Speed–Curvature Relations in Speech Production Challenge the 1/3 Power Law

Pascal Perrier¹ and Susanne Fuchs²

¹Institut de la Communication Parlée, Laboratoire Grenoble Image Parole Signal Automatique, Centre National de la Recherche Scientifique, Grenoble Institut National Polytechnique, Grenoble, France; and ²Center for General Linguistics, Typology and Language Universals, Berlin, Germany

Submitted 9 October 2007; accepted in final form 17 June 2008

Perrier P, Fuchs S. Speed–curvature relations in speech production challenge the 1/3 power law. J Neurophysiol 100: 1171–1183, 2008. First published June 18, 2008; doi:10.1152/jn.01116.2007. Relations between tangential velocity and trajectory curvature are analyzed for tongue movements during speech production in the framework of the 1/3 power law, discovered by Viviani and colleagues for arm movements. In 2004, Tasko and Westbury found for American English that the power function provides a good account of speech kinematics, but with an exponent that varies across articulators. The present work aims at broadening Tasko and Westbury’s study 1) by analyzing speed–curvature relations for various languages (French, German, Mandarin) and for a biomechanical tongue model simulating speech gestures at various speaking rates and 2) by providing for each speaker or each simulated speaking rate a comparison of results found for the complete set of movements with those found for each movement separately. It is found that the 1/3 power law offers a fair description of the global speed–curvature relations for all speakers and all languages, when articulatory speech data are considered in their whole. This is also observed in the simulations, where the motor control model does not specify any kinematic property of the articulatory paths. However, the refined analysis for individual movements reveals numerous exceptions to this law: the velocity always decreases when curvature increases, but the slope in the log–log representation is variable. It is concluded that the speed– curvature relation is not controlled in speech movements and that it accounts only for general properties of the articulatory movements, which could arise from vocal tract dynamics or/and from stochastic characteristics of the measured signals.

INTRODUCTION

Studying of both the kinematic characteristics of speech movements and the detailed properties of the articulatory trajectories is an approach that is widely used in speech communication research to infer the underlying control mechanisms of speech gestures. From this perspective, the so-called 1/3 power law, originally proposed by Viviani and Terzuolo (1982) to characterize the relations between speed and curvature of human movements, is of particular interest. Indeed, since the original study in 1982, numerous investigations have been concerned with the validation of this law and with its potential explanations. These studies greatly contributed to debates about fundamental issues in human motor control, which are also crucial for speech communication research, such as the use of kinematic properties in perception mechanisms or the role of centrally planned optimal motor strategies versus physically based factors in the patterning of movement kinematics.

In spite of this potential contribution to speech motor control issues, we are aware of only one study (Tasko and Westbury 2004) testing the validity of the 1/3 power law for speech movements. Its general observation is that speech movements tend to confirm the 1/3 power law, although with some discrepancies. The authors interpreted their results in terms of articulators’ specific behavior and proposed hypotheses about the underlying motor control. In line with Tasko and Westbury’s work, the present study tests the validity of the 1/3 power law for speech. We will, however, broaden the scope to an analysis of different languages and supplement experimental results with simulations carried out using a realistic model of speech production for which movements are controlled on a target-to-target basis without any specific control of the trajectory shapes or of the velocity profiles. This approach was chosen to test the universality of the speed–curvature relations across languages and to look for possible physical explanations of the 1/3 power law.

Our work is structured in the following way: in the first section, a summary of the main findings related to the 1/3 power law for human movements in general and for speech production in particular is given. Subsequently, the methodology of our study is presented, involving the description of an experimental study and a modeling approach. In the RESULTS section, the adequacy of the 1/3 power law is tested for both the experimental and the simulated data, and a comparison of these two sets of data is proposed. In the final section, interpretations and conclusions for the 1/3 power law and its potential application to speech motor control are provided.

Rationale: the 1/3 power law

A number of empirical studies have shown that there is a strong coupling between the speed and the trajectory curvature of human movements: when curvature increases, speed decreases and vice versa. A major formalization of this coupling was provided by Viviani and colleagues (Lacquaniti et al. 1983; Viviani and Terzuolo 1982), who found for planar drawing hand movements that the angular velocity α(t) and the trajectory curvature c(t) of the end effector obey the so-called 2/3 power law.

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
When considering the tangential velocity $v(t)$ and the trajectory curvature $c(t)$ this relation becomes

$$v(t) = kc(t)^{-1/3} \quad \text{[since } v(t) = a(t)c(t)]$$

which is known as the “1/3 power law.” Factor $k$ is defined as the “velocity gain factor,” accounting for differences in average movement velocity. The terms “1/3 power law” and “power law” will be used interchangeably throughout.

**Major findings and potential explanations**

Further evidence supporting the findings of Viviani and colleagues was found for complex arm movements at various rates (Viviani and Cenzato 1985), for locomotion (Vieilledent et al. 2001) and for eye motion (de’Sperati and Viviani 1997), suggesting that the power law could be a fundamental universal characteristic of human movements. It was even suggested that this characteristic was so typical of human movements that it could strongly influence human perception in terms of naturalness and classification (Viviani and Stucchi 1992) as well as in terms of perceptual anticipation of impending events (Kandel et al. 2000).

