7th SMC
July 5-8, 2017
Groningen

DYNAMICS AND KINEMATICS OF REPETITIVE SPEECH MOVEMENTS

Stephan R. Kuberski, Adamantios I. Gafos
University of Potsdam, Linguistics Department and Excellence Area in Cognitive Science, Potsdam, Germany

AIMS

Several dynamical models of speech movements and gestures have been proposed in the field of speech motor control. Some others, from the field of general motor control, have not been evaluated nor applied in the domain of speech yet.

One significant distinction among all the models consists in terms of their topology. In the past, models using solely fixed point dynamics tended to dominate the field of dynamical modeling of (discrete) movements. More recently, a shift towards models containing more sophisticated types of topology (e.g., multiple fixed points, limit cycle attractors) is noticeable.

We aim to assess these models using a paradigm with a more systematic speech rate control than in previous attempts. Here we present an overview of our current assessment methods based on two datasets of repetitive speech.

MODEL EVALUATION

Using the **standard model** (Saltzman and Munhall, 1989), an isolated single-articulator movement can be described by a linear second order system with critical damping:

$$\ddot{x} = -kx - b\dot{x}\,, \quad b = 2\zeta\sqrt{k}\,, \quad \zeta = 1\,. \tag{1}$$

Dynamics and kinematics of these movements are fully determined by initial conditions and model parameters of **stiffness** (k) and **damping** (b). Analytical treatment yields the following proportions between peak velocity v^*, movement amplitude A and movement duration T (with substitution $\omega_0 = \sqrt{k}$):

$$\frac{v^*}{A} = \frac{c\pi}{T}\,, \quad T = \text{const.}\,, \quad v^* = c\omega_0 A\,. \tag{2}$$

Certain model **predictions are exploitable** to check consistency with experimental data. First, the proportionality constant c is maximal for the undamped case ($c = 0.5$, see also Fuchs et al., 2011). Second, numerical simulation (Figure 1) for varying initial conditions, natural frequencies and damping ratios reveals a maximal relative time to peak velocity value (RTTP) of 0.5 (also in case of absent damping, bottom left panel).

