DYNAMICS AND KINEMATICS OF REPETITIVE SPEECH MOVEMENTS

Adamantios I. Gafos, Stephan R. Kuberski

University of Potsdam
Linguistics Department
Excellence Area in Cognitive Science
Potsdam, Germany

3rd International Workshop on Dynamical Modeling
18-19 July, 2017 Cologne, Germany

Linear model kinematics (1)

- using the **standard model** of speech gestures (Saltzman and Munhall, 1989) we can model a (one-dimensional) **single articulator movement** by the linear second order system

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

- model parameters **stiffness** k and **damping** b can be defined in terms of a dimensionless damping ratio $\zeta \geq 0$ and natural frequency $\omega_0 = \sqrt{k}$
- we can solve the differential equation yielding expressions for **displacement**, **velocity** and **acceleration** (dependent on damping ratio)
- e.g., for **critical damping** ($\zeta = 1$) with initial conditions $(x_0, \dot{x}_0)$

$$\begin{aligned} & x(t) = \mathrm{e}^{-\omega_0 t}\,(x_0 + [\omega_0 x_0 + \dot{x}_0]t) \\ \Rightarrow \quad & \dot{x}(t) = \mathrm{e}^{-\omega_0 t}\,(\dot{x}_0 - \omega_0[\omega_0 x_0 + \dot{x}_0]t) \\ \Rightarrow \quad & \ddot{x}(t) = \mathrm{e}^{-\omega_0 t}\left(-2\omega_0\dot{x}_0 - \omega_0^2 x_0 + \omega_0^2[\omega_0 x_0 + \dot{x}_0]t\right) \end{aligned}$$

Linear model kinematics (1)

- using the **standard model** of speech gestures (Saltzman and Munhall, 1989) we can model a (one-dimensional) **single articulator movement** by the linear second order system

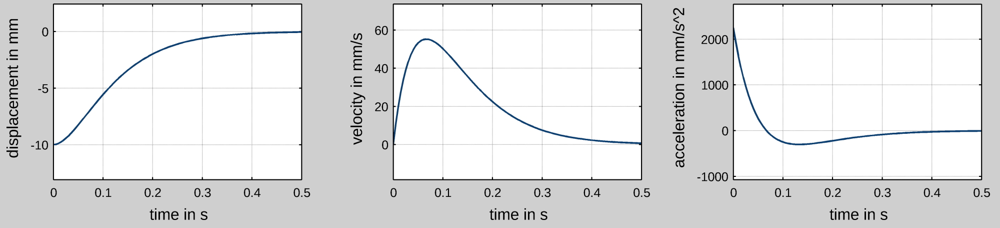
$$\ddot{x} = -kx - b\dot{x}\,, \quad b = 2\zeta\sqrt{k}$$

- model parameters **stiffness** k and **damping** b can be defined in terms of a dimensionless damping ratio $\zeta \geq 0$ and natural frequency $\omega_0 = \sqrt{k}$
- we can solve the differential equation yielding expressions for **displacement**, **velocity** and **acceleration** (dependent on damping ratio)
- e.g., for **critical damping** ($\zeta = 1$) with initial conditions (x_0, $\dot{x}_0$)

$$\begin{aligned}
& x(t) = \mathrm{e}^{-\omega_0 t}\,(x_0 + [\omega_0 x_0 + \dot{x}_0]t) \\
\Rightarrow \quad & \dot{x}(t) = \mathrm{e}^{-\omega_0 t}\,(\dot{x}_0 - \omega_0[\omega_0 x_0 + \dot{x}_0]t) \\
\Rightarrow \quad & \ddot{x}(t) = \mathrm{e}^{-\omega_0 t}\,\big(-2\omega_0\dot{x}_0 - \omega_0^2 x_0 + \omega_0^2[\omega_0 x_0 + \dot{x}_0]t\big)
\end{aligned}$$

Linear model kinematics (2)