There have been many debates in the literature about the possible origin of the relation between speed and curvature and whether it reflects the underlying motor control of the CNS. It has been suggested that the power law results from optimization principles underlying human movement production, such as jerk minimization (Viviani and Flash 1995), maximum smoothness principles (Todorov and Jordan 1998), or minimizing the impact of neural noise on target reaching accuracy (Harris and Wolpert 1998). In contrast, it was also proposed that the 1/3 power law could arise from biomechanical properties of the peripheral motor system (Gribble and Ostry 1996). More recently, Maoz et al. (2005) suggested that it could, at least partly, arise from the noise that is inherently added to the kinematic data because of measurement inaccuracy and/or the stochastic characteristics of the motor system (including neural noise).

Arguments against the hypothesis of a major role for physical factors in the power law are mainly based on the work of Massey et al. (1992) who demonstrated that the same coupling between speed and curvature still held true when subjects were drawing curved patterns in isometric conditions with a joystick (i.e., when the motor system did not move). However, Gribble and Ostry (1996) moderated the relevance of this finding, since they showed by means of their model that under isometric conditions force variations also obey the same law. Another concern about the viability of the biomechanical explanation comes from the fact that the power law was found in mechanical systems as different as upper limb, lower limb, and eyes. A possible answer to this can be found in Gribble and Ostry (1996) and Schaal and Sternad (2002). The former suggest that basic springlike characteristics, shared by all motor systems, would justify the impact of biomechanics. The latter assume that these springlike characteristics naturally ensure the smoothness of the trajectory necessary for the power law to apply.

Criticism of the hypothesis of a dominant influence of centrally planned factors in the power law is mainly based on the observation that many exceptions to this law exist in complex human movements. Schaal and Sternad (2002) observed for instance that during the drawing of large three-dimensional (3D) elliptical patterns, the exponent of the non-linear relation between speed and curvature often greatly deviates from the theoretical value ($-1/3$). They concluded that the 1/3 power law should be seen as “an epiphenomenon of smooth oscillatory trajectory generation in joint space.”

**Preliminary results for speech production**

This summary suggests that, despite its controversy, the 1/3 power law is a major experimental finding that raises fundamental issues about the control and perception of human movements: What kinematic properties of human movements are directly related to centrally planned factors? To what extent are properties of the motor system used to perceive human movements?

In this context, speech movements should have been of particular interest. Indeed, speech biomechanics is very specific because two of the main vocal tract articulators (tongue and lips) are nonrigid bodies interacting with hard structures (a complex biomechanical case to study) and because time-variable boundary conditions (tongue–palate and tongue–teeth interaction; upper lip–lower lip and lip–jaw interactions) probably constrain speech gestures to a great extent. In addition, speech task specification is likely to vary across the world’s languages because of the well-known differences in size and complexity of phonological inventories. Moreover, speech gestures in the oral cavity are to a large extent not perceived visually, but auditorily via the emitted speech signal after a nonlinear transformation from the articulatory to the acoustic domain.

In spite of these interesting specificities, to our knowledge, only one study has addressed the power law issue in the context of speech production. It was carried out by Tasko and Westbury (2004) who took advantage of the large X-ray microbeam speech production database to analyze the kinematic characteristics of speech for 18 American English speakers. They found an exponent value that was near, but not exactly, ($-1/3$). The exponent value, the velocity gain factor, and the strength of the speed–curvature relation varied systematically across subjects and subjects. For the whole set of data, the exponent values ranged from $-0.44$ to $-0.34$. Therefore the absolute values of the exponent are slightly above the range expected from the original power law for planar drawing movements. However, they are much smaller than the absolute values found by Pollick and Ishimura (1996) for 3D straight-ahead point-to-point movements (ranging from 0.52 to 0.67). According to Tasko and Westbury (2004), this range of variation in the exponent values suggests that speech gestures are closer to planar drawing movements for which the whole trajectory is controlled than to target-directed movements.

On average, the variance of speech velocity explained by the power law ranged from 65 to 70%, which is smaller than the values observed for drawing and visual-tracking movements. This suggests a weaker speed–curvature coupling in speech movements than that in other human movements. The authors ascribe this to the fact that their speech material included consonantal segments, where the tongue is in contact with the palate, which constrains its movement. Finally, Tasko and
Westbury proposed that interarticulator differences found in the exponent and in the velocity gain factor values reflect the behavioral specificities of each orofacial structure.

The present work aims at broadening the study by Tasko and Westbury (2004) in essentially two directions: 1) the potential influence of language-specific constraints is assessed by studying the speed–curvature relationships for speakers of relatively unrelated languages (French, German, Mandarin Chinese); 2) the contribution of biomechanical factors is tested by analyzing the speed–curvature relationship in artificial tongue movements generated with an anthropomorphic speech production model compared with the experimental data. In both cases, special attention has been devoted to the analysis of the power law adequacy at the level of each singular tongue gesture. Indeed, if Viviani and colleagues’ hypothesis is true, and if the power law results from motor control strategies that they can be perceptually identified, it should be systematically observed for each single gesture. Thus in our study speech sequences were split into small-movement segments delimited by two successive velocity zero crossings, and speed–curvature relations were calculated in parallel for the whole set of data and for each segment separately.

METHODS
Experimental data
To evaluate the universality of the power law, it was decided to analyze articulatory data from various languages and various speakers, collected in the context of studies with various objectives. These data were gathered at ICP (Institut de la Communication Parlé) and at ZAS (Center for General Linguistics, Typology and Language Universals) in the last 10 years. Similar experimental setups were used in both labs (2D Electromagnetic Midsagittal Articulograph AG100; Carstens Medizin Electronics), with the same data postprocessing and with similar experimental protocols.

