HIGHER-ORDER KINEMATIC ANALYSIS OF SPEECH MOVEMENT DATA

Stephan R. Kuberski and Adamantios I. Gafos
University of Potsdam, Department of Linguistics and Cognitive Sciences

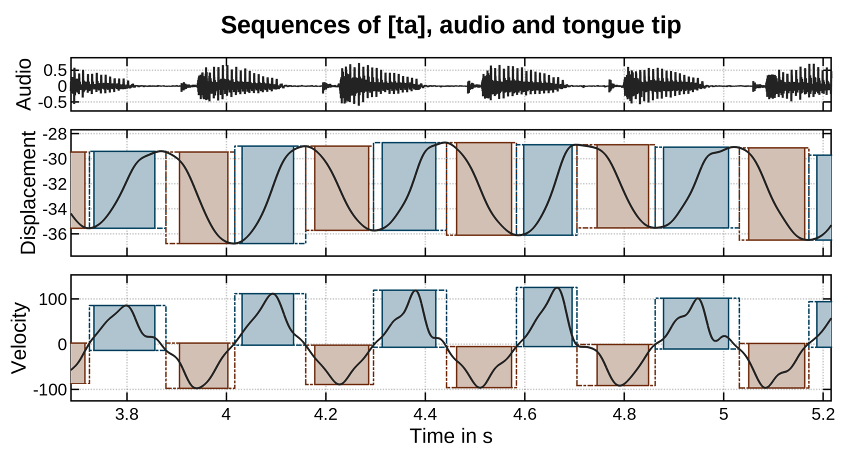
SFB 1287 Limits of Variability in Language — Cognitive, Grammatical, and Social Aspects

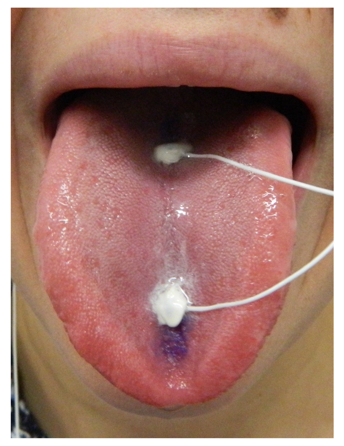
Universität Potsdam

INTRODUCTION

We present techniques and results of a higher-order kinematic analysis of speech movement data registered by contemporary Electromagnetic Articulography (EMA). In particular, we first demonstrate the applicability of a well-established in the human movement field (but in speech rarely used) spline-smoothing approach and illustrate its superiority over traditional signal representations. Second, using a heptic spline-smoothing approach, we reveal a so far unknown set of acceleration-based kinematic relations in data of repetitive speech.

In a task where participants repeated syllables in time with a metronome, we registered speech movement data of approximately 10.000 /ka/ and /ta/ syllables from 10 native speakers of German and English (3+3 female, 2+2 male). We implemented an extensive speech rate manipulation (5 metronome rates, 150–480 bpm, covering slow, normal and fast speech) to elicit vocal tract actions from a wider kinematic spectrum than before (cf. Kelso et al., 1985; Ostry et al., 1987; Patel et al., 1999).





Latest generation Electromagnetic Articulography (AG501, Carstens) was used to track the three-dimensional articulatory motion at high spatial and temporal resolution (0.3–0.5 mm RMS at 1250 Hz). For each of the two primary articulators involved (tongue body for /ka/, tongue tip for /ta/ syllables), we performed a kinematic analysis using two distinct smoothing approaches.

SMOOTHING TECHNIQUES

Despite the outstanding spatio-temporal resolution of EMA, articulatory displacement data returned by this method are substantially contaminated by noise. Moreover, kinematic analysis requires the evaluation of quantities not directly measurable by the device (e.g., velocity, acceleration). Hence, there is need for an appropriate smoothing and approximation approach of noisy displacement data and their derivatives.

