitations, aerodynamics and biomechanics in order to produce the relevant sounds. Third, experimental and modelling results are described supporting the idea that the three properties are crucial in shaping speech kinematics for selected speech phenomena. Hence, caution should be taken when interpreting kinematic results based on experimental data alone.

1 Studying kinematic behaviour: Evidence from experiments and models

Based on a number of previous articles in the motor control and speech production literature, this article intends to show the complex nature of speech

kinematics and the difficulty of its interpretation in terms of speech motor control. Interpretations discussed here either explain kinematics with respect to fine tuned control by the Central Nervous System (CNS) or they take into account a combination of influences arising from motor control strategies at the CNS level and from the complex physical properties of the peripheral motor system. We hypothesize that the latter is more realistic and ecological. We also suggest that speech motor control makes use of specific physical properties of the speech production apparatus to achieve its goals. These physical properties are vocal tract limitations, aerodynamics, and biomechanics. By means of experimental and modelling results evidence will be provided that aerodynamics, vocal tract limitations, and biomechanics¹ play a crucial role and shape the kinematics of speech.

1.1 Controversies in motor control: Examples investigating the 1/3 power law

In this section we will mainly focus on the ‘1/3 power law’ in order to show by means of a well-known characteristic how kinematic behaviour has been interpreted in very different ways with respect to its underlying motor control mechanism.

The 1/3 power law has been extensively described by Viviani and colleagues (e.g. Viviani & Terzuolo 1982, Viviani & Schneider 1991, Viviani & Flash 1995) based on experimental data of subjects tracing or perceiving planar movements. Evidence has been given that there is a relationship between the degree of curvature and the speed of movement, with slower movements in the more curved parts and faster movements in the less curved parts of the trajectory. The power law can be described by the formula:

$$V(t) = k * R(t)^{\beta}$$

where V is the tangential velocity, R is the radius of the curvature, k is a velocity gain factor, and β has been estimated from experimental data as being close to 1/3. Therefore, the rule was called the 1/3 power law (1/3 corresponds to the movement in a Cartesian coordinate system – if movements are described in an angular space than it becomes the 2/3 power law).

Viviani and colleagues have not only shown that this law systematically applies for planar human arm movements, they also demonstrated that it deeply influences the perception of these movements. Indeed, if artificial movements do

¹ Biomechanics as such are not a peculiarity of speech motor control, they are also found in limb systems. However, biomechanics are included here since the tongue is a muscular hydrostat and has very specific biomechanical properties.

not respect this law, human subjects will not perceive them correctly. Since the power law can be discussed with respect to the action-perception interaction in speech (although it has not been intensively studied for speech), it is worth asking whether it is inherent to the human motor system or whether it is specifically controlled by the Central Nervous System (CNS).

Viviani and Flash (1995) defended the hypothesis that this rule is used by the CNS in order to optimise gestures (jerk minimization, see below), a strategy to limit the excess of degrees of freedom. In addition, these authors rejected the hypothesis that the $1/3$ power law could be explained by muscle properties and/or muscle dynamics since it also holds under isometric conditions (Massey et al. 1992), where the subjects were asked to draw planar patterns by grasping a 3D handle and pushing on it without moving their arms.

However, Gribble and Ostry (1996) noted that even under isometric conditions, muscle properties can affect force development and, additionally, that under these conditions it is difficult to separate between centrally planned strategies and effects due to the periphery. They suggested that muscle properties can account for the $1/3$ power law. Thus, using a biomechanical planar arm model they simulated elliptical tracings of arm movement. The model included the shoulder and the elbow joints and is generally controlled by means of shifting equilibrium points (Feldman 1986, Feldman et al. 1990). Simulations were run with control signals specifying an equilibrium shift at a constant tangential velocity (i.e. no relation between curvature and speed existed in the modelled control signals). They also simulated arm trajectories under an isometric condition and circular movements with control signals of constant speed and radius of curvature. In all the different conditions Gribble and Ostry found the simulated movements respected the $1/3$ power law relationship. The authors propose that *“muscle properties and dynamics can play a significant role in the emergence of this relationship”* (p. 2859).

Lebedev, Tsui and Gelder (2001) tried to explain the $1/3$ power law by means of the principle of least action, a principle known in theoretical physics. They report: *“From the principle of least action it follows that the CNS does not impose the power law directly, but follows the strategy of accomplishing the desired goal in a preset time with the minimum mechanical work required”* (p. 50). From their point of view the CNS follows a strategy minimizing the amount of mechanical work. A similar principle - although not related to the $1/3$ power law - was introduced to speech by Nelson (1983). He pointed out that the kinematic movement of speech articulators would be the result of a centrally controlled optimisation process, aiming at minimising the jerk, the third derivative of displacement over time. However, he also noted that the resulting velocity profiles are similarly bell-shaped as in a simple undamped linear mass-spring system with constant stiffness.

Another explanation for the 1/3 power law has been given by Harris and Wolpert (1998) who suggested that it would be the result of a strategy minimizing positional variability due to signal dependent neuronal noise which is neurobiologically more plausible than a centrally planned strategy minimizing the jerk.

The power law has been found in movement production and to be a determinant in perception (for speech production first results are reported by Tasko and Westbury 2004).

To summarise, the 1/3 power law describes the kinematic relation between speed of movement and degree of curvature for different human motor systems and it is integrated in human perception processes. Interpretations of the law are manifold. On the one hand researchers tried to show that this law can be explained by a centrally planned mechanism, e.g. minimizing the jerk, which would involve a complex internal representation of the motor system in the CNS. On the other hand, it has been shown that the 1/3 power law may purely be the consequence of muscle properties and dynamics and there is no need for a complex control of the phenomenon. Similar controversies have also been observed with respect to articulatory movements in speech production, which have been described by means of central control strategies or by means of specific characteristics of the speech production apparatus.

1.2 Controversies in speech motor control: Kinematic variations due to speech rate and loudness differences, and their underlying control

One of the most cited references investigating speech rate effects on kinematics is Adams, Weismer and Kent (1993). The authors recorded 5 speakers by means of the x-ray microbeam system producing a single sentence several times at 5 different speech rates. Changes in speech rate had a different impact on movement duration for opening and closing gestures, and for lower lip and tongue movements. Additionally, the number of velocity peaks as well as parameter c , an index of the velocity profile's geometry ($c=V_{max}/V_{mean}$), increased with decreasing speech rate and measures of the symmetry of the velocity profile changed across speaking rate. Adams, Weismer and Kent discussed these kinematic results with respect to the following motor control principles suggested in the literature:

- (1) Opening and closing gestures are differently controlled and reflect different muscle synergies (based on Gracco's kinematic results and discussion in 1988).
- (2) Asymmetries of the velocity profile due to changes in speech rate are a consequence of feedback mechanisms providing spatial information about the articulator (based on Bullock and Grossberg's VITE model 1988).

With decreasing speech rate the potential role of this feedback mechanism increases and would result in a more asymmetrical velocity profile. Changes in symmetry may also reflect a shift of the motor control system from an open-looped control (=without using feedback information) to a closed-loop control (=using feedback information) with slower movements.

(3) In agreement with Wiencke, Janssen and Belderboss (1987) the greater number of velocity peaks at slower rates may be a universal mechanism of speech motor control. Adams, Weismer and Kent also suggested in accordance with Milner and Ijaz' (1989) findings for hand movements that multiple peaks could originate from overlapping submovements, since it may be difficult to generate longer movements with one motor command only.

Thus, these authors clearly interpreted kinematic changes due to speech rate as a consequence of active control mechanisms from the CNS.

Although some of the hypotheses mentioned here may be true, we would like to note the following two points: First, most people who compared opening and closing gestures in repetitive CV-syllables (including Gracco 1988) used oral stops in their speech material. The main goal in oral stop production is to produce an airtight seal for the oral closure and a following perceptually salient burst. It has been hypothesized that the articulatory movement is planned towards a target located above the actual vocal tract limit (e.g. for bilabials see Löfqvist and Gracco 1997, for alveolars see Fuchs et al. in press). In terms of stability and simplicity, such a control strategy seems to be extremely efficient in comparison with the control of a fine positioning. The impact of the lower lip on the upper for bilabials or of the tongue at the palate for alveolars is likely to influence the velocity profile of the closing gesture, but has no or less influence on the opening gestures.

Second, double velocity peaks can also occur for instance in a single /y/-/u/ movement without any underlying submovements. Payan and Perrier (1997) found that the origin of such a double peaked pattern is due to muscle anatomy inducing a certain time sequencing of the activation/deactivation of the Styloglossus and the Genioglossus posterior muscles. Therefore, any inference from kinematics alone about the underlying motor control mechanisms should be considered with caution.