SUBJECTS. Three rather unrelated languages were selected: French, German, and Mandarin Chinese. Two speakers were considered for each language (F1 and F2 for French, G1 and G2 for German, C1 and C2 for Mandarin Chinese). They were male subjects except F1. Their ages ranged from 25 to 30 yr except for C1 who was 45. They were scientists working in our labs except C2. All the speakers had no history of speech, language, or hearing pathologies. The French and the Mandarin Chinese data were collected in the context of a cross-linguistic study of coarticulation mechanisms (Ma et al. 2006), whereas the German data were obtained for a study of interspeaker token-to-token variability (Mooshammer et al. 2004). Details about the experimental procedures and the data processing are given in Ma et al. (2006) and Mooshammer et al. (2004). Only the most immediately relevant aspects of the method are described in the following text.

SPEECH TASKS. The cross-linguistic study of coarticulation of Ma et al. (2006) aimed at characterizing in Vowel1–Vowel2 and Vowel1–Consonant–Vowel2 sequences how the production of the second vowel Vowel2 influences the articulation of the preceding sounds (anticipatory coarticulation). The authors of this study were especially interested in looking at potential interactions between the phonological properties of the language and anticipatory coarticulation strategies. Mandarin Chinese and French were chosen because the phonological status of the syllable differs across the two languages.

The same sequences were recorded for both languages. They were carefully chosen to respect the phonotactical rules of each language. In Mandarin Chinese all syllables were pronounced with a high-level tone (i.e., flat tone). The vowels were one of /a, u/ and the consonant one of /t, k/. The sequences were recorded both as nonsense words in meaningful carrier sentences (example for French: “C’est akì ça?”; i.e., “Is that akì?”) or as parts of meaningful words or groups of words within meaningful sentences (example for French: “C’est Harry qui t’a touché?”; i.e., “Did Harry touch you?”). For the current study, we selected the latter part of the corpus for both languages because it offers a larger variety of sounds. Ten repetitions of five different sentences lasting for around 1 s each were considered for subjects F1, F2, and C2. For subject C1, who was the subject for the pilot study, the corpus was slightly different and corresponds to four repetitions of 20 different sentences.

The German material was collected to study the patterns of articulatory variability measured for each speaker in different repetitions of a vowel in the same phonetic context. Moreover, the authors analyzed the potential influences of perceptual constraints, morphological characteristics of the vocal tract, and linguistic factors on the patterns of variability. In its whole the corpus consisted of C1VC2V/ nonsense words with C1 and C2 being either the velar stops /k/ and /g/ or the bilabial stops /p/ and /b/ and V being one of the 14 tense or lax German vowels. The initial stop C1 was phonologically voiced and the medial C2 was phonologically voiceless. V is in a stressed position. All nonsense words were embedded in the carrier sentence “Sage...bitte” (“Say...please”) and they were repeated 10 times. The whole set of data yet consisted of 280 sentences for each subject. For the current study, a limited subset of these data was analyzed, consisting of 50 randomly selected sentences for each speaker.

DATA ACQUISITION. For all speakers, the articulatory data were collected with the AG100 system. With this device measurements are based on the electromagnetic induction phenomenon. Alternating magnetic fields are generated by three transmitter coils that are mounted on the corners of a helmet that has the form of an equilateral triangle. Small transducer coils (hereafter called sensors) are attached to the articulators in the midsagittal plane. Each sinusoidal magnetic field induces—in the subject's mouth—sinusoidal current in each sensor, the intensity of which is proportional to the inverse of the cube of the Euclidian distance separating the transmitter and the sensor. The measurement of the intensity of the three currents generated in each sensor by the three transmitters, after some corrective preprocessing, yields measures of this sensor’s locations as a function of time. To minimize measurement errors, the sensors have to be accurately located in the midsagittal plane of the head. Two sensors, located on fixed parts of the head, are used as reference sensors to compensate for possible head rotation in the midsagittal plane. For each measurement, a tilt factor is calculated that tells whether the measurements can be trusted. Further information about this technique can be found in Perkell et al. (1992) and in Hoole and Nguyen (1999).

For all subjects, four sensors were located on the tongue (note that for speaker F2 the Tback sensor stopped working during the experiment; it was not at all considered for this study). They were quasi evenly distributed from about 1 to 5 cm from the tongue tip. Figure 1 (top) shows a schematic view of these sensor locations on the tongue contour in the midsagittal plane. Sensors were designated (from front to the back) as Tip, Thade, Tdors, and Tback. Reference coils were glued to the bridge of the nose and above the upper incisors.

The horizontal direction of the speaker’s bite plane was also measured at the end of each session. To do so a Plexiglas sheet was inserted in the subject’s mouth. Two sensors were located on this sheet along the anterior–posterior direction. The line joining these sensors, when the speaker is asked to bite against the sheet, is considered to be the horizontal direction. The articulatory data were then rotated around the reference sensor on the upper incisor, to ensure that the x and y directions of the rotated data correspond...
Simulated data

To study the potential impact of biomechanical factors on the speed–curvature relation, tongue movements were simulated with a biomechanical model controlled on a target-to-target basis (Payan and Perrier 1997; Perrier et al. 2003). Earlier works in our group (see in particular Payan and Perrier 1997; Perrier et al. 2003, 2005) have demonstrated that the tongue trajectories simulated with this model are realistic and similar to real trajectories measured on human speakers.