Traditional filtering

In speech, one of the most often used smoothing techniques is that of digital filtering. Numerical differentiation is generally carried out by means of finite differences. As a typical exemplar of this approach, we filtered our data by a 3rd-order Butterworth low-pass with 20 Hz cutoff. Velocity and acceleration estimates were computed using a central difference scheme with an additional intermediary 5-point average filter, in case of acceleration.

Spline-smoothing

As an alternative smoothing and differentiation approach, we approximated our data by means of smoothing splines. More specifically, we took recourse to Woltring (1986)'s classic spline-smoothing and differentiation code. In other fields of human motion analysis, spline-smoothing is considered to be the adequate choice when performing kinematic analyses (see Wood, 1982; Woltring, 1985; Medved, 2001 with overviews of different smoothing techniques). Concisely summarized, a smoothing spline $s_p(t)$ is a piecewise polynomial approximation of n noisy measurements x_i at times t_i which minimizes the criterion function C_p for a suitably selected regularization parameter p

$$C_p = \sum_{i=1}^{n} [x_i - s_p(t_i)]^2 + p \int_{t=-\infty}^{+\infty} |s_p^{(m)}(t)|^2 dt.$$

The parameter p can be chosen such that the resulting spline $s_p(t)$ shows deviation from (x_i, t_i) with a fixed predicted mean-squared error (Wahba, 1979; RMS, as specified by the manufacturer of the EMA device). For this presentation, we chose a heptic spline representation (i.e., half order $m = 4$) resulting in 8th-order polynomials with smooth derivatives up to 6th order.

PROJECT FUNDING

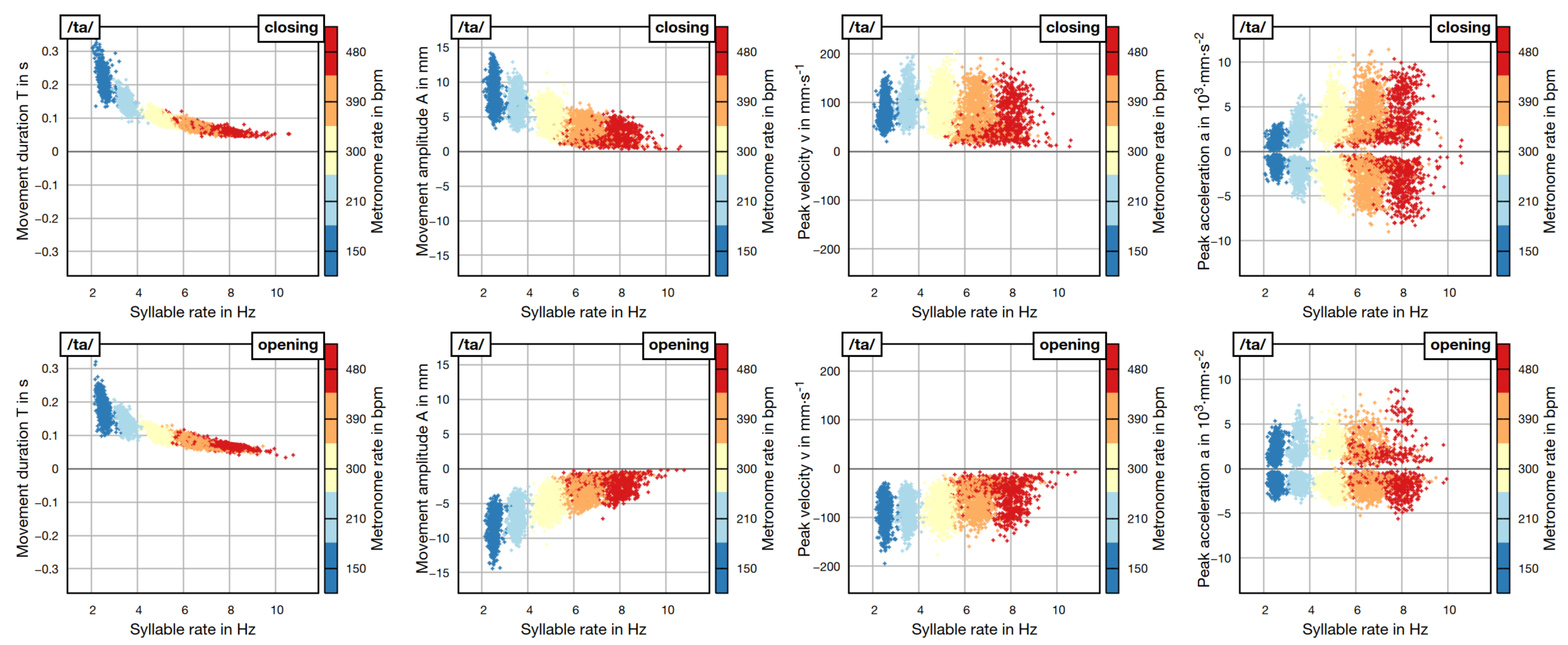
This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): Project ID 317633480, SFB 1287, Project C04. Additional financial support has been received from the European Research Council (ERC): Grant agreement ID 249440.