McClellan and Clay (1995) studied the relation between lower lip kinematics and their underlying single motor unit activity by means of EMG. They proposed that at least three different mechanisms may contribute to an increase in movement velocity: changes of the rate of firing motor units, changes in motor unit recruitment, and changes of the stiffness of the relevant articulator. Their aim was to observe the first two - firing rate and recruitment patterns of

single lower lip motor units simultaneously with the corresponding lip kinematics under varying speech rate conditions and phonetic structure. Three subjects were recorded by means of EMG of the following muscles: orbicularis oris inferior (OOI, active during lip closing), depressor labii inferior (DLI, active during lip opening) and mentalis (MENT, active during lip closing). The speech material consisted of repetitive CV syllables with C being /p/, /v/ or /f/ and V being /æ/ or /ʌ/. The firing rate was defined as the number of spikes per second, and spikes have been determined operationally by means of a threshold criterion. Kinematic results for lower lip movements at a higher speech rate exhibit differences with respect to opening and closing gestures. In general, closing gestures showed a significant increase in velocity whereas opening gestures did not. However, in the opening gestures differences in the average number of spikes per syllable were observed for the DLI in dependence of the vowel context. The average number of spikes were positively correlated with the amplitude of the velocity peaks in the kinematic signal. A similar correlation could not be found for the closing gestures. The authors suggest that an increase in speech rate from very slow to fast is associated with an increase of the firing rate of single motor units. According to these authors and contrary to Adams et al.'s (1993) suggestion, increase in speech rate is produced without changing the control strategy from multiple submovements at a slow rate to a unique movement at a fast rate, at least for opening gestures. According to McClean (personal communication), this finding sheds also new light on the interpretation of the variation of the number of velocity peaks with speech rate as observed by Adams et al. (1993): it is related to the motor units firing rate that determines the overlap between the successive parts of the movement associated with each motor unit activation.

In a follow-up study McClean and Tasko (2003) proposed again: *“Although the relationship between neural input to motoneurons and kinematics is extremely complex, kinematic analysis can provide a partial window to the neural processes underlying speech production”* (p. 1388). They investigated average lower lip and jaw muscle activities (mentalis MENT, depressor labii inferior DLI, masseter MAS (jaw opener)), anterior belly of digastric ABD (jaw opener) by means of broad-field EMG recordings simultaneously with kinematic data by means of EMA. Speech rate varied in 5 conditions and intensity in 2 different levels. Results for variations of loudness showed a strong positive correlation between muscle activation level with mean movement speed and movement distance. Concerning variations in speech rate similar results could not be found. Most consistently across speakers a general negative correlation of muscle activation levels with movement duration has been detected. In summary, observations of kinematics with varying speech rate are manifold. Authors often discuss kinematic results with respect to concepts from motor

control in general. One of the most cited references has been mentioned, discussing lip and tongue kinematics with respect to different underlying control strategies counting for variations in speech rate (Adams et al. 1993).

There is little experimental evidence in the literature directly linking kinematics of speech movements and their underlying motor unit/s activity. McClean and Clay (1995) and McClean and Tasko (2003) investigated variations of speech rate and loudness by simultaneously recording lip and jaw movements and lip and jaw muscle activity. Parts of their findings provide evidence on the link between kinematics and muscle activity and demonstrate that no particular control strategy is necessary when switching from slow to fast (for opening gestures in single motor units McClean & Clay 1995) or from normal to loud speech (in multiple motor units McClean and Tasko 2003). In order to further understand speech motor control mechanisms it seems therefore indispensable either to directly investigate kinematics and muscle activity, or/and to compare experimental data with simulations using different motor control models generating muscle activation. Other appropriate methodologies would be to perturb the motor system and compare compensatory movements with the unperturbed condition or to compare normal and pathological speech. The two latter will not be taken into account here.

1.3 The implementation of motor control models in speech: Some notes, our assumptions and methodology

Speech production involves the precise control of fast articulatory actions in a task specific manner and is therefore characterised by neural and muscular activities. However, there are at least two particular properties of the speech production mechanism which seem to be speech specific in comparison to other human motor systems (e.g. arm or eye movements):

- (1) the most flexible articulator, the tongue, moves in a narrow space delimited by the palate, pharyngeal walls, teeth, cheeks and lips, and
- (2) expiratory air coming from the lungs, passing the glottis propagates through the vocal tract with certain characteristics depending on the changing vocal tract configurations and the corresponding changes in the perturbation of air.

The consequences of these two properties are challenging since the speech motor control system may integrate vocal tract limitations or certain aerodynamic information in the planning of sounds. Hoole et al. (1998) for instance, investigated the potential role of aerodynamics onto kinematics in order to explain the forward movements during oral closure in velar stop production (for further discussion on this topic see 2.3.2.). By means of EMA and intraoral pressure measurements they carried out an experiment where three speakers pronounced

the target words in an egressive and in an ingressive condition. Although forward movement for velars considerably decreased in the ingressive condition, it was not eliminated. The authors suggest “*that the motor planning system may be anticipating the aerodynamic forces and planning movement trajectories to take advantage of the direction and magnitude of the force vector*” (p.136).

We suppose that during speech acquisition the infant develops simple internal models (for a review on the simplicity versus complexity of internal representations, also called ‘internal models’ see Perrier in press). It establishes relations between (1) motor commands and perceptual outputs, (2) motor commands and proprioception (including the limits of tongue movement due to vocal tract boundaries), (3) motor commands and aerodynamics (in particular the magnitudes of subglottal pressure), and (4) biomechanical properties of the articulators such as mass, inertia, muscle force directions etc. These relations are mapped with respect to their auditory consequences and within the learning process integrated in planning sounds. Given the fast nature of some articulatory movements, which are sometimes below 50 ms, such internal models are required since this duration is below the minimum delay necessary for the cortex to monitor the ongoing speech act. We also assume that many kinematic patterns are consequences of the physical properties of the complex speech apparatus tuned by the CNS. They are not due to complex internal models that would precisely determine articulatory movements at each point of the trajectory. In this framework we want to test to what extent speech characteristics can be explained by simple control strategies and simple internal models. Our methodology is therefore to compare experimental results with results from simulations by means of a complex biomechanical model, and by controlling targets, not trajectories. Of particular interest for the current work is the hypothesis that vocal tract limitations, aerodynamics and biomechanics can affect the kinematics of articulatory movements. Their potential role will be discussed below.

2 Potential underlying factors of speech kinematics

Before providing examples on how biomechanics, aerodynamics and vocal tract limits can shape speech movements, we like to introduce briefly the model we used.

2.1 Introduction to the complex peripheral model

Since we assume a complex peripheral speech apparatus plus a simple control due to the CNS and our methodology is to compare experimental data with data from simulations of a model, a complex biomechanical tongue model has been built previously.

Elastic properties of tongue tissues are accounted for by finite element (FE) modelling. Muscles are modelled as force generators that (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 contractions within tongue tissues. Curves (see Figure 1) representing the contours of the lips, palate and pharynx in the midsagittal plane are added to specify the limits of the vocal tract.

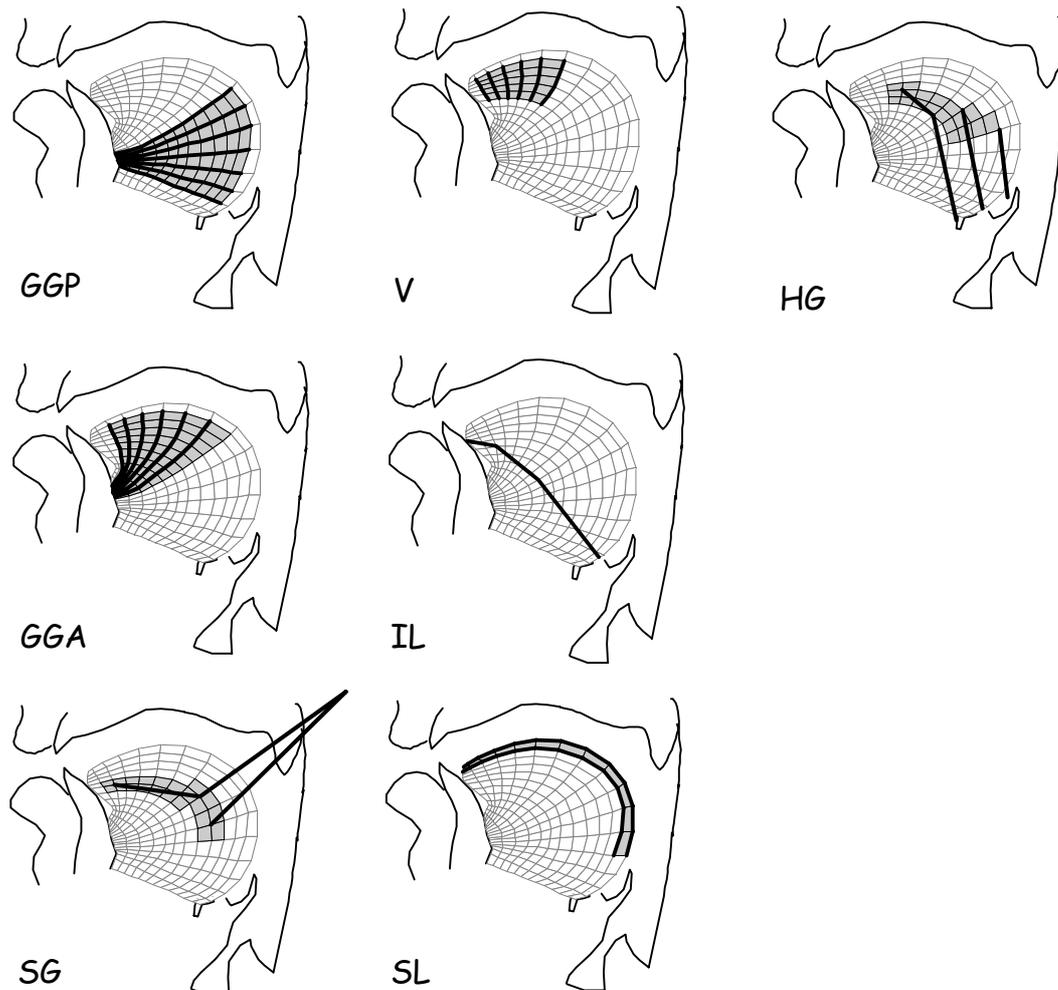


Figure 1: Defined tongue muscles of the FE structure, black lines correspond to macrofibres: GGP (genioglossus posterior), GGA (genioglossus anterior), SG (styloglossus), V (verticalis), IL (inferior longitudinalis), SL (superior longitudinalis), HG (hyoglossus)