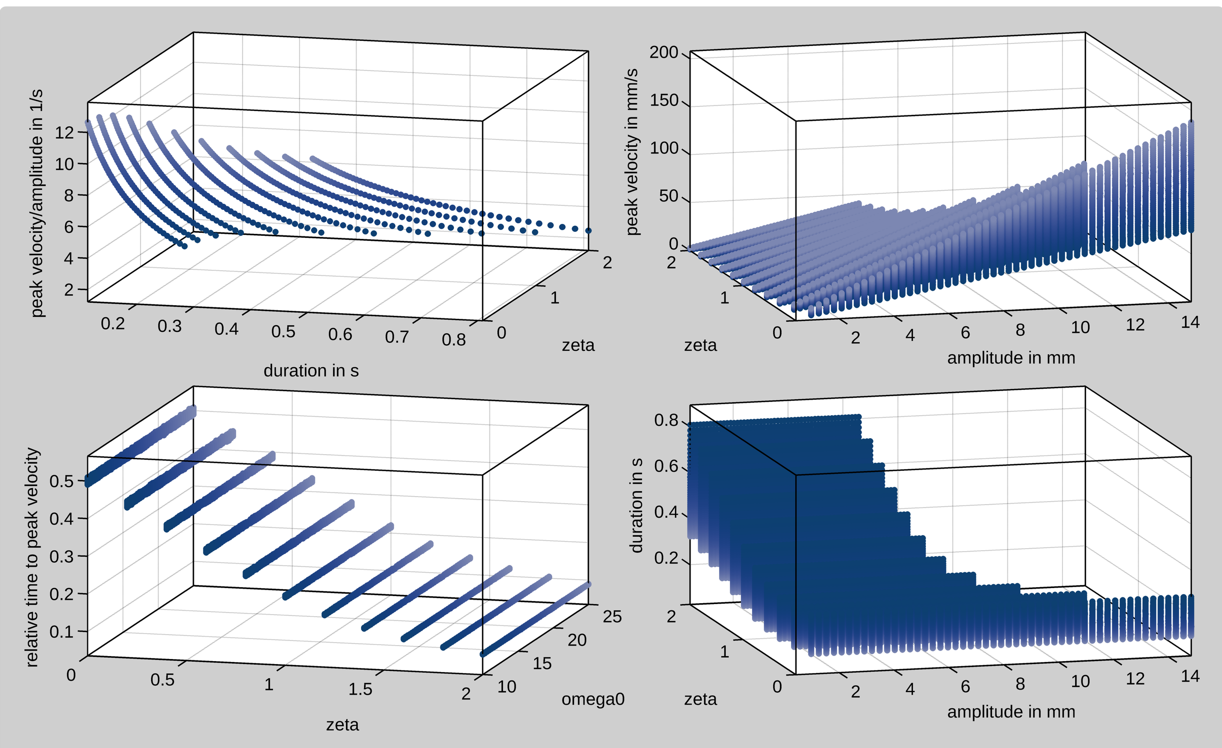

Figure 1: Simulation of the linear second order model with predicted kinematic relations as in (2). Natural frequency $\omega_0 = \sqrt{k}$ is color-coded from dark (small values) to light blue (large values).

Other, nonautonomous or nonlinear, fixed point models have been proposed (e.g., Kröger et al., 1995; Sorensen and Gafos, 2016). In terms of model evaluation, for current purposes these can be treated equivalently to the linear second order model.

In contrast to single fixed point models, there also exist models with **more advanced dynamical topologies** (containing additional limit cycle regimes: e.g., Schöner, 1990). These models, able to inherently render both sequential and repetitive movements, are in principle applicable to speech but have not been applied there yet.

One major prediction of multi-stability models is a **qualitative change in model behaviour** (bifurcation) at some critical parameter value. Using a finger flexion-extension paradigm and simulations with the Jirsa and Kelso (2005) model, Huys et al. (2008) argued that fast rhythmic (finger) movements are governed by a qualitatively different dynamical regime, a limit cycle, than discrete movements and that the **switch to the limit cycle** can be induced by increasing the pace of their finger flexion-extension task. So far, evidence for such a transition in a speech task is lacking. We pursued this prediction in the domain of speech in our Potsdam KORSA pilot study.

HARVARD-HASKINS DATASET

The Harvard-Haskins dataset of **regularly timed speech** (Patel et al., 1999) contains electromagnetic articulometer (EMA) displacement data of tongue, lips and jaw movements during production of various sequences of the form /baCa/. For the subset of our interest (/baba.../), recordings were taken from three adult native English speakers, each articulating 11 syllables in 4 trials at a **self-paced rate** being trained priorly at a rate of 2 Hz = 120 bpm.

Principal component analysis of **jaw movements** of this subset reveals a strong asymmetry between opening and closing movements, namely:

	openings	closings
duration	low	high
peak velocity	high	low
stiffness	high	low
RTTP variability	low	high

1. Opening movements are shorter (Figure 2, left panel) and faster (middle panel) than closing movements.
2. Standard model stiffnesses differ significantly for openings ($k = 448.1\,\text{s}^{-2}$, $c = 0.513$) and closings ($k = 134.4\,\text{s}^{-2}$, $c = 0.503$) (left and middle panels, cf. equation (2)).
3. Relative time to peak velocity values (RTTP) show higher variability for closing than for opening movements (right panel).

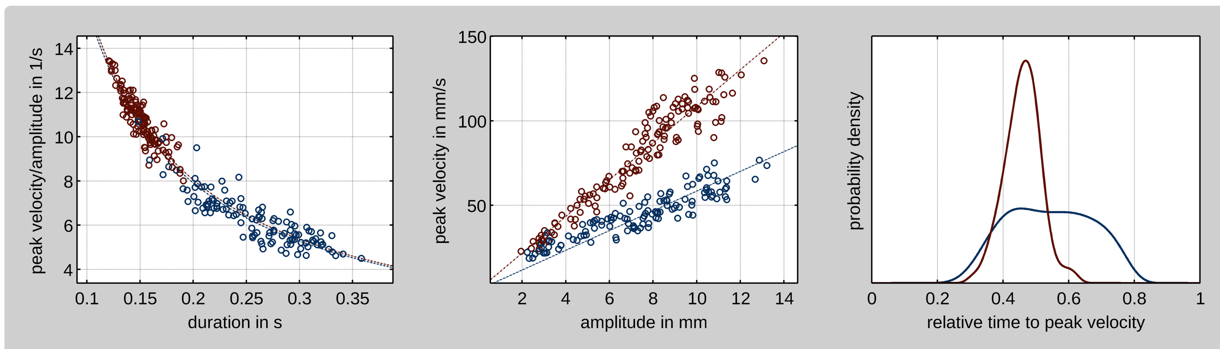

Figure 2: Kinematic characteristics of jaw movements (openings: red, closings: blue) during regularly self-paced production of /baba.../ (approx. 2 Hz = 120 bpm).