- example curves of **displacement** $x(t)$, **velocity** $\dot{x}(t)$ and **acceleration** $\ddot{x}(t)$ for the **critically damped** linear system ($x_0 = -10\,\mathrm{mm}$, $\dot{x}_0 = 0\,\mathrm{mm/s}$, $k = 225\,\mathrm{s}^{-2}$)

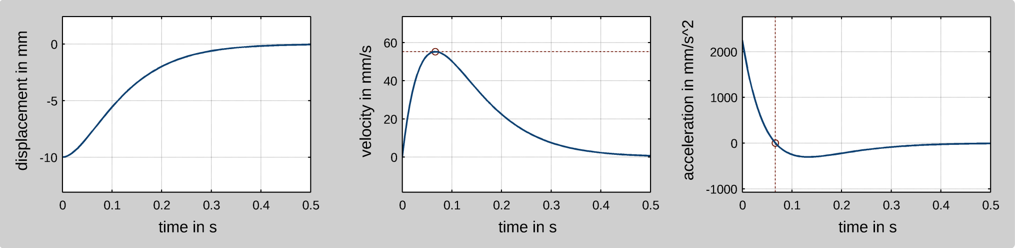


- a simple model of an articulatory movement defines the **kinematic variables**:

1. **peak velocity** v^* at instant of zero acceleration
2. movement onset and offset at instants of a certain fraction of peak velocity
3. movement **amplitude** A and **duration** T follow straight forward
4. as well as **relative time to peak velocity** (RTTP)

Linear model kinematics (2)

- example curves of **displacement** $x(t)$, **velocity** $\dot{x}(t)$ and **acceleration** $\ddot{x}(t)$ for the **critically damped** linear system ($x_0 = -10\,\text{mm}$, $\dot{x}_0 = 0\,\text{mm/s}$, $k = 225\,\text{s}^{-2}$)

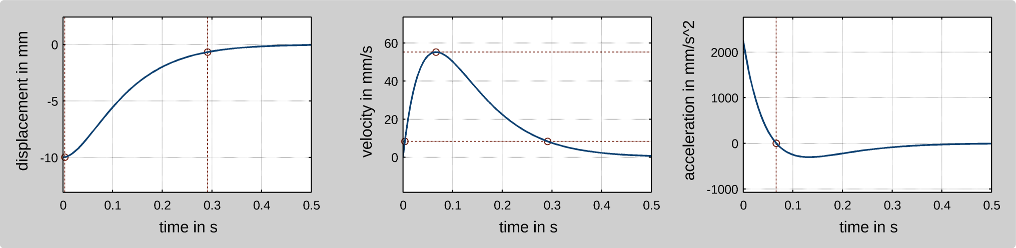


- a simple model of an articulatory movement defines the **kinematic variables**:

1. **peak velocity** v^* at instant of zero acceleration
2. movement onset and offset at instants of a certain fraction of peak velocity
3. movement **amplitude** A and **duration** T follow straight forward
4. as well as **relative time to peak velocity** (RTTP)

Linear model kinematics (2)

- example curves of **displacement** $x(t)$, **velocity** $\dot{x}(t)$ and **acceleration** $\ddot{x}(t)$ for the **critically damped** linear system ($x_0 = -10\,\text{mm}$, $\dot{x}_0 = 0\,\text{mm/s}$, $k = 225\,\text{s}^{-2}$)

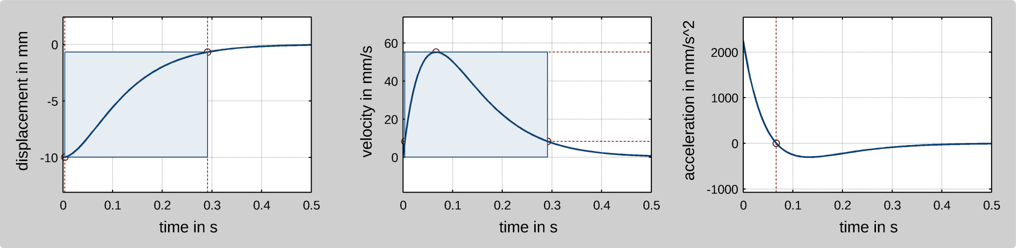


- a simple model of an articulatory movement defines the **kinematic variables**:

1. **peak velocity** v^* at instant of zero acceleration
2. movement onset and offset at instants of a certain fraction of peak velocity
3. movement **amplitude** A and **duration** T follow straight forward
4. as well as **relative time to peak velocity** (RTTP)

Linear model kinematics (2)

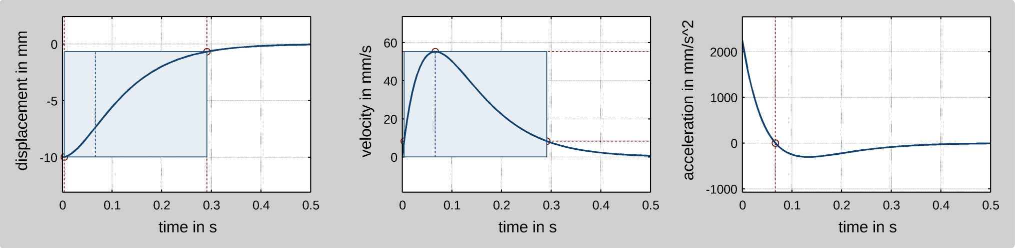
- example curves of **displacement** $x(t)$, **velocity** $\dot{x}(t)$ and **acceleration** $\ddot{x}(t)$ for the **critically damped** linear system ($x_0 = -10\,\mathrm{mm}$, $\dot{x}_0 = 0\,\mathrm{mm/s}$, $k = 225\,\mathrm{s}^{-2}$)



- a simple model of an articulatory movement defines the **kinematic variables**:
  1. **peak velocity** v^* at instant of zero acceleration
  2. movement onset and offset at instants of a certain fraction of peak velocity
  3. movement **amplitude** A and **duration** T follow straight forward
  4. as well as **relative time to peak velocity** (RTTP)

Linear model kinematics (2)

- example curves of **displacement** $x(t)$, **velocity** $\dot{x}(t)$ and **acceleration** $\ddot{x}(t)$ for the **critically damped** linear system ($x_0 = -10\,\mathrm{mm}$, $\dot{x}_0 = 0\,\mathrm{mm/s}$, $k = 225\,\mathrm{s}^{-2}$)



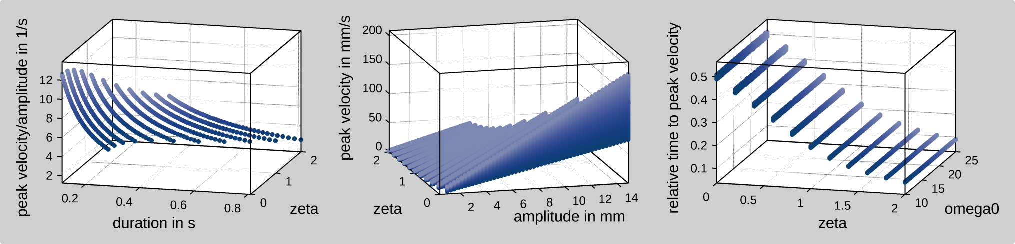
- a simple model of an articulatory movement defines the **kinematic variables**:
  1. **peak velocity** v^* at instant of zero acceleration
  2. movement onset and offset at instants of a certain fraction of peak velocity
  3. movement **amplitude** A and **duration** T follow straight forward
  4. as well as **relative time to peak velocity** (RTTP)

Linear model kinematics (3)

- analytical treatment of the **standard model** leads to the **kinematic relations**:

$$\frac{v^*}{A} = \frac{c\pi}{T} \quad \text{and} \quad v^* = c\omega_0 A$$

- **constant of proportionality** c depends on damping ratio ζ and is maximal ($c = 1/2$) in undamped case (simple harmonic oscillator)



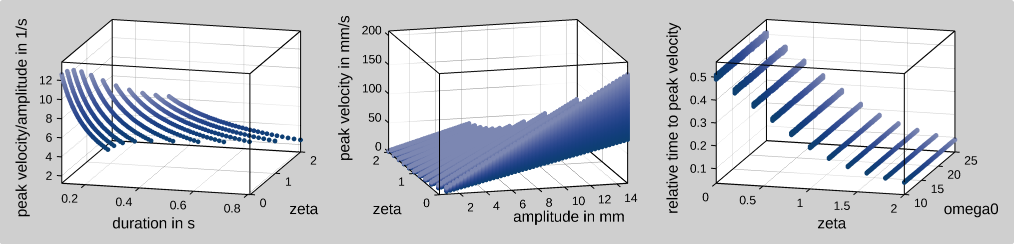
- numerical simulation reveals further a maximal value of 0.5 for **relative time to peak velocity**
- predicted extremal values can be exploited for **model consistency checks** with experimental data (e.g., Fuchs et al., 2011)

Linear model kinematics (3)

- analytical treatment of the **standard model** leads to the **kinematic relations**:

$$\frac{v^*}{A} = \frac{c\,\pi}{T} \quad \text{and} \quad v^* = c\,\omega_0 A$$

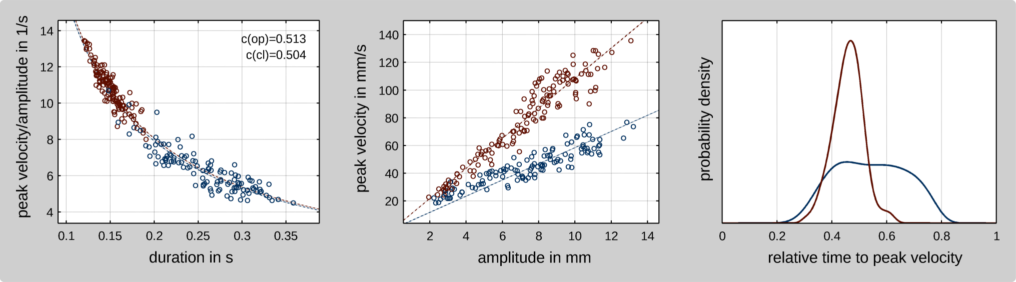
- **constant of proportionality** c depends on damping ratio ζ and is maximal ($c = 1/2$) in undamped case (simple harmonic oscillator)



- numerical simulation reveals further a maximal value of 0.5 for **relative time to peak velocity**
- predicted extremal values can be exploited for **model consistency checks** with experimental data (e.g., Fuchs et al., 2011)

Harvard-Haskins database

- the Harvard-Haskins database of **regularly timed speech** (Patel et al., 1999) contains articulatory data of **self-paced** tongue, lips and jaw movements in /baCa.../ sequences at approx. 2 Hz
- the principal component analysis of **jaw movements** in /baba.../ (3 native English speakers, 4 trials, 11 syllables each) shows the following **kinematic characteristics**:

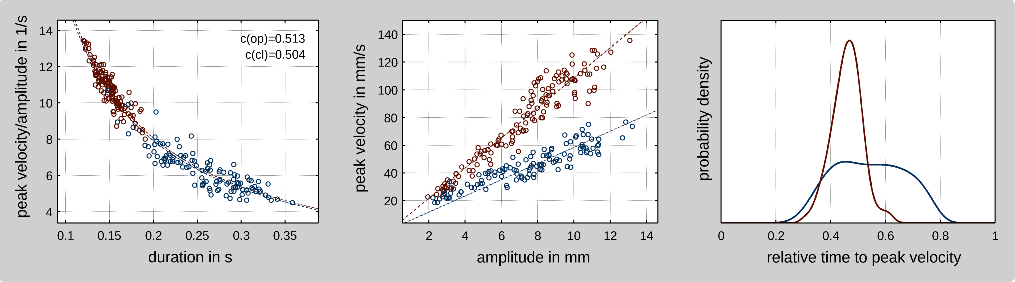


- aside from the asymmetry between opening (red) and closing movements (blue)

1. **kinematic relations** (left and middle) are in general agreement with the **standard model predictions**
2. high **relative time to peak velocity values** (> 0.5, right) cannot be rendered by it

Harvard-Haskins database

- the Harvard-Haskins database of **regularly timed speech** (Patel et al., 1999) contains articulatory data of **self-paced** tongue, lips and jaw movements in /baCa.../ sequences at approx. 2 Hz
- the principal component analysis of **jaw movements** in /baba.../ (3 native English speakers, 4 trials, 11 syllables each) shows the following **kinematic characteristics**:



- aside from the asymmetry between opening (red) and closing movements (blue)
  1. **kinematic relations** (left and middle) are in general agreement with the **standard model predictions**
  2. high **relative time to peak velocity values** (> 0.5, right) cannot be rendered by it

Beyond linear fixed point dynamics

- some improvements to **fixed point dynamics** of the standard model in speech domain:
  1. explicit time dependencies/nonautonomy (e.g., gestural force field by Kröger et al., 1995)
     $$\ddot{x} = -\omega^2 x - 2\omega\dot{x}\,, \quad \text{gestural force function } \omega(t)$$
  2. additional nonlinear terms (e.g., soft spring model by Sorensen and Gafos, 2016)
     $$\ddot{x} = -kx - b\dot{x} + dx^3$$
- there exist other models from general motor control with **qualitatively different dynamics**:
  3. nonlinear oscillator by Schöner (1990) ($\omega = \omega_0 +$ "behavioral information")
     $$\ddot{x} = -(a^2 + \omega^2)x + 2a\dot{x} - 4bx^2\dot{x} + 2abx^3 - b^2x^5$$
  4. excitable system by Jirsa and Kelso (2005) with timing constant τ and external input I
     $$\begin{cases} \dot{x} = (x + y - \frac{1}{3}x^3)\tau \\ \dot{y} = -(x - a + by - I)/\tau \end{cases}$$
- 1 and 2 here can be treated equivalently in terms of kinematic relations
- 3 and 4 have not been applied to speech yet

Beyond linear fixed point dynamics

- some improvements to **fixed point dynamics** of the standard model in speech domain:
  1. explicit time dependencies/nonautonomy (e.g., gestural force field by Kröger et al., 1995)
  $$\ddot{x} = -\omega^2 x - 2\omega\dot{x}\,, \quad \text{gestural force function } \omega(t)$$
  2. additional nonlinear terms (e.g., soft spring model by Sorensen and Gafos, 2016)
  $$\ddot{x} = -kx - b\dot{x} + dx^3$$
- there exist other models from general motor control with **qualitatively different dynamics**:
  3. nonlinear oscillator by Schöner (1990) ($\omega = \omega_0 +$ "behavioral information")
  $$\ddot{x} = -(a^2 + \omega^2)x + 2a\dot{x} - 4bx^2\dot{x} + 2abx^3 - b^2x^5$$
  4. excitable system by Jirsa and Kelso (2005) with timing constant τ and external input I
  $$\begin{cases} \dot{x} = (x + y - \frac{1}{3}x^3)\tau \\ \dot{y} = -(x - a + by - I)/\tau \end{cases}$$
- 1 and 2 here can be treated equivalently in terms of kinematic relations
- 3 and 4 have not been applied to speech yet

Beyond linear fixed point dynamics

- some improvements to **fixed point dynamics** of the standard model in speech domain:

1. explicit time dependencies/nonautonomy (e.g., gestural force field by Kröger et al., 1995)

$$\ddot{x} = -\omega^2 x - 2\omega\dot{x}\,, \quad \text{gestural force function } \omega(t)$$

2. additional nonlinear terms (e.g., soft spring model by Sorensen and Gafos, 2016)

$$\ddot{x} = -kx - b\dot{x} + dx^3$$

- there exist other models from general motor control with **qualitatively different dynamics**:

3. nonlinear oscillator by Schöner (1990) ($\omega = \omega_0 +$ "behavioral information")

$$\ddot{x} = -(a^2 + \omega^2)x + 2a\dot{x} - 4bx^2\dot{x} + 2abx^3 - b^2x^5$$

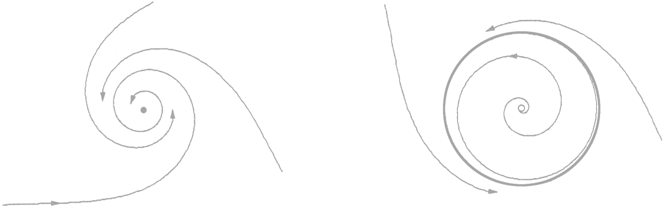
4. excitable system by Jirsa and Kelso (2005) with timing constant τ and external input I

$$\begin{cases} \dot{x} = (x + y - \frac{1}{3}x^3)\tau \\ \dot{y} = -(x - a + by - I)/\tau \end{cases}$$

- 1 and 2 here can be treated equivalently in terms of kinematic relations
- 3 and 4 have not been applied to speech yet

Stability and bifurcation

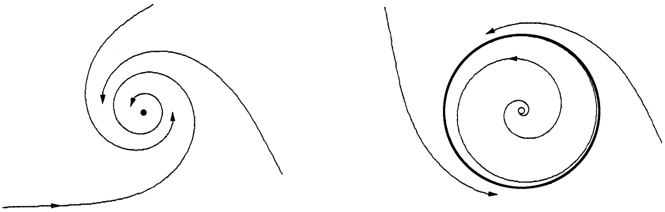
- all these systems are **second order systems**; a **two-dimensional phase space** is conjectured to be sufficient to render both discrete and cyclic movements (Vatikiotis-Bateson and Kelso, 1990)
- some of these models admit limit cycle dynamics (which have not been applied to speech yet)
- **fixed point** and **limit cycle attractors** differ with respect to their topology (Strogatz, 2014):



- fixed point dynamics characterizes **targeted, discrete movements**, whereas limit cycle dynamics characterizes **stationary, periodic movements**
- systems containing multiple attractors show the phenomenon of **bifurcation**, that is a **qualitative change in model behaviour** on change of some model parameter
- there are hints that (speech) rate is such a bifurcation parameter (Tuller and Kelso, 1991; Goldstein et al., 2007; Huys et al., 2008)

Stability and bifurcation

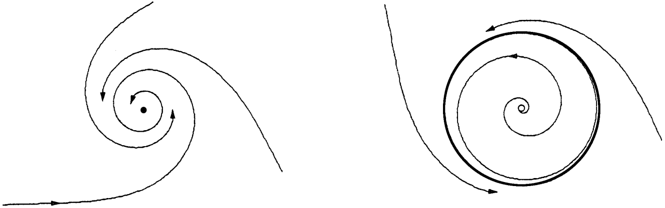
- all these systems are **second order systems**; a **two-dimensional phase space** is conjectured to be sufficient to render both discrete and cyclic movements (Vatikiotis-Bateson and Kelso, 1990)
- some of these models admit limit cycle dynamics (which have not been applied to speech yet)
- **fixed point** and **limit cycle attractors** differ with respect to their topology (Strogatz, 2014):



- fixed point dynamics characterizes **targeted, discrete movements**, whereas limit cycle dynamics characterizes **stationary, periodic movements**
- systems containing multiple attractors show the phenomenon of **bifurcation**, that is a **qualitative change in model behaviour** on change of some model parameter
- there are hints that (speech) rate is such a bifurcation parameter (Tuller and Kelso, 1991; Goldstein et al., 2007; Huys et al., 2008)

Stability and bifurcation