The jaw and the hyoid bone are represented in this plane by static rigid structures to which the tongue is attached. Changes in jaw height can be simulated through a single parameter that modifies the vertical position of the whole FE structure. The model is controlled according to the λ model (Feldman 1986) that specifies for each muscle a threshold length, λ , where active force

starts. If the muscle length is larger than λ , muscle force increases exponentially with the difference between the two lengths. Otherwise there is no active force. Hence, muscle forces are typically non-linear functions of muscle lengths. The control space is called the λ space (for more details see Payan & Perrier 1997, Perrier et al. 2003). Muscle forces are applied to the FE structures via macrofibers that specify the insertions and the main force directions of each muscle (see black lines in Figure 1).

The modelling of the fluid-wall interaction and the action of the tongue at the palate are integrated in the complex biomechanical model. The first implies at each time step the specification of: (1) the area function from the sagittal distances generated by the 2D tongue model and their coronal correspondence (see Perrier et al. 1992), and (2) the volume velocity of the airflow through the vocal tract. These steps are followed by the computation of: (1) the distribution of pressure within the vocal tract, and (2) the pressure forces at each node of the tongue model, which are then added to the muscle forces to calculate the global forces shaping the tongue.

The area function is computed using an adapted version of the original $\alpha\beta$ -model (Heinz and Stevens 1965), where $\beta=1.5$. In order to provide realistic vocal tract cross-sectional areas, α varies from the glottis to the lips according to a division of the vocal tract into 7 sections and to the value of the sagittal distance (Perrier et al. 1992).

Flow velocity and pressure distribution can be calculated with a flow model (Pelorson et al. 1995) based on a simple 2D potential flow theory, accounting for viscous losses as a perturbation of the inviscid solution. In addition, flow separation effects within a constriction are taken into account. For the sake of simplicity, the flow separation position is estimated as the point downstream of the constriction where the cross-sectional area is 20% larger than the minimum area of the constriction (Pelorson et al. 1995). The flow model is driven by a single parameter: the pressure difference $\Delta P = P_0 - P_{out}$, where P_0 and P_{out} are respectively the pressure past the glottis and at the lips.

The action of the tongue at the palate is modelled in two steps: (1) detection of contact of the tongue at the palate, and (2) generation of the resulting contact forces by means of the penalty method (Marhefka & Orin 1996). More specifically, if a node of the FE structure moves beyond the limit (palate) a repulsion force is generated to move this node back as a non-linear function of the penetration distance (for further details on the method, see Perrier et al. 2003).

The model is currently under development in order to improve it from two dimensions to three dimensions (Gerard et al. in press).

2.2 Aerodynamics

Precise articulatory movements on their own do NOT produce speech sounds, the propagation of air with a certain density and speed of particles through the vocal tract is required. The source (e.g. vocal fold vibrations) and the vocal tract interact with each other during the production of different sounds.

Experimental evidence from simultaneous aerodynamic and articulatory movement measures are rather rare. However, kinematic results have often been interpreted with respect to aerodynamics, especially in explaining phenomena like devoicing in oral stops or voicing/voicelessness in consonant clusters. The same holds true for aerodynamic results, where underlying articulatory movements have been inferred without any experimental evidence. One exception is for instance the study of Svirsky et al. (1997) who observed tongue displacement in relation to intraoral pressure changes. Both measurements were used to assess the validity of a tongue compliance model. Results for tongue displacement for the consonants in /aba/, /apa/, and /ama/, and differences in intraoral pressure were investigated to shed some light on the question whether cavity enlargement is an active or passive mechanism. Cavity enlargement has been discussed as one mechanism to sustain the transglottal pressure difference, necessary for the production of vocal fold vibrations during oral closure. It turned out that the magnitudes of peak tongue dorsum displacement recorded by means of EMA were significantly larger during the production of voiced bilabials compared to smaller magnitudes in voiceless bilabials, even though the intraoral pressure was higher for the voiceless. It seemed surprising that such tongue dorsum differences occurred during a bilabial when surrounded by the same unrounded vowel context. The displacement was close to zero in the sequence involving the nasal. Svirsky et al. reported: *“It is interesting to observe that the relatively sharp, fast downward tongue dorsum displacements during /apa/ or /aba/ were generally close to the rise in intraoral pressure”*(p.565). Using a lumped parameter circuit model Svirsky et al. estimated tongue compliance and found much higher values for the voiced stops than for the voiceless. However, relaxation of the tongue for the voiced stop did not explain all the results. Hence, the authors proposed a combination of intentional relaxation of tongue muscles with an active displacement of the tongue. For the voiceless they suggested an active stiffening process of the tongue. In the context we discuss here their findings also provide evidence that intraoral pressure, i.e. aerodynamics, can affect tongue displacement since differences do not occur in the nasal (Nasals obviously do not involve a high intraoral pressure, since the air can escape via the nasal cavity). However, a separation between changing tongue positioning due to pressure and/or due to an active mechanism

cannot be based on their results and the amount of change the pressure causes is rather speculative.

Another example for the potential influence of aerodynamics on tongue kinematics are looping patterns, which will be discussed in more details in section 2.3.2. (see the contribution of Hoole et al. 1998, Perrier et al. 2000a).

2.3 Biomechanics

Two different examples will be given here in order to point out the potential role of biomechanics for the speech motor control process and its implications for kinematics. The first one is related to the explanation of the limited degrees of freedom of tongue movements in vowel production², and the second to the explanation of the so called ‘looping patterns’ in velar stop production.

2.3.1 Degrees of freedom for tongue movements during speech production

The production of speech requires the simultaneous control of at least thirty different muscles. However, at the same time classical articulatory descriptions of vowel production are limited to a small number of parameters such as high versus low, front versus back for the tongue, and rounded versus spread for the lips. Hence, the understanding of speech motor control requires a reduction of the dimensionality from the muscle control space to a more functional, speech-related control space. The functional, speech-related control space will hereafter be called the degrees of freedom of the vocal tract. The reduction in dimensionality is a desired aim since it allows generalisations independent of speakers’ specific mechanisms. Even more broadly, it might show to what extent specific muscles are coordinated to produce meaningful sounds in the different languages. Previous work on the degrees of freedom were mainly based on statistic analyses of kinematic data. Harshman et al. (1977), Jackson (1988), Nix et al. (1996) and Hoole (1999) applied a PARAFAC analysis to x-ray or EMA data for English, Icelandic, and German. Maeda (1990) ran a guided principal component analysis (PCA) on x-ray data of French. Sanguineti and colleagues (Sanguineti et al. 1997, 1998) used the same corpus as Maeda, but additionally they provided a projection of the data set in a modelled muscle space by means of a biomechanical model of the tongue, jaw and hyoid bone. These authors were not only able to present a reduction in dimensionality, but

² It should be noted that parts of this work (but with different simulations) have been presented at ICSLP Beijing, see Perrier et al. (2000b).

also a description of the muscular correlates of the degrees of freedom during vowel production.

Although four different languages were analysed in these studies, most of the results presented show that more than 90% of the variance observed in the tongue shapes can be ascribed along two main degrees of freedom: (1) a movement of the tongue body along a high-front to low-back axis (called 'front raising' in Harshman et al. 1977) and (2) a bunching of the tongue along a high-back to low-front axis (called 'back raising' in Harshman et al., 1977). Jackson (1988) found that the number of degrees of freedom were language specific, i.e. different for English and Icelandic. However, his PARAFAC analysis was then proved to be degenerate by Nix et al. (1996), who reanalysed the same data set.

The results of these studies lead to questions about the origin of the two main degrees of freedom: are they learned, speech-specific actions, or are they due to basic properties of the speech production mechanism? In the following we will explore the hypothesis that the two main degrees of freedom have their origin in the anatomical and biomechanical properties of the speech production apparatus. Toward this aim, the bio-mechanical model of the tongue was used to generate a large set of tongue configurations, on which a PCA was ran in order to extract the main axes of deformation.