Some of the kinematic relations (left and middle panels) are **compatible with the standard model predictions** in equation (2) and Figure 1 (both top panels). In contrast, the large RTTP values (> 0.5) found specifically for closing movements conflicts with model predictions (Figure 1, bottom left panel). Perkell et al. (2002) previously also reported such **high RTTP values** (corresponding to unbalanced velocity profiles) and associated these in general with movements at speech rates below fast.

POTSDAM KORSA PILOT

In our recent pilot study, using the general paradigm of **repetitive speech** (cf. Kelso et al., 1985; Ostry et al., 1987; Patel et al., 1999), we sought evidence (in speech) of the sort Huys et al. (2008) have uncovered for finger movements. We recorded EMA data of tongue and jaw movements at **systematically controlled speech rates** (by metronome). A single adult native English speaker repeated sequences of 15 to 30 syllables in 4 trials for each rate condition. Sequences were of the form /CVCV.../ and /CVC.../.

Principal component analysis of **tongue tip movements** in /tata.../ at low and common speech rates (below 210 bpm = 3.5 Hz) corroborates the asymmetry between opening and closing movements also found in the Harvard-Haskins dataset. Additionally, analysis reveals a rate dependency of kinematic properties and a **qualitative change of disappearing asymmetries** at a rate of approx. 210 bpm (Figure 3):

1. Velocity profiles' dip test p-values (a measure for probability of unimodality, top panel) change from bi-/multimodal (low p-values) to unimodal profiles (high p-values) at critical rate.
2. Below critical rate, opening and closing movements substantially differ with respect to peak velocity and RTTP (middle and bottom panels). Above critical rate, these properties approximate each other.
3. Below critical rate, RTTP variability of closing movements is substantially higher than that of opening movements (bottom panel).
4. Both openings and closings converge to the extremal RTTP value of 0.5 of the undamped harmonic oscillator (bottom panel) indicating symmetrical velocity profiles and minimal energy consumption (Nelson, 1983).