In this model, the kinematic properties of the movements are not specified at the motor control level. Tongue movements are generated as sequences of submovements between targets. In the complete speech production model, called GEPPETO (GEstures shaped by the Physics and by a PErectually Oriented Targets Optimization), these targets are related to the phonological structure of the speech sequence (Perrier et al. 2005): targets are considered as physical correlates of the phonological inputs. However, in the current study, this motor control layer was not used and the successive targets underlying the production of tongue movements are considered to be just intermediate spatial objectives within a larger tongue movement, without any explicit links to the phonological level.

For the production of a movement, the motor control system specifies via the motor control variables (see following text) the mechanical properties and the timing (in terms of transition time between targets and hold duration of each target) of the successive targets. The shift of the motor control variables between the successive target values is made at a constant rate and its timing is set up independently of any consideration related to the shape of the articulatory path or to articulatory kinematics. Thus in this model, the kinematic properties of the movements (in particular the trajectory curvature and the tangential velocity) are neither explicitly controlled nor are they the consequences of an optimal motor control strategy. On the contrary, the kinematic properties are the results of a combination of central and peripheral influences: 1) those of the motor command values and their timing, 2) those of the muscle force generation mechanisms (see following text), and 3) those of the dynamical and anatomical properties of tongue tissues (inertia, elasticity, muscle fiber arrangements).

In this model, elastic properties of tongue tissues are accounted for by finite-element (FE) modeling with a mesh defined by 221 nodes and 192 elements. The FE mesh represents a two-dimensional (2D) projection of the tongue in the midsagittal plane. It is inserted inside a projection of the vocal tract in this plane, characterized by curves representing the contours of the lips, palate, and pharynx. The jaw and the hyoid bone are also represented in this plane by static rigid structures to which the tongue is attached. Mechanical contacts of the tongue with the palate and the teeth are also modeled. The main muscles responsible for shaping and moving the projection of the tongue in the midsagittal plane are taken into account: posterior and anterior parts of the genioglossus, styloglossus, hyoglossus, inferior and superior longitudinalis, and verticalis. They are represented in the model at two different levels (see Fig. 2). First, their action on the tongue body is accounted for by “macrofibers” (the bold lines on Fig. 2) that specify the direction of the forces and the nodes of the FE mesh to which the forces are applied. Second, since the activation of a muscle modifies the elastic properties of the muscle tissues, muscles are also represented in the model by a number of selected elements within the FE structure (gray-shaded elements in Fig. 2), whose mechanical stiffness increases with muscle activation.

Muscle force generation mechanisms are modeled in a functional way according to Feldman’s Equilibrium Point Hypothesis of motor control (Feldman 1986). This model reflects the claim that α motoneuron activation, which generates force, is not centrally controlled, but is the consequence of the interaction between a central command (called λ) and afferent inputs related to muscle length and velocity. The activation is zero if the muscle length is shorter than λ. Otherwise, the activation is a function of the difference between the muscle length and λ and of the muscle length change rate. The relation between active muscle force and muscle activation is approximated by an exponential function. In addition, the sliding filament theory is taken into account in the model: the force-generation capability of a muscle depends on muscle shortening or lengthening velocity (for more details see Payan and Perrier 1997).
For the purpose of this study, 600 sequences consisting of an initial schwa (/ə/, corresponding to the tongue position at rest) followed by three different articulatory targets were generated. The motor control variables used for the three articulatory targets following the initial schwa were selected using a Monte Carlo method based on a uniform sampling of the motor command space. Consequently, the targets used in the simulations correspond to tongue shapes that are randomly selected among all the possible tongue shapes that can be generated by the model, without any intention to replicate the corpora used for the experimental data and without any attempt to make any link between these targets and the phonemes of a given language. Movements can include parts where the tongue is in contact with the palate, but for the majority of the sequences this did not happen.

To also evaluate to which extent the velocity gain factor of the power law accounts properly for movement velocity variations, simulations were carried out for three speaking rates: normal, fast, and slow. These sequences will be referred to hereafter as simulated data compared with the experimental data. The time courses of the motor control variables were defined by a hold duration, for which the motor commands at targets were kept constant, and by a transition duration, for which motor commands were shifted at a constant rate from one target value to the next. At a slow speaking rate the hold duration and the transition duration were both 120 ms; at a normal speaking rate they were respectively 100 and 80 ms; and at a fast speaking rate they were both 60 ms. Durations at a normal speaking rate were chosen because they allowed generating movements with durations and velocity amplitudes consistent with experimental data.

The trajectories of three flesh points located on the tongue surface in the model were considered for analysis, similar to the tongue tip, tongue blade, and tongue dorsum sensor in the experimental data. In addition, we also considered a flesh point at the surface of the tongue located in the pharyngeal region, since movements from this region could not be obtained on the basis of the experimental data. The trajectories for the simulated sequences were sampled at 125 Hz.

Data analyses

To analyze the relation between speed and curvature, the same methods were applied to experimental and simulated data. The data were first low-pass filtered with a linear-phase Remez filter (Gain 0 dB from 0 to 15 Hz; Gain < −46 dB above 30 Hz). As shown by Maoz et al. (2005), such a low-pass filter has the advantage of potentially decreasing the contribution of noise to the emergence of the power law. After filtering, the first and the second derivatives (dx/dt and d²x/dt²) were computed for each sample using the finite-differences approximation. To be accurate, this method requires a sampling frequency that is extremely large compared with the bandwidth of the signal (15 Hz for our data). To ensure accuracy, the low-pass filtered data were oversampled by a factor 10 (corresponding to a sampling frequency varying from 1,250 Hz for the simulated data to 5,000 Hz for the experimental data collected from the French
Coefficient $a$ corresponds to the exponent of the power law and coefficient $b$ to $\log_{10}(k)$, where $k$ is the velocity gain factor. The log–log linear regression approach was criticized by Schaal and Sternad (2001), who showed for high velocities the existence of a bias in the estimation of $a$ compared with nonlinear regression techniques. In spite of this, we chose to use the log–log linear regression for two reasons: 1) to facilitate the comparison with other studies, which use this method in their large majority, and 2) because maximal velocities in speech are relatively small, essentially less than 50 cm/s.