First results were presented in Perrier et al. (2000b). They were based on a gaussian sampling of the motor control space with the commands around the rest position as an average vector. These simulations were limited to the analysis of tongue configurations during vowel production, excluding those which were too close to the palate. In this paper we propose an extension of the previous work, covering a very broad range of tongue shapes. We adopted a uniform sampling method and included tongue configurations in slight contacts with the palate. In doing so our simulations cover the whole range of tongue shapes that can be generated by the model. Thus, 9000 tongue configurations were simulated and analysed with the classical PCA procedure (see Perrier et al. 2000b for details). The results of the PCA are depicted in figure 2 for a variation of +/-1 standard deviation around the mean value along each of the principal axes. The first and second factors clearly correspond to the typical front and back raising patterns. The third factor can be associated with a vertical downward movement of the tongue body and results for the fourth factor are rather marginal.

In the majority of the studies based on statistical analyses of articulatory data, more than 90% of the variance observed for a subject were described by the first two factors, while in our study 3 factors are necessary to reach approximately the same level of description. Results are as follows: the first factor explains 69 % of the variance, the first two factors 88 %, the first three factors 96 % and the first four factors 99 %. The slightly greater number of factors is in agreement with Nix et al.'s (1996) findings, which showed that

when the tongue shapes of 6 speakers were analysed together, 4 factors were necessary to reach the same level of description in comparison to the 2 factors extracted from the data of a single subject. Since our data were generated from a variety of random muscle commands relevant for vowel production, they may be more general, analogous to the combined data from 6 speakers.

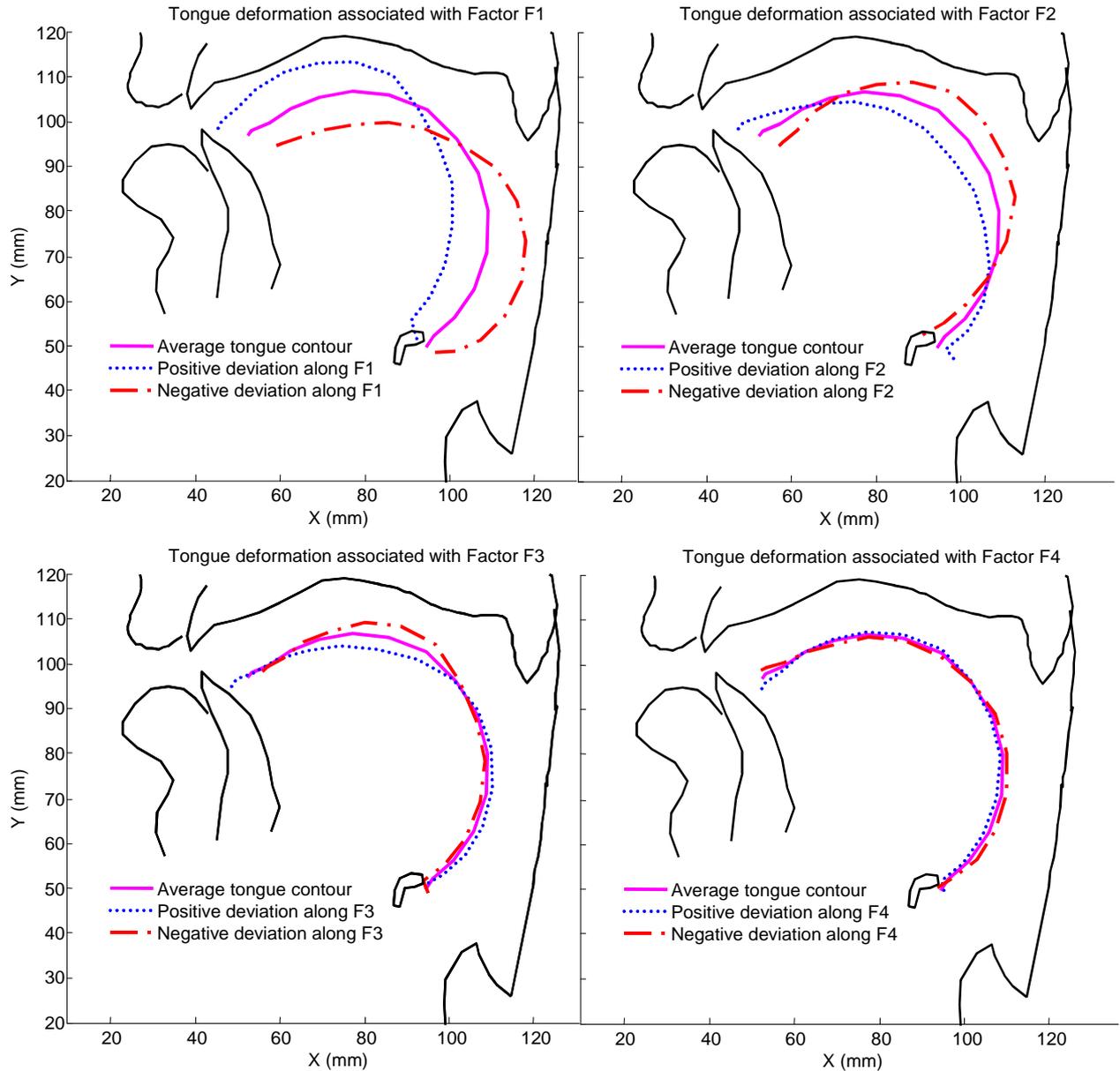


Figure 2: Tongue deformations based on a PCA for the first four factors (from upper left to lower right), solid line: average contours, dotted lines: positive deviations from the average, dashed-dotted lines: negative deviations from the average; for further details see text

We conclude that the degrees of freedom in vowel production extracted from our simulations for French, and found in several studies for German, Icelandic, and English are due to the anatomical and biomechanical properties of the tongue and therefore not language-specific. Speech motor control uses these degrees of freedom to determine and differentiate speech articulations with respect to the various sounds of a language.

2.3.2 On looping patterns

In a series of papers (e.g. Houde 1968, Ohala 1983, Mooshammer et al. 1995, Hoole et al. 1998, Löfqvist and Gracco 2002, Geng et al. 2003, Perrier et al. 2000a, Perrier et al. 2003, Brunner et al. 2004, Brunner 2005) researchers were interested in explaining the striking movement trajectories occurring during velar stop production. The trajectories have been called ‘looping patterns’ since they are reminiscent of ellipses. Loops can be found during V1CV2-sequences with C being a velar stop. Depending on the surrounding vowel context with V1 being a back vowel and V2 a front vowel one would expect a forward movement during the oral closure, simply as a consequence of coarticulation. Such forward sliding movements are however also found for V1=V2 as for instance in /aka/ where one could assume comparable movements towards oral closure and back.

The explanations for the phenomenon are manifold: due to aerodynamics, biomechanics³ or cost minimization. Aerodynamics is in most cases mentioned, but for different reasons: Houde (1968) assumed that the forward movement of the tongue along the palate in a voiced velar stop would be due to the increased intraoral air pressure. Ohala (1983) attributed looping patterns to a strategy enlarging the oral cavity in order to maintain voicing for the voiced velar stop. Mooshammer et al. (1995) rejected this hypothesis since they found larger forward movements for the voiceless in comparison to the voiced stops. In order to test the impact of intraoral pressure onto tongue kinematics quantitatively, Hoole et al. (1998) observed looping patterns in normal and ingressive speech. Although they found smaller loops in ingressive speech, they were also directed forwards so that an increased intraoral pressure can not capture the whole phenomenon alone. Modeling work has been carried out by Perrier et al. (2000a) using a combination of a biomechanical model and an airflow model. They investigated looping patterns in low back and high front vowel contexts and found that biomechanics have a major impact on the kinematic patterns while aerodynamics play a negligible role when the velar stop is produced during low

³ Since our previous work on loops was mainly related to biomechanics (Perrier et al. 2003), we have included the example at this point. However, it could also be included at the section on aerodynamics or vocal tract limits.

back vowel context. For /aki/ and /iki/ sequences the authors mentioned comparable influences of biomechanics and aerodynamics on the loops. These patterns were sensitive to the onset of pressure rise in the closing gesture and to the amount of pressure. With an earlier onset of the pressure rise and with a higher pressure, larger movement amplitudes were simulated.

A totally different perspective explaining looping patterns has been given by Löfqvist and Gracco (2002). They state that neither aerodynamics nor biomechanics alone would account for the observed patterns. Hence they suggest that loops are a result of a general motor control principle - the cost minimization process. This principle is associated with holding the third derivative of the movement, the jerk, as small as possible; this corresponds to a general smoothing strategy (Hogan 1990). According to the cost minimization principle, the whole trajectory of the loops would be controlled by the CNS.

In contradiction to this idea, Perrier et al. (2003) simulated looping patterns by means of a biomechanical model without any cost minimization strategy: Consonants and vowels in VCV-sequences have been specified in terms of targets. The consonant was always /k/ and for the two surrounding vowels /a, i, u/ have been used. Based on the findings of their simulations they suggested that biomechanical properties of the tongue explain looping patterns for all sequences where the first vowel was /a/ or /u/, independent of the second vowel. When the first vowel was /i/ they found variable forward or backward loop patterns, depending on the position of the target specified for /i/. This finding was consistent with the variability of experimental data reported in the literature. Consequently, no central processes seem to be necessary to control the whole trajectory of these sequences.