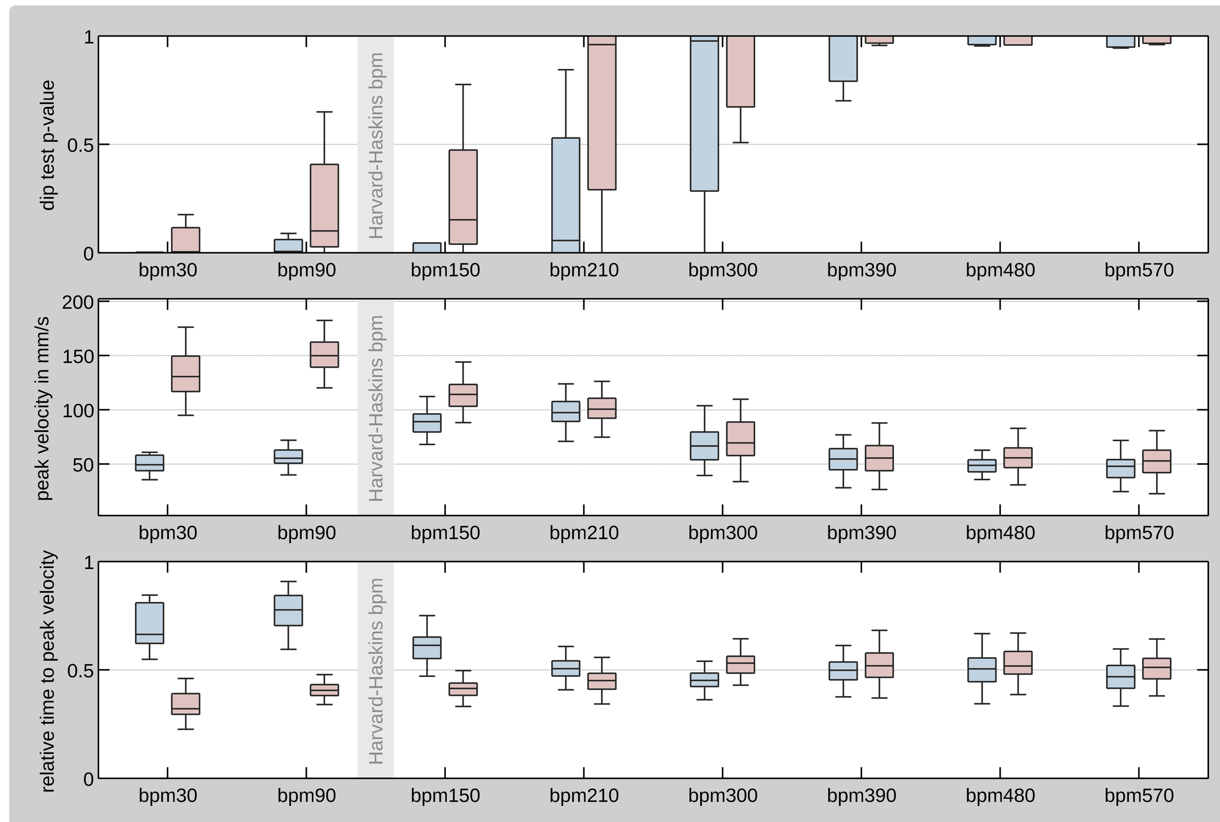

Figure 3: Rate dependencies of tongue tip movements (openings: red, closings: blue) in velocity profile form (top panel), peak velocity (mid panel) and RTTP (bottom panel) during metronome-paced production of /tata.../ (60 bpm = 1 Hz).

Observed asymmetries between openings and closings up to the critical speech rate follow those also found in the Harvard-Haskins dataset. Beyond the critical rate, kinematic properties converge and asymmetries disappear. This qualitative change might indicate the **existence of a topological bifurcation** in the underlying dynamical model.

There is evidence of similar patterns in other subsets of our pilot data (e.g., /kaka.../, /titi.../). One of our aims in future work is the development of tools to reliably diagnose the existence of a rate dependent bifurcation and determine its topological structure.

REFERENCES

Fuchs, Susanne, Pascal Perrier, and Mariam Hartinger (2011). "A critical evaluation of gestural stiffness estimations in speech production based on a linear second-order model". In: *Journal of Speech, Language, and Hearing Research* 54 (4).

Huys, Raoul, Braenna E. Studenka, Nicole L. Rheaume, Howard N. Zelaznik, and Viktor K. Jirsa (2008). "Distinct timing mechanisms produce discrete and continuous movements". In: *PLoS Computational Biology* 4 (4).

Jirsa, Viktor K. and J. A. Scott Kelso (2005). "The excitator as a minimal model for the coordination dynamics of discrete and rhythmic movement generation". In: *Journal of Motor Behaviour* 37 (1).

Kelso, J. A. Scott, Eric Vatikiotis-Bateson, Elliot L. Saltzman, and Bruce Kay (1985). "A qualitative dynamic analysis of reiterant speech production: phase portraits, kinematics, and dynamic modeling". In: *The Journal of the Acoustical Society of America* 77 (1).

Kröger, Bernd J., Georg Schröder, and Claudia Opgen-Rhein (1995). "A gesture-based dynamic model describing articulatory movement data". In: *The Journal of the Acoustical Society of America* 98 (4).

Nelson, Winston L. (1983). "Physical principles for economies of skilled movements". In: *Biological Cybernetics* 46 (2).

Ostry, David J., James D. Cooke, and Kevin G. Munhall (1987). "Velocity curves of human arm and speech movements". In: *Experimental Brain Research* 68 (1).

Patel, Aniruddh D., Anders Löfqvist, and Walter Naito (1999). "The acoustics and kinematics of regularly timed speech: a database and method for the study of the P-center problem". In: *Proceedings of the 14th International Congress of Phonetic Sciences*. (San Francisco). Vol. 1.

Perkell, Joseph S., Majid Zandipour, Melanie L. Matthies, and Harlan Lane (2002). "Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues". In: *The Journal of the Acoustical Society of America* 112 (4).

Saltzman, Elliot L. and Kevin G. Munhall (1989). "A dynamical approach to gestural patterning in speech production". In: *Ecological Psychology* 1 (4).

Schöner, Gregor (1990). "A dynamic theory of coordination of discrete movement". In: *Biological Cybernetics* 63 (4).

Sorensen, Tanner and Adamantios I. Gafos (2016). "The gesture as an autonomous nonlinear dynamical system". In: *Ecological Psychology* 28 (4).