KINEMATIC PARAMETERS

The continuous motion of the two primary articulators involved (tongue body for /ka/, tongue tip for /ta/ syllables) was segmented into successive closing (constriction formation) and opening (constriction release) movements using a zero-velocity criterion. A speech movement was thus defined to be a period of an articulator's motion from one position of rest (quasi-steady state) to another, during which velocity progressively rises (acceleration phase) up to some maximal value (peak velocity) and declines again (deceleration phase) homing in on the aimed target. For each of the so-determined movements, we computed the following set of kinematic parameters (cf. Ostry et al., 1983; Ostry and Munhall, 1985; Vatikiotis-Bateson and Kelso, 1993):

Quantity	**Symbol**	**Unit**	**Sign convention** closing	opening
Movement duration	T	s	+	+
Movement amplitude	A	mm	+	−
Peak velocity	v	mm/s	+	−
Peak acceleration/deceleration	a	mm/s^2	+/−	−/+

From durations of neighboring pairs of movements, we additionally estimated the instantaneous rate of syllable production (syllables per second in Hz) as the factual outcome of the imposed metronome rate. The following figure gives an overview of the kinematic parameters for syllables of /ta/ in the spline-smoothing approach, separated by closing (top) and opening direction (bottom row):



In order to compare the kinematic parameter estimates (whose true values are a priori unknown) of the two considered smoothing techniques, we examined the relative percent difference (RPD) of each pair of measurement values. The RPD of two measurements x and y is defined as $2(x - y)/(x + y)$ and expresses their relative difference with respect to the mean of x and y. The result of this comparison is given in the following table:

Average RPD in %	**/ka/ syllables** closing	opening	**/ta/ syllables** closing	opening
Movement duration	1.00	-1.07	-0.64	0.24
Movement amplitude	-0.49	-0.55	0.42	0.34
Peak velocity	2.11	2.92	7.96	3.84
Peak acceleration	4.65	10.09	12.42	15.23
Peak deceleration	14.54	7.49	26.59	18.59