A generally accepted explanation for the combination of underlying factors and their strength contributing to these kinematics is still missing. In order to further support the biomechanical explanation of looping patterns we assume that loops may not only be found in the production of velar stops in e.g. /a/-context, but also in any other movement directed to the velar region. Thus, a sequence such as /aua/ should also show looping patterns to a certain extent. We therefore simulated 3 different sequences: /aua/, /aka/ with the impact of the tongue at the palate included in the model, and /aka/ with no palate in place, i.e. no impact of the tongue at the palate. The muscle activation patterns are given in table 1.

In all cases (see figure 3) it can be observed that in the upper part of the trajectory slight forward movements occur. The size of the loop is clearly larger for the /aka/-sequence than for /aua/. This is consistent with Perrier et al.'s (2003) findings that the relative position of the consonant and first vowel target has an incidence on the size of the loop. Finally, the trajectory goes further back in the absence of the palate (compare /aua/ and /aka/ without the palate). This

movement is due to the major influence of the Styloglossus that pulls the tongue back high in the velar region.

Table 1: Muscle activation patterns for the three simulations: - no activation, + slight activation, ++ clear activation)

	GGP	GGA	HYO	SG	VER	SL	IL
/a/	-	+	+	-	-	-	+
/k/	++	-	-	++	-	-	-
/u/	+	-	-	++	-	-	-

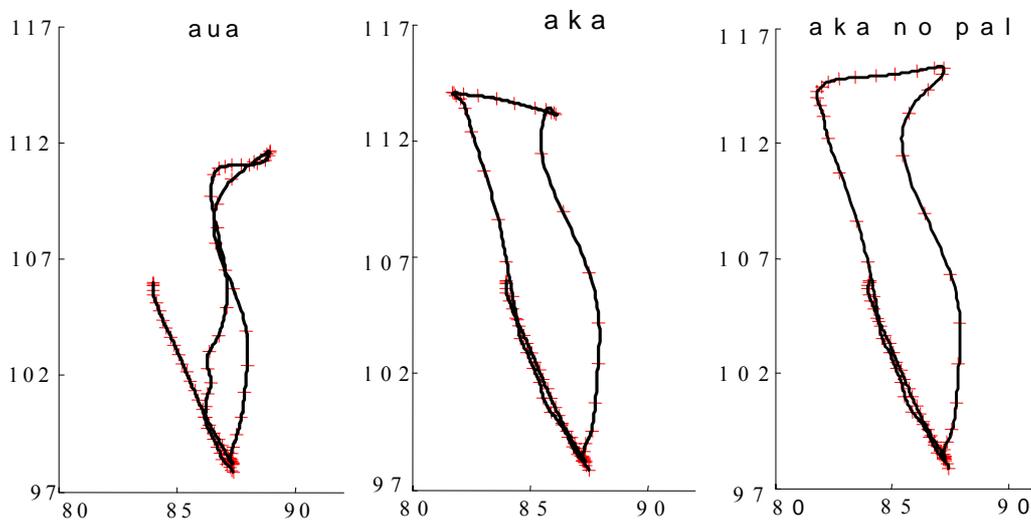


Figure 3: Trajectories of the simulated sequences in mm: /a u a/, /a k a/ (with palate) and /a k a/ (without palate) – from left to right; all simulations start at a rest position (at 83.5mm, 106mm) move downwards to the first /a/ than upwards for the /k/ or /u/, forwards and downwards again for the second /a/

We would like to note briefly another factor which might influence the shape of the looping patterns and maybe due to the anatomical properties of the relevant speaker. Figure 4 plots EMA trajectories for two speakers repetitively producing /ka/-sequences for a period of 10 seconds. Their task was to realise the syllables as quickly and as intelligible as possible (for methodology see Hartinger 2005). The x-y coordinates of the tongue back coil (in grey) and the tongue dorsum coil (in black) are displayed. The first and the last tokens are discarded for visualisation purpose. The bold black line on top corresponds to the palate trace. Speaker 1 clearly exhibits larger looping patterns for n=39

repetitions, especially due to the larger forward movement of the tongue in comparison to speaker 2 (n=35 repetitions). It seems implausible to explain the speaker dependent differences due to differences in speech rate since these are minor. Additional to possible biomechanical and aerodynamic factors which may contribute to the different looping patterns, one can also notice differences in the palate shape for the two, with speaker 1 exhibiting a dome shaped palate (see figure 4) and speaker 2 a flat palate shape from a sagittal perspective (see figure 5). The palate shapes for the two speakers are not only known due to the palate trace of the EMA recording, but they have also been analysed on the basis of their EPG palates. It is hypothesised that the variations in palate shape, the planned consonant target (for speaker 1 it is further backward), and the angle of incidence between tongue trajectory and palate contour are partly responsible for the different trajectories (for the general idea of the latter see Brunner et al. 2005). However, this hypothesis needs further verification by implementing different palate shapes in the biomechanical model.

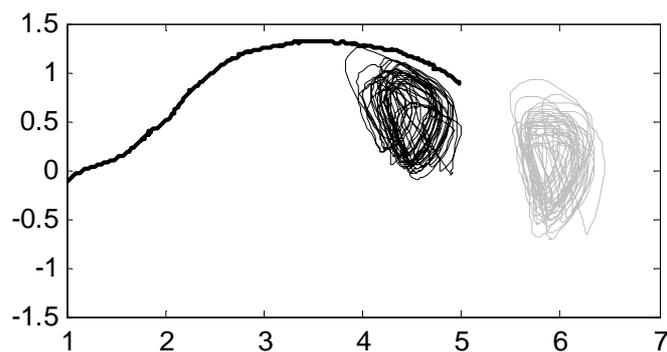


Figure 4: Articulatory trajectories during repetitive /ka/- productions for speaker 1. black: tongue dorsum coil, grey: tongue back coil

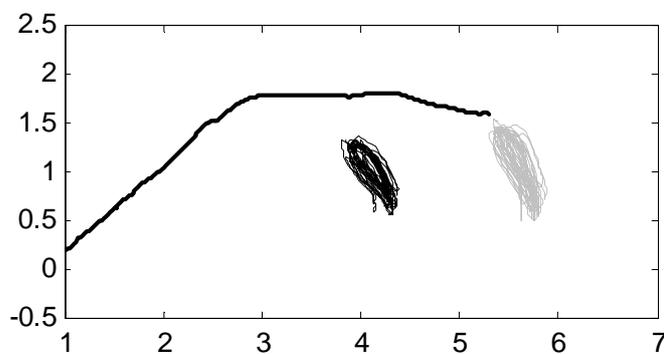


Figure 5: Articulatory trajectories during repetitive /ka/- productions for speaker 2, black: tongue dorsum coil, grey: tongue back coil

2.4 Limits due to vocal tract borders

It is one of the peculiarities of speech that the most flexible articulator the tongue, moves in a narrow space delimited by soft tissues (lips, cheeks, soft palate, pharyngeal walls) and hard tissues (the hard palate and the teeth)⁴. We mainly focus on the upper limit for the tongue's action, the palate and assume that:

(1) these vocal tract borders influence kinematic patterns and their token to token variability especially for those sounds which are realised very close to the vocal tract borders such as high front vowels (see Mooshammer et al. 2004, Brunner et al. this volume).

(2) the tongue's action at the palate is taken into account in the speech motor control process in terms of limiting the degrees of freedom for tongue movement and supporting the tongue's shaping. As far as consonant production is concerned, Stone (1991) for instance suggested that some tongue shapes, particularly those in the production of alveolar fricatives, could not be produced by a free-standing tongue position, i.e. without the palate as a reference.

In previous studies (Fuchs et al. 2001, Fuchs et al. in press) we investigated the production strategies of alveolar stops and fricatives by means of simultaneous EMA and EPG recordings. For alveolar stops versus fricatives, two different control strategies were hypothesized: a target above the contact location for alveolar stops resulting in a collision of the tongue tip at the palate as opposed to a precise positioning of the tongue at the lateral margins of the palate for alveolar fricatives. Results for both strategies were evident in tongue tip kinematics and tongue palate contact patterns. The large deceleration peak in /t/ during the closing gesture in comparison to a smaller peak in the preceding opening gesture supports the hypothesis for a collision of the tongue tip at the palate (in agreement with Hoole 1996, Fuchs et al. 2001). Additionally, the movement amplitude and the velocity for the closing gesture in /t/ were larger as opposed to the alveolar fricative, although the closing gesture duration was significantly shorter (/a/-context). The stop also showed more anterior palatal contact patterns than the fricative which may be interpreted as a consequence of the collision of the tip against the palate in comparison to a precise positioning. Further evidence for this hypothesis was provided by measuring the amplitude of movement during the acoustically defined closure or constriction. The tip sensor moved to a greater extent for the stop than for the fricative.

Based on our previous results we suppose that the collision of the tongue at the palate has an impact on the duration of the deceleration phase of the closing gesture, i.e. the stronger the impact, the shorter the deceleration phase,

⁴ The effects of soft tissues on tongue kinematics may be different in comparison to effects due to an action of the tongue at hard tissues.

resulting in an asymmetrical velocity profile. Since the acceleration phase should not be affected and the deceleration duration shortens, the profile should become skewed to the right. In order to test this hypothesis we carried out the following experiment⁵. Three speakers were recorded by means of EMA producing repetitive CV-syllables (/ta/ and /ka/) as fast and as intelligible as possible within a 10s time interval. On average between 35 and 40 syllables were produced. So far only closing gestures have been taken into account. The acceleration and the deceleration duration for the tongue back sensor in /ka/ and for the tongue tip sensor in /ta/ were measured in the tangential velocity signal. The acceleration phase was defined as the duration between closing gesture onset (velocity minimum) and the velocity peak and the deceleration duration as the time interval between peak velocity and closing gesture offset (following velocity minimum). Figure 6 shows the results of this measurement.