The linear regressions were calculated with the statistical software SPSS for Windows, version 15.0 (Function: Regression, linear model), for each speaker or simulation condition and each sensor or tongue region separately. For the whole set of data (All), a regression was computed for all the samples of the corpus taken together. When data were split in subsets (sentences or segments), regressions were calculated for each subset separately. Only the regressions that were statistically significant ($P < 0.01$) were taken into account for the rest of the study. For the whole set of data (All) and for the set Segment, all regressions were significant. For the data set Segment, a certain percentage of regressions (<15% in all cases) was not significant. Table 1 summarizes for each sensor the number of segments selected for each speaker or for each simulation condition, together with their mean path lengths and their mean durations.

The results were further analyzed for the different sets of data (All, Segment, Subsegment) separately. For the whole set of data (All) the analysis focused on two specific aspects: 1) the exponent values $a$ and the velocity gain factor $k$ and 2) the amount of variance ($R^2$) explained by the log–log linear regressions. When data were split in subsets (sentences or segments), the distributions of the exponent values and of the velocity gain factors calculated for all subsets (i.e., all sentences or all segments) were characterized for each sensor and for each speaker or each simulation condition separately, via their mean values and their variance. The comparison between the two sets of data, Segment and Subsegment, was based on these measures.

**RESULTS**

**Complete set of data (All)**

Distributions of complete sets of the experimental data were plotted in the $[\log_{10} c(t), \log_{10} v(t)]$ plane, split by speaker and sensor. The same representation was also done for the simulated data, split by tongue region, and by speaking rate condition. Thus 35 scatterplots (4 sensors × 5 speakers, 3 sensors for speaker F2, 4 flesh points × 3 speaking rates for the model) were available that summarized the global relations.
between speed and curvature in the log–log plane. These plots showed that the data distributions were very similar across speakers, languages, and tongue sensors for the experimental data, as well as across tongue regions, and speaking rate conditions for the simulated data. Moreover, strong similarities exist between the experimental and the simulated data. Figure 3 shows two examples of these distributions in the log–log plane, one for the simulated data (Fig. 3, top) and one for the experimental data (Fig. 3, bottom). The data are represented in light gray, whereas the best linear fit is depicted with a bold solid line. The exponent value $a$ and the explained variance ($R^2$) of the linear regression are also indicated in each plot. This figure illustrates well the main trends observed in our data. Both scatterplots look similar. The ranges of variation of speed and curvature are very close, although slightly larger for the simulated data. The exponent values and the explained variance are also very similar. In both cases it can also be observed that the proximity of the data to the linear regression line decreases when curvature becomes very small. This can be easily explained by the inaccuracy of the curvature estimation when the path becomes close to a straight line.

Figure 4 displays the results of the log–log regressions for the exponent values $a$ and for the velocity gain factors $k$, respectively, split by experimental and simulated data, by speaker, tongue sensor, tongue region, and speaking rate condition. Table 2 provides for each of the regressions the degrees of freedom ($df$) and the explained variance ($R^2$). All regressions are statistically significant ($P < 0.005$).
The exponent values are all negative, as shown for one example in Fig. 3. Figure 4 represents their absolute values. In the left plots of each panel, a horizontal dotted line marks the position of the (-1/3) key value proposed by Viviani and colleagues. Obviously, in all cases the exponent values found in our data are very close to this key value. The exact range of variation is [-0.376; -0.291] for the experimental data and [-0.377; -0.312] for the simulated data. A trend exists for the data simulated at a slow speaking rate to have slightly larger exponent values than the data simulated at normal and fast speaking rates. However, these values are still very close to (-1/3). Thus strong similarities exist for the exponent values between experimental data and simulations carried out with the biomechanical model. As observed by Tasko and Westbury (2004), differences exist for each speaker among the exponent values calculated for the different tongue sensors. However, the direction of variation across sensors is speaker dependent and no general trend can be found, except for the fact that the tongue tip sensor tends to correspond to higher exponent values than the other sensors (true for all speakers except C2). Variations among tongue regions are also systematically observed for the simulated data, but the tongue tip does not show the highest exponent values.

The velocity gain factor (Fig. 4, right panels) varies from 3.90 (subject G1, blade sensor) to 11.12 (subject F2, tip sensor) for the experimental data and from 6.25 (tip region at slow speaking rate) to 11.42 (dorsum region at fast speaking rate). Two points become evident from these observations: 1) the velocity gain factors computed for the simulated data are within the range of variation measured for the experimental data; and 2) there is a large variability across speakers and the differences between two given sensors are similar for all the tongue sensors. These findings suggest that interspeaker variability of the velocity gain factor reflects differences in speaking rate among speakers. This assumption is supported by the results obtained for the simulated data, since the velocity gain factor increases when the speaking rate increases. However, the range of variability observed for the velocity gain factor across speakers, who were all supposed to speak at a normal rate, is much larger than the range of variability observed for the simulated data, for which the speaking rate was explicitly modified.