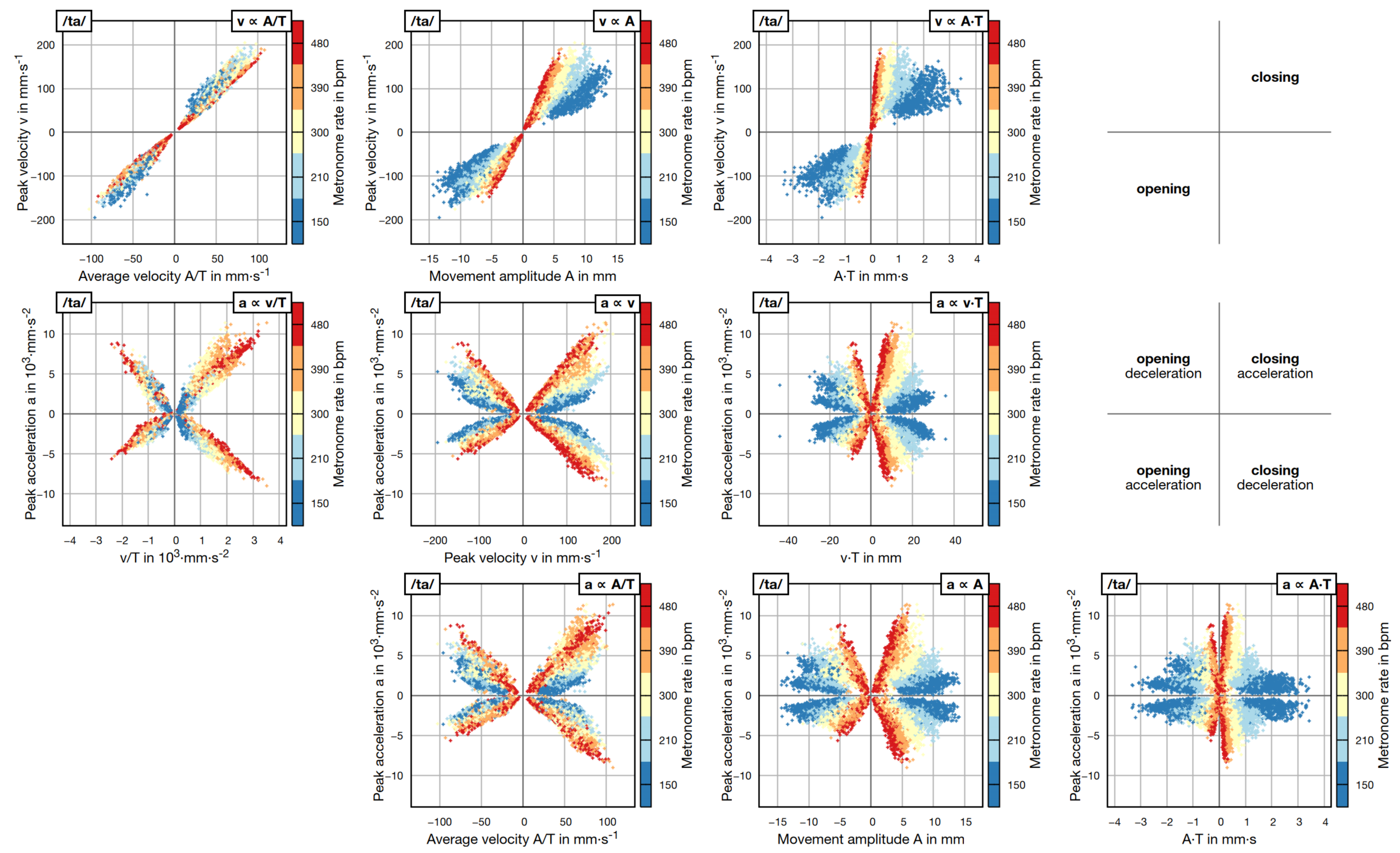
Overall, magnitudes of the time derivative estimates show consistently higher values (positive RPD) for traditional filtering than spline-smoothing: acceleration and deceleration estimates are substantially larger by a range of 5–15% RPD (corresponding to a factor of 1.05–1.16) and 7–27% RPD (factor 1.07–1.31), respectively. Velocity estimates are moderately larger: 2–8% RPD (factor 1.02–1.08). Duration and amplitude estimates are approximately on par (RPD in the range of ±1%).