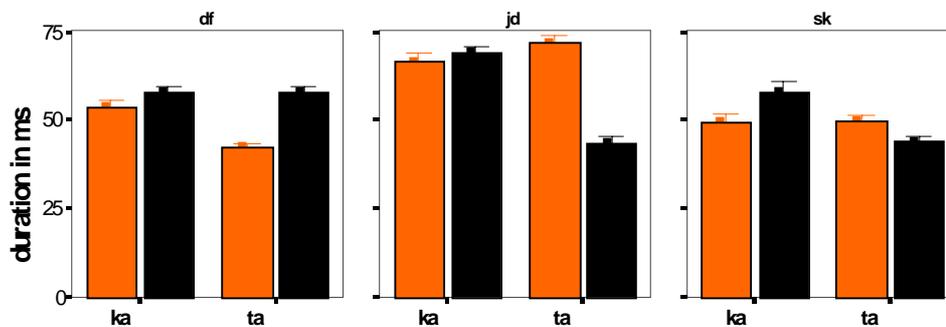


Figure 6: Means of acceleration and deceleration duration in ms with +/-1 standard error for the three subjects (df, jd and sk from left to right), grey bars correspond to the acceleration phase and black bars to the deceleration phase; left two bars /ka/, right two bars /ta/

At first glance, the results in figure 6 do not support the predicted patterns. For /ka/ none of the subjects shows the differences we supposed, since the deceleration duration is longer for sk, and rather similar to the acceleration duration for jd and for df. For /ta/ the results for two subjects (jd and sk) are in agreement with our assumptions, but df shows the reverse.

When looking into the details it becomes apparent that speaker df produces a small loop in the alveolar stop (see figure 7, right upper graph), i.e. the longer deceleration phase is due to a small forward sliding of the tip at the alveolars, starting at the marked dotted line.

⁵ Thanks to Mariam Hartinger and Jörg Dreyer for the recordings.

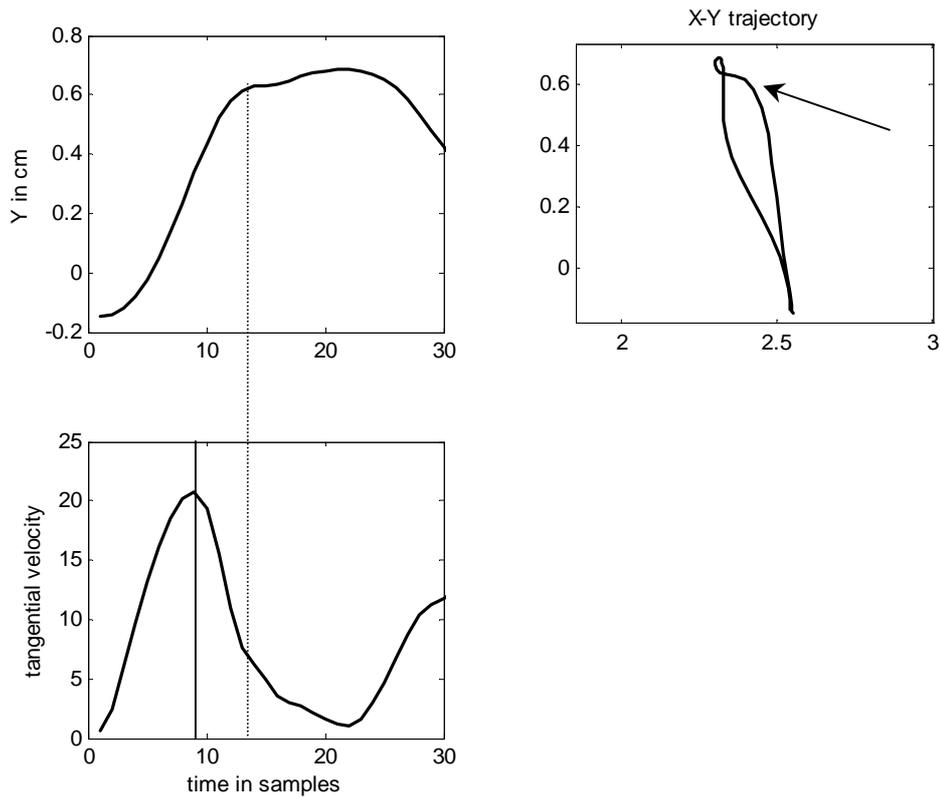


Figure 7: Example for DF's /ta/ production, left upper graph: upward vertical tongue tip movement, right upper graph: XY trajectory, left lower graph: tangential velocity profile, bold line: velocity peak, dotted line: beginning of forward movement corresponds to the array in the right upper graph

The corresponding tangential velocity profile (left lower graph) decelerates more slowly after tongue-palatal contact was made. The deceleration phase can be divided into two different parts, one where the tongue makes first contact with the palate and the second, where it continues to move along the palate in forward direction. Here the deceleration phase becomes longer than the acceleration. It is a typical pattern in the results for the velar stop and to some extent in the production for the alveolar stop for DF.

We used the tongue model in order to study the impact of the tongue's movement at the palate. The model has the advantage that simulations can be run with and without the palate in place and both conditions can be compared. Since the model is most realistic for /ka/ we chose the velar sequence and simulated 100 sequences with varying Styloglossus activity (lambda values between 61 and 91 in 0.3 steps; the lower the value, the greater the activity and the likelihood for an impact). We used two conditions, one with palate in place and one without the palate.

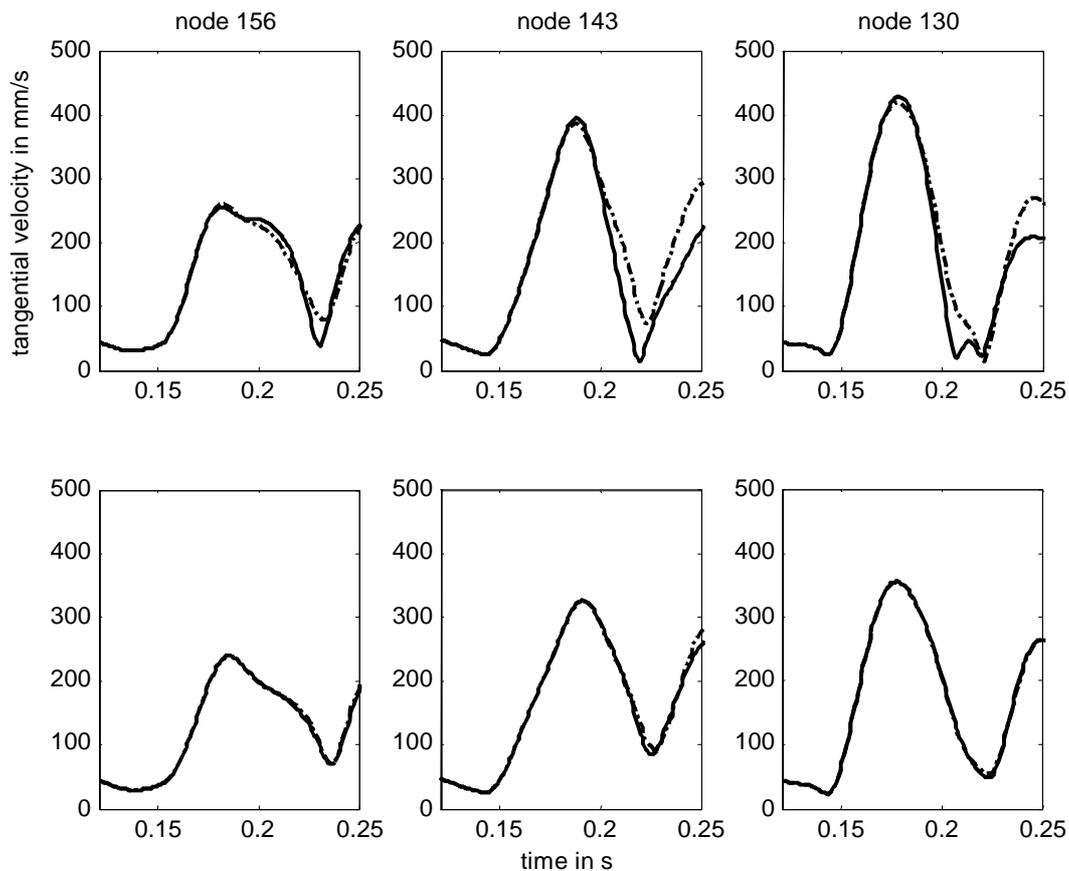


Figure 8: Simulated data showing the tangential velocity of the closing gesture for three different nodes (left: 156, middle: 143, right: 130 corresponding to: further front (156) to a further back (130) articulation in the vocal tract; bold lines: simulations with the palate in place, dashed-dot lines: simulations with no palate; upper track: high SG activity (73), lower track: lower SG activity (79)

Figure 8 shows a comparison of tangential velocities between simulations with and without the palate in place. In the upper track a higher Styloglossus activity has been chosen than in the lower track. It corresponds to a target which

is higher above the palate. These simulations clearly exhibit an impact of vocal tract limits on tongue kinematics when comparing simulations with and without palate. Differences are less strong for the more anterior articulation (node 156) in comparison to the more posterior articulation (node 130). The deceleration phase is noticeably shorter due to the impact. In the simulations with lower Styloglossus activity differences are rather marginal and won't be taken into account.