To further analyze this question, we compared more specifically the results obtained for the different speaking rates of the simulated data, with those obtained for speakers F1 and F2, for whom the corpora used in this study were exactly the same. From speaker F1 to speaker F2 the sentence duration is multiplied by 0.82 (1.01 s for speaker F1 and 0.83 s for speaker F2), whereas the ratio for velocity gain factor is between 1.56 and 1.68, depending on the tongue sensor (see Fig. 4). For the simulated data, when the duration of the speech sequence is divided by 2 (from the slow to the fast speaking rate), the velocity gain factor is multiplied by a factor ranging from 1.3 to 1.44 (depending on the tongue region). Obviously, the relation between change in speaking rate and change in velocity gain factor is very different. This statement is consistent with the classical observation that speaking rate is not systematically related to the velocity of vocal tract articulators. Indeed, numerous studies have shown that speakers may adjust speaking rate either by modifying mainly the overall velocity of articulators or by modifying mainly the distance covered during a segment. In our simulations the first strategy was used and this is well accounted for by velocity gain factor variations. As shown by sentence durations, speakers F1 and F2 did speak at fairly similar speaking rates, but the kinematic analysis reveals that their articulators moved at different speeds: for speaker F1 the average velocities of the tongue tip, tongue blade, and tongue dorsum were respectively 6.44, 5.24, and 5.25 cm/s, whereas they were respectively 11, 10.84, and 9.27 cm/s for F2. This corresponds to a speed ratio varying between 1.71 and 2.07, which is also fairly well accounted for by velocity gain factors computed for each speaker. Thus although the velocity gain factor seems to be a reliable parameter to assess interspeaker differences in articulators’ speed, it is not adapted to measuring interspeaker differences in speaking rate because of the variety of strategies available to adjust speaking rate.

The variance explained by the log–log linear regression varies in the range from 0.522 to 0.665 for the experimental data and in the range from 0.486 to 0.580 for the simulated data. The fit is slightly better for the experimental data. However, in all cases the regressions are statistically highly significant (P < 0.005) and, given the very large number of degrees
of freedom (df, in Table 2), the amount of explained variance can be considered as fairly large.

In summary, the analysis of the complete sets of data, split by speaker, speaking rate, and tongue sensor/region suggests that the 1/3 power law applies to tongue movements. It also shows that the characteristics of the movements generated with the biomechanical tongue model controlled on a target-to-target basis, without any centrally controlled specification concerning the trajectory between the targets, closely match those of the experimental data. In addition, our simulations demonstrate that the tongue behavior in the pharyngeal region, which could not be studied experimentally, does not differ from that in the palatal region.

Data split by sentences (Sent) and segments (Seg)

It should be acknowledged that the amount of variance explained by the log–log regression in our data (~50%) is smaller than the amount of variance found in other human movements (see, e.g., de’Sperati and Viviani 1997). There are two possible and compatible explanations for this limitation: 1) the velocity gain factor varies among single movements or 2) the exponent value itself varies among single movements. To further study this question, log–log linear regressions were computed for the data split according to sentences and segments (see details in Data analyses).

In the experimental data, sentences lasted on average for 1.5 s, whereas duration of the three target sentences of the simulated data varied between 0.36 s (fast speaking rate) and 0.72 s (slow speaking rate). Because of this temporal discrepancy, no reliable comparison between simulated and experimental data was possible for the sentences. This is why data split only by segment are presented for the simulations.

Typical examples of the exponent distributions are given in Fig. 5 for the simulated data (tongue blade sensor at normal speaking rate; Fig. 5, top) and for the experimental data (speaker F1, tongue blade sensor; Fig. 5, bottom). Looking at these distributions, three general observations can be made: 1) the values are distributed around the (~1/3) key value proposed by Viviani and colleagues; 2) in the experimental data the distribution range is wider for the segments (Seg) than for the sentences (Sent) with more data points for values closer to zero (greater than −0.2); 3) for the segments, the width of the distributions of the simulated data are similar to that of the experimental data, but the exponent values tend to be more negative (with minimal values less than −0.45). Figure 6 (left and central plots of each panel) gives a more detailed account of these general observations.

The absolute values of the average exponent calculated for all the sentences (Sent) (total number of sentences: 1,253, with df varying between 40 and 1,225) and segments (Seg) (total number of segments: 11,632, with df varying between 4 and 682) are presented for each speaker or simulation condition, and for each sensor or tongue region separately (left plots). The corresponding SDs are presented in the center plots. The results are plotted with light markers for the sentences and with dark markers for the segments. As suggested by Fig. 5, it can be observed that all the average values remain in the range of the (~1/3) key value. Comparing Fig. 4 with Fig. 6, it can be noted for the experimental data that the average exponent values calculated for the sentences are very close to the exponent values calculated for the complete set of data (All). There is a trend for the average exponents calculated for the segments to have smaller absolute values than those calculated for the sentences. For the experimental data the SDs calculated for the segments are noticeably larger than those calculated for the sentences. In the majority of cases, a multiplying factor in the range of 2 to 3 is observed between the SDs of these two sets of data. For the segments the SD can be 30% of the average value.

Similar observations can be made for the velocity gain factor. The average values presented in the right plots of Fig. 6 are close to the velocity gain factors calculated for the complete sets of data (All). The SDs (given in Table 3) are generally much larger for the segments than for the sentences, with a multiplying factor of around 2 between the two sets of data.
For the sentences, the SDs of the exponent values and velocity gain factors distributions are small, in the range of 10 to 15% of their average values. These average values are very close to the values found for the complete sets of data, which describe the global speed–curvature relations of our data. $R^2$, the amount of variance explained for the sentences by the log–log linear regressions, varies between 0.265 (df = 62) and 0.857 (df = 40), with a median value at 0.596. Thus for the

FIG. 6. Absolute values of the exponent $a$ (left plots), SDs of the exponent (central plots), and velocity gain factors $k$ (right plots) computed for the data split by sentences (Sent; light symbols) and by segments (Seg; dark symbols). See Fig. 4 for additional details.
data split by sentences, $R^2$ values are quite variable, although their median value is in the range of the values calculated for the complete set of data. Thus the general trend observed for the complete sets of data also applies to each sentence individually. In other words, a power law with an exponent close to $(-1/3)$ accounts for the data at the level of individual sentences as well.

For the segments, such a power law is also found in the majority of cases, as attested by the fact that the average exponent values are close to $(-1/3)$. However, the variability of the exponent and velocity gain factor values is much larger than that for the sentences, which means that for a non-negligible number of segments the speed–curvature relation cannot be accounted for by a power law with an exponent value close to $(-1/3)$. This is true both for experimental and for simulated data, with similar amounts of variability of the exponent and velocity gain factor values within each kind of subset. $R^2$, the amount of variance explained for the segments by the log–log linear regressions, varies between 0.104 (df = 89) and 0.990 (df = 6), with a median value at 0.602. Thus contrary to our expectations, the amount of explained variance is not larger for the data split by segments than for the data split by sentences.

Do segments that do not follow the $1/3$ power law have any peculiarities that could explain this phenomenon? To answer this question, segments having exponent values larger than or equal to $-0.15$ or smaller than or equal to $-0.45$ were extracted and analyzed with respect to their duration, their average and maximal tangential velocity, and the length of their articulatory path. No evidence could be found for any of these characteristics that would make them different from the segments for which the $1/3$ power law applies. Figure 7 gives a few examples of these segments (with an exponent value larger than or equal to $-0.15$ in the two top rows and smaller than $-0.45$ in the two bottom rows). It can be seen that they feature a variety of shapes. Some of these segments are heavily curved, whereas other segments are rather straight. For some segments the curvature is quite constant along the trajectory, whereas it varies for other segments. In addition to the articulatory path, Fig. 7 presents the variation of the tangential velocity along the trajectory with the thickness of the line: the thicker the line, the faster the movement. This presentation allows one to see the detailed relation between curvature and speed along the path. It becomes evident that for all the segments the velocity decreases when the curvature increases, which is consistent with a power law with a negative exponent in the form of Eq. 2, but with an exponent value different from $(-1/3)$. The systematic analysis of their kinematic characteristics revealed that these segments often correspond to small movements with small path lengths and small tangential velocity, especially those having an exponent value greater than $-0.15$. However, this is not systematic and many segments with similar small path lengths are well described by the $1/3$ power law. Thus it is very unlikely that the inability of the $1/3$ power law to describe the speed–curvature relation in these segments could be the consequence of an artifact due to the small length of the path followed during this segment.

Finally, going back to the question that was raised at the beginning of this section, it seems that the reduced amount of explained variance found in our experimental and simulated speech data, compared with the values classically published in the literature for other human movements, could have its origin in the variability of the speed–curvature relation across individual segments. This variability is found both in the exponent values and in the velocity gain factor values.

**DISCUSSION**

On the basis of our results we suggest for speech movements three main conclusions that will be developed in the following text.

1) The $1/3$ power law accounts for general trends of speed–curvature relations; this is true in the experimental data independently of the language considered and in the movements simulated with a biomechanical model of the tongue without any centrally controlled specification of trajectory properties.

2) The $1/3$ power law does not account for the variety of the detailed time-to-time relation between speed and curvature.

3) There are results supporting the hypothesis that the global properties accounted for by the $1/3$ power law do not reflect any property of the underlying motor control strategies.

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Tongue Tip</th>
<th>Tongue Blade</th>
<th>Tongue Dorsum</th>
<th>Tongue Back/Pharynx</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Seg Sent</td>
<td>Seg Sent</td>
<td>Seg Sent</td>
<td>Seg Sent</td>
</tr>
<tr>
<td>C1</td>
<td>2.56 1.05</td>
<td>1.95 0.81</td>
<td>1.84 0.86</td>
<td>1.68 0.88</td>
</tr>
<tr>
<td>C2</td>
<td>1.97 1.01</td>
<td>1.83 1.02</td>
<td>1.79 0.91</td>
<td>1.87 0.87</td>
</tr>
<tr>
<td>G1</td>
<td>1.01 0.50</td>
<td>1.25 0.51</td>
<td>1.31 0.54</td>
<td>1.47 0.47</td>
</tr>
<tr>
<td>G2</td>
<td>1.74 0.85</td>
<td>1.31 0.88</td>
<td>1.54 0.94</td>
<td>1.47 0.71</td>
</tr>
<tr>
<td>F1</td>
<td>1.47 1.00</td>
<td>1.44 0.77</td>
<td>1.56 1.00</td>
<td>1.91 1.54</td>
</tr>
<tr>
<td>F2</td>
<td>2.66 1.45</td>
<td>2.27 1.12</td>
<td>2.03 0.95</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 3.** SDs of the distribution of the velocity gain factor calculated for the sentences (Sent) and the segments (Seg) for the experimental data and for the segments only for the simulated data

---

*J Neurophysiol* • VOL 100 • SEPTEMBER 2008 • www.jn.org
Experimental devices. The simulated data provided evidence that this behavior is consistent across subjects and for the three languages that were analyzed as well as for the interspeaker comparison. The speed–curvature relations for the sentences and for the complete sets of data are consistent with the law. In addition, for each subject or modeling condition, and for each tongue sensor or tongue region, the average exponent value is negative, which means that the velocity decreases with the curvature of the articulatory path. This supports Viviani and colleagues’ hypothesis that this coefficient accounts for the nearly second-order dynamical characteristics of the articulatory paths and without any centrally controlled specification of the trajectory properties or velocity profiles. The simulations carried out with the biomechanical tongue model for speech support Gribble and Ostry’s hypothesis that the power law can be explained by the characteristics of the data simulated with our model were controlled on a target-to-target basis without any optimization based on kinematic characteristics of the articulatory paths and without any centrally controlled specification of the trajectory properties or velocity profiles. The detailed time-to-time relations between speed and curvature in speech movements can be analyzed on the basis of the log–log linear regressions carried out separately for each individual segment (Figs. 5 and 6, and Table 3). Both for experimental and simulated data, the distributions of the slope of the log–log regression, which corresponds to the exponent value of the power law relation, show large SDs. Obviously, this observation is not compatible with a law assuming a constant exponent value that applies for each speech movement on a time-to-time basis.

**1/3 power law and motor control strategies**

Our results show that the general trends accounted for by the 1/3 power law are observed both for the experimental and for the simulated data. It is important to recall here that the simulations carried out with the biomechanical tongue model were controlled on a target-to-target basis without any optimization based on kinematic characteristics of the articulatory paths and without any centrally controlled specification of the trajectory properties or velocity profiles. The simulations carried out with the biomechanical tongue model for speech support Gribble and Ostry’s hypothesis that the power law can be explained by the nearly second-order dynamical characteristics of the motor system. Tongue stiffness (i.e., the Young modulus of tongue tissues) varies in the model during the course of a movement due to muscle activation (for details see Payan et al. 2005).
and Perrier 1997). It could, at least in part, be at the origin of the variability observed in the exponent value across individual segments around the theoretical (−1/3) value. Indentation measurements carried out on a fresh human cadaver tongue have shown that even more complex nonlinear relations exist between stress and strain in real tongue tissues (Gérard et al. 2005), which can induce noticeable variations in the stiffness during tongue displacements. Larger variation in the elastic characteristics of the tongue compared with other motor systems could explain why the variability of the exponent values measured in this study for speech is larger than that previously observed for other human movements.

According to Maoz and colleagues (2005), noise in the data could largely contribute to the emergence of the 1/3 power law. Their demonstration is based on the formula used to compute the curvature (see Eq. 4) that, de facto, induces that \( \log_{10} [\nu(t)] \) and \( \log_{10} [c(t)] \) are linked by the equation

\[
\log_{10} [\nu(t)] = -\frac{1}{3} \log_{10} [c(t)] + \frac{1}{3} \log_{10} \left( \frac{d^2x}{dt^2} \frac{dy}{dt} - \frac{d^2y}{dt^2} \frac{dx}{dt} \right) \tag{6}
\]

If, due to noise, the term

\[
\frac{1}{3} \log_{10} \left( \frac{d^2x}{dt^2} \frac{dy}{dt} - \frac{d^2y}{dt^2} \frac{dx}{dt} \right)
\]

is statistically not correlated with \( \log_{10} [\nu(t)] \), the 1/3 power law directly emerges from Eq. 6. Given that the accuracy of a correlation estimate between two stochastic processes increases with the number of measurements, the fact that the exponent values estimated from the complete sets of data tend to be closer to (−1/3) than the values estimated from the segments (Seg) supports Maoz and colleagues’ suggestion.

In conclusion, our study combining the analysis of articulatory data from three different languages and the analysis of articulatory data generated with a biomechanical tongue model provides evidence that the 1/3 power law is a very global characteristic of the speed–curvature relations in speech movements and that it is language independent. The fact that it applies to movements simulated with a biomechanical model of the tongue—driven without any centrally controlled specification of trajectory properties—suggests that the 1/3 power law does not result from any motor control strategy. In addition, this law does not apply to every single movement description. Thus no reliable inference can be made regarding speech motor control on the basis of the 1/3 power law.

ACKNOWLEDGMENTS

We thank C. Savariaux and J. Dreyer for technical expertise during the EMMA (electromagnetic midsagittal articulometer) recordings, at Institut de la Communication Parlée and at the Center for General Linguistics, and the two anonymous reviewers for very helpful comments. Thanks to our colleagues C. Mooshammer, J. Dang, and L. Ma who helped collect the experimental data for other research purposes, to L. Koenig for suggestions and proofreading, to U. Bielh for technical assistance, and to D. Fuchs.

GRANTS

This work was supported by the French-German “POPAART” project funded by the Centre National de la Recherche Scientifique, the French Ministry of Foreign Affairs in the framework in the P2R program, and German Research Council Grant GWZ 4/8, P1.

REFERENCES


J Neurophysiol • VOL 100 • SEPTEMBER 2008 • www.jn.org

Downloaded from http://jn.physiology.org/ by IP 10.220.33.4 on October 28, 2016