KINEMATIC RELATIONS

Let $\omega = \pi/T$ be the natural frequency of the articulatory motion (the alternation of closing and opening movements). From first principles of dimensional analysis (e.g., Buckingham, 1914), we derived the following set of theoretically possible relations between the kinematic parameters of T, A, v and a with slopes proportional to different powers of ω:

ω^0-group	ω^1-group	ω^2-group	ω^3-group
$v \propto A/T$	$v \propto A$	$v \propto A \cdot T$	
$a \propto v/T$	$a \propto v$	$a \propto v \cdot T$	
	$a \propto A/T$	$a \propto A$	$a \propto A \cdot T$

This set of relations greatly extends the number of kinematic relations well-known and well-studied in speech (which are $v \propto A/T$ and $v \propto A$, e.g., in Ostry et al., 1983; Ostry and Munhall, 1985; Vatikiotis-Bateson and Kelso, 1993; Fuchs et al., 2011). For our data, the following figure demonstrates the empirical presence of all these theoretically predicted relations (shown is only /ta/ in the spline-smoothing approach; for /ka/ and the filtering approach the results are qualitatively the same):



The clear empirical structure of these relations (linear relations with slopes proportional to different powers of ω) provided an effective way to rigorously assess the performances of the two distinct smoothing techniques. In a linear regression analysis, we tested both approaches for differences in their regression standard errors (which are the root mean squares of the regression residuals) and obtained the following result:

Regression standard error	ω^0-group filter	spline	ω^1-group filter	spline	ω^2-group filter	spline	ω^3-group filter	spline
/ka/ syllables	0.2703	0.2426	2.113	1.781	20.97	16.74		
	0.9197	0.6880	7.233	5.368	71.71	53.31		
			17.63	13.14	159.7	118.4	1779	1334
/ta/ syllables	0.2775	0.2310	2.506	1.973	27.88	21.12		
	1.063	0.7680	9.223	6.478	100.9	70.30		
			23.28	16.56	243.4	171.0	2998	2060

For all the kinematic relations tested, the regression standard errors are significantly smaller (25% on average) for the smoothing splines. This result markedly underpins the superiority of the spline-smoothing approach over traditional filtering in the domain of speech.

REFERENCES

Buckingham (1914). "On physically similar systems; illustrations of the use of dimensional equations." *Phys Rev* 4(4).

Fuchs, Perrier, and Hartinger (2011). "A critical evaluation of gestural stiffness estimations in speech production based on a linear second-order model." *J Speech Lang Hear Res* 54(4).

Kelso, Vatikiotis-Bateson, Saltzman, and Kay (1985). "A qualitative dynamic analysis of reiterant speech production: phase portraits, kinematics, and dynamic modeling." *J Acoust Soc Am* 77(1).

Medved (2001). *Measurement of human locomotion*. CRC.

Ostry, Cooke, and Munhall (1987). "Velocity curves of human arm and speech movements." *Exp Brain Res* 68(1).

Ostry, Keller, and Parush (1983). "Similarities in the control of the speech articulators and the limbs: kinematics of tongue dorsum movement in speech." *J Exp Psychol Hum Percept Perform* 9(4).

Ostry and Munhall (1985). "Control of rate and duration of speech movements." *J Acoust Soc Am* 77(2).

Patel, Löfqvist, and Naito (1999). "The acoustics and kinematics of regularly timed speech: a database and method for the study of the p-center problem." *ICPhS*.

Vatikiotis-Bateson and Kelso (1993). "Rhythm type and articulatory dynamics in English, French and Japanese." *J Phon* 21.

Wahba (1979). "Smoothing and ill-posed problems." *Solution Methods for Integral Equations*. Springer.

Woltring (1985). "On optimal smoothing and derivative estimation from noisy displacement data in biomechanics." *Hum Mov Sci* 4(3).

Woltring (1986). "A Fortran package for generalized, cross-validatory spline smoothing and differentiation." *Adv Eng Softw* 8(2).

Wood (1982). "Data smoothing and differentiation procedures in biomechanics." *Exerc Sport Sci Rev* 10(1).