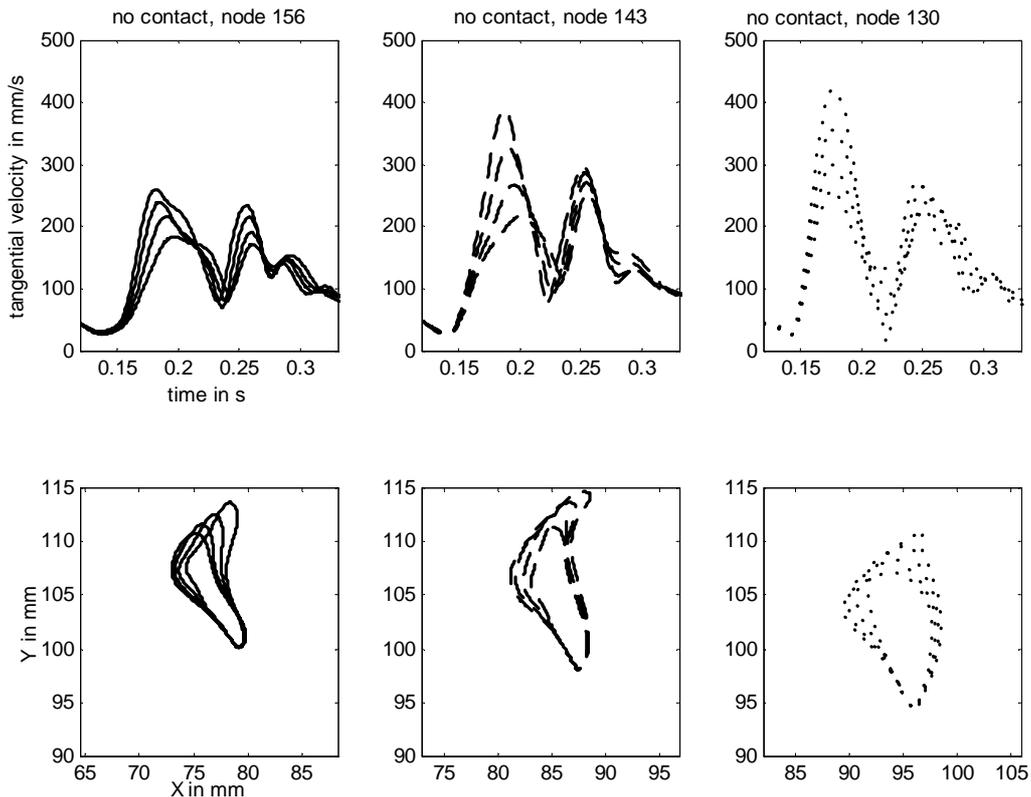


Figure 9: Tangential velocities in mm/s (upper track) and XY-trajectories in mm (lower track) for simulations with increasing Styloglossus activity (91, 85, 79, 73); no palate in place, 3 columns correspond to the 3 nodes (156, 143, 130) from left to right, see text for further details

Figure 9 exhibits some further interesting results for the simulations without the palate (similar effects exist also in the simulations with the palate and are therefore not included here): (1) higher SG activity coincides with higher peak velocity, with an overall shorter duration and further backward movement during the closing gesture. (2) Asymmetries in the velocity profile of the closing gestures (first velocity profile in the upper tracks) vary with respect to the different nodes. Node 156 shows a left shaped pattern with a longer deceleration phase than the acceleration phase. Nodes 143 and 130 which are further

backward in their placement show rather symmetrical velocity profiles with some variations depending on the SG activity.

This means that the speaker dependent differences we found in the experimental data may be a consequence of biomechanics (muscle cocontraction) and differences in the tongue back sensor placement, with sk having the tongue back sensor located more anteriorly than speakers jd and df. Additionally, they do not contradict our hypothesis that the impact of the tongue at the palate reduces the deceleration phase since all the simulations with a non-negligible impact (see figure 8) exhibit this pattern. However, the velocity profile serving as a reference may be asymmetrical due to biomechanical reasons and not bell-shaped as supposed. Therefore, the deceleration duration is not generally shorter than the acceleration duration.

3 Conclusions

In this paper, we have presented a number of examples showing that the interpretation of speech kinematics in terms of underlying speech motor control is not straightforward. We have demonstrated that the parameters such as peak velocity, asymmetry of the velocity profile, and trajectory curvature, which have frequently been used in the literature to infer hypotheses about the underlying speech motor control strategies, are in fact the result of complex and non linear combinations of different factors. These factors are obviously linked with high level motor control strategies including optimal planning and listener oriented control, but they are also linked with physical phenomena such as speech articulators' muscle anatomy, biomechanics and dynamics, mechanical interactions between articulators (tongue-palate or tongue-teeth contacts), and interactions between airflow and soft tissues. As exemplified by the controversial debates about the origin of the so-called 1/3 power law corresponding to a non linear relation between tangential velocity and trajectory curvature of human arm movements, this statement holds not only true for speech, but also for every kind of skilled human movement. However, it is particularly acute for speech, since speech production necessitates the control of a complex motor system, coupling hard bodies, soft tissues and aerodynamics under time-varying mechanical constraints via the coordination of more than 30 potentially independent muscles. It involves the control of movements realized in sometimes very short durations, discarding any potential on-line feedback mechanism going up to the cortex.

And last but not least, speech production is a peculiar motor activity which essence is semiotic. Hence, hypotheses about speech motor control must

not only give an account of the observed kinematics, but also of the link between the observed kinematics and the underlying semiotic code.

Studies of speech kinematics have been the main basis of speech production research for many years. They have been extremely fruitful and have allowed to develop major hypotheses about speech motor control and its interaction with speech perception in the semiotic framework of speech communication.

To continue this kind of investigation is a necessity and justifies the remarkable effort that many of our colleagues have put in the development and enhancement of new data acquisition techniques using the most recent developments in physical measurement technologies. In parallel, investigations in the broad domain of human motor control have made noticeable progress, in particular in the domain of learning, of internal representations, and in the way to integrate low-level short delay feedback loops. Interpretations of speech kinematics in terms of motor control have been inspired from these findings. However, the main trend in speech production studies has been to relate observations of speech kinematics directly to high level motor control strategies involving complex internal models. They have often overseen the important role of physics in shaping the patterns of speech kinematics.

With this paper we propose that the physical properties of the peripheral speech production apparatus should be put into the center of our investigations in order to account for the complex nature of speech kinematics. We suggest that the complex nature of speech kinematics is for a large part due to the complex peripheral speech apparatus and that it may not systematically be found in higher level motor control strategies.

Acknowledgements

This work has been supported by a grant from the German Research Council (DFG) GWZ 4/8-1, P.1. Thanks to Joe Perkell for the collaboration on a previous version of the degrees of freedom in vowel production, to Mariam Hartinger for recording the EMA data on repetitive syllables, for Phil Hoole, Christian Abry, and the editors of this volume to provide useful comments on an earlier version of this paper.

References

Adams, S.G., Weismer, G. & Kent, R.D. (1993): Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research*, 36: 41-54.

- Bullock, D. & Grossberg, S. (1988): Neural dynamics of planned arm movements: Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review*, 95: 49-90.
- Brunner, J., Fuchs, S., Perrier, P. & Kim, H.-Z. (2004): Correlation between angle of incidence and sliding patterns of the tongue along the palate in Korean velar stops. *147th Meeting of the Acoustical Society of America, New York* [abstract & poster].
- Brunner, J. (2005): Supralaryngeal mechanisms of the voicing contrast in velars. *ZAS Papers in Linguistics* 39.
- Brunner, J., Fuchs, S. & Perrier, P. (2005): The influence of the palate shape on articulatory token-to-token variability. *ZAS Papers in Linguistics*, 42: 43-66
- Feldman, A.G. (1986): Once more on the Equilibrium-Point Hypothesis (λ Model) for motor control. *Journal of Motor Behavior*, 18(1): 17-54.
- Feldman, A.G., Adamovich, S.V., Ostry, D.J. & Flanagan, J.R. (1990): The origin of electromyograms – explanations based on the equilibrium point hypothesis. In Winters, J. & Woo, S.: *Multiple muscle systems: Biomechanics and movement organization*. Springer Verlag: New York: 195-213.
- Fuchs, S., Perrier, P. & Mooshammer, C. (2001): The role of the palate in tongue kinematics: an experimental assessment in VC sequences from EPG and EMMA data. *Proceedings of Eurospeech Aalborg*, 3: 1487-1490.
- Fuchs, S., Perrier, P., Geng, C. & Mooshammer, C. (in press): What role does the palate play in speech motor control? Insights from tongue kinematics for German alveolar obstruents. Harrington, J. and Tabain, M. (Eds.): *Towards a better understanding of speech production processes*. Psychology Press: New York.
- Geng, C., Fuchs, S., Mooshammer, C., Pompino-Marschall, B. (2003): How does vowel context influence loops? *Proceedings of the 6th Speech Production Seminar Sydney*: 1-6.
- Gerard, J., Perrier, P. & Payan, Y. (in press): 3D biomechanical tongue modelling to study speech production. In Harrington, J. and Tabain, M. (eds.): *Towards a better understanding of speech production processes*. Psychology Press: New York.
- Gracco, V.L. (1988): Timing factors in the coordination of speech movements. *The Journal of Neuroscience*, 8: 4628-4639.
- Gribble, P.L. & Ostry, D.J. (1996): Origins of the power law relation between movement velocity and curvature: Modeling the effects of muscle mechanics and limb dynamics. *Journal of Neurophysiology*, 76(5): 2853-2860.
- Harris, C. M., & Wolpert, D. M. (1998): Signal-dependent noise determines motor planning. *Nature*, 394: 780–784.
- Harshman, R. A., Ladefoged, P. N., & Goldstein, L. (1977): Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*, 62: 693–707.
- Hartinger, M. (2005): *Untersuchungen der Sprechmotorik von Poltereren mit Hilfe der elektromagnetischen mediosagittalen Artikulographie (EMMA)*. Unpublished PhD thesis at Martin Luther University Halle (Saale).

- Heinz, J.M. & Stevens, K.N. (1965): On the relations between lateral cineradiographs, area functions and acoustic spectra of the speech. *Proceedings of the 5th International Congress of Acoustic*, A44.
- Hogan, N. (1990): Mechanical impedance of single- and multi-articulator systems. In Winters, J.M. & Woo, S.L.-Y.(eds.): *Multiple muscle systems. Biomechanics and movement organization*. Springer: Berlin, New York: 149-164.
- Hoole, P. (1996). Theoretische und methodische Grundlagen der Artikulationsanalyse in der experimentellen Phonetik. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM)*, 34: 3-173.
- Hoole, P. (1998): Modelling tongue configuration in German vowel production. *Proceedings of the 5th ICSLP Sydney*, 5: 1867-1870.
- Hoole, P. (1999): On the lingual organization of the German vowel system. *Journal of the Acoustical Society of America*, 106(2): 1020-1032.
- Hoole, P., Munhall, K. & Mooshammer, C. (1998): Do air-stream mechanisms influence tongue movement paths? *Phonetica*, 55(3): 131-146.
- Houde, R. (1968): A study of tongue body motion during selected consonant sounds. *Speech Communications Research Laboratory, Santa Barbara, SCRL Monograph 2*.
- Jackson, M.T.T. (1988): Analysis of tongue positions: Language-specific and cross-linguistic models. *Journal of the Acoustical Society of America*, 84(1): 124-143.
- Lebedev, S., Tsui, W.H. & Van Gelder, P. (2001): Drawing movements as an outcome of the principle of least action. *Journal of Mathematical Psychology*, 45: 43-52.
- Löfqvist, A. & Gracco, V.L. (1997): Lip and jaw kinematics in bilabial stop consonant production. *Journal of Speech, Language, and Hearing Research*, 40(4): 877-893.
- Löfqvist, A. & Gracco, V. L. (2002): Control of oral closure in lingual stop consonant production. *Journal of the Acoustical Society of America*, 111(6): 2811–2827.
- Maeda, S. (1990): Compensatory articulation during speech: evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In Hardcastle, W.J. & Marchal, A. (eds.): *Speech production and speech modelling*. Dordrecht: Kluwer Academic Publishers: 131-150.
- Marhefka, D. W. & Orin, D. E. (1996): Simulations of contact using a non-linear damping model. *Proceedings of IEEE International Conference on Robotics and Automation, Minneapolis, MN*, 2: 1662–1668.
- Massey, J., Lurito, J., Pellizzer, G. & Georgopoulos, A. (1992): Three-dimensional drawings in isometric conditions: relations between geometry and kinematics. *Experimental Brain Research*, 88: 685-690.
- McClellan, M.D., Clay, J.L. (1995): Activation of lip motor units with variations in speech rate and phonetic structure. *Journal of Speech, Language and Hearing Research*, 38: 772-782.

- McClellan, M.D. & Tasko, S.M. (2003): Association of orofacial muscle activity and movement during changes in speech rate and intensity. *Journal of Speech, Language and Hearing Research*, 46: 1387-1400.
- Milner, T.E. & Ijaz, M. (1990): The effect of accuracy constraints on three-dimensional movement kinematics. *Neuroscience*, 35: 365-374.
- Mooshammer, C., Hoole, P. & Kühnert, B. (1995): On loops. *Journal of Phonetics*, 23: 3–21.
- Mooshammer, C., Perrier, P., Fuchs, S., Geng, C. and Pape, D. (2004). An EMMA and EPG study on token-to-token variability. *Arbeitsberichte Institut für Phonetik und digitale Sprachverarbeitung Universität Kiel (AIPUK)*, 36: 46-63.
- Nelson, W.L. (1983): Physical principles for economies of skilled movements. *Biological Cybernetics* 46: 135-147.
- Nix, D. A., Papcun, G., Hogden J., & Zlokarnik, I. (1996): Two cross-linguistic factors underlying tongue shapes for vowels. *Journal of the Acoustical Society of America*, 99: 3707-3717.
- Ohala, J.J. (1983): The origin of sound patterns in vocal tract constraints. In MacNeilage, P.F. (ed.): *The Production of Speech*. Springer Verlag: New York, Heidelberg, Berlin: 189-216.
- Ostry, D.J. & Feldman, A.G. (2003): A critical evaluation of the force control hypothesis in motor control. *Experimental Brain Research*, 153: 275-288.
- Payan, Y. & Perrier, P. (1997): Synthesis of V–V sequences with a 2D biomechanical tongue model controlled by the equilibrium point hypothesis. *Speech Communication*, 22: 185-205.
- Pelorson, X., Hirshberg, A., Wijnands, A.P.J. & Bailliet H.M.A. (1995): Description of the flow through in-vitro models of the glottis during phonation. *Acta Acustica*, 3: 191-202.
- Pelorson X., Liljencrants J., Kröger B. (1995): On the aeroacoustics of voiced sound production. *Proceedings of the 15th International Congress on Acoustics, Trondheim, Norway*, 4: 501-504.
- Perrier P., Boe L.J. & Sock R. (1992): Vocal tract area function estimation from midsagittal dimensions with CT scans and a vocal tract cast: Modelling the transition with two sets of coefficients. *Journal of Speech and Hearing Research*, 35: 53–67.
- Perrier, P., Payan, Y., Perkell, J., Zandipour, M., Pelorson, X., Coisy, V. & Matthies, M. (2000a): An attempt to simulate fluid-walls interactions during velar stops. In *Proceedings of the 5th Seminar on Speech Production: Models and Data, Kloster Seeon*: 149-152.
- Perrier, P., Perkell, J., Payan, Y., Zandipour, M., Guenther, F. & Khalighi, A. (2000b): Degrees of freedom of tongue movements in speech may be constrained by biomechanics. *Proceedings of the ISCLP Beijing*, 2: 162-165.
- Perrier, P., Payan, Y., Zandipour, M. and 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(3): 1582-1599.

- Perrier P. (in press) About speech motor control complexity. In Harrington, J. and Tabain, M. (eds.): *Towards a better understanding of speech production processes*. Psychology Press: New York.
- Sanguineti, V., Laboissière, R., & Payan, Y. (1997): A control model of human tongue movements in speech. *Biological Cybernetics*, 77(1): 11–22.
- Sanguineti, V., Laboissière, R., & Ostry, D.J. (1998): A dynamic biomechanical model for neural control of speech production. *Journal of the Acoustical Society of America*, 103(3): 1615-1627.
- Stone, M. (1991): Toward a model of three-dimensional tongue movements. *Journal of Phonetics*, 19: 309-320.
- Svirsky, M., Stevens, K., Matthies, M., Manzella, J., Perkell, J. and Wilhelms-Tricarico, R. (1997) Tongue surface displacement during bilabial stops. *Journal of the Acoustical Society of America*, 102: 562-571.
- Tasko, S.T. & Westbury, J.R. (2004): Speed–curvature relations for speech-related articulatory movement. *Journal of Phonetics*, 32: 65-80.
- Viviani, P., & Terzuolo, C. (1982) Trajectory determines movement dynamics. *Neuroscience*, 7, 431-437
- Viviani, P. & Schneider, R. (1991): A developmental study of the relationship between geometry and kinematics in drawing movements. *Journal of Experimental Psychology. Human Perception and Performance*, 17: 198-218.
- Viviani, P. & Flash, T. (1995): Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *Journal of Experimental Psychology. Human Perception and Performance*, 21: 32-53.
- Wiencke, G., Janssen, P. & Belderbos, H.(1987): The influence of speaking rate on the duration of jaw movement. *Journal of Phonetics*, 15:111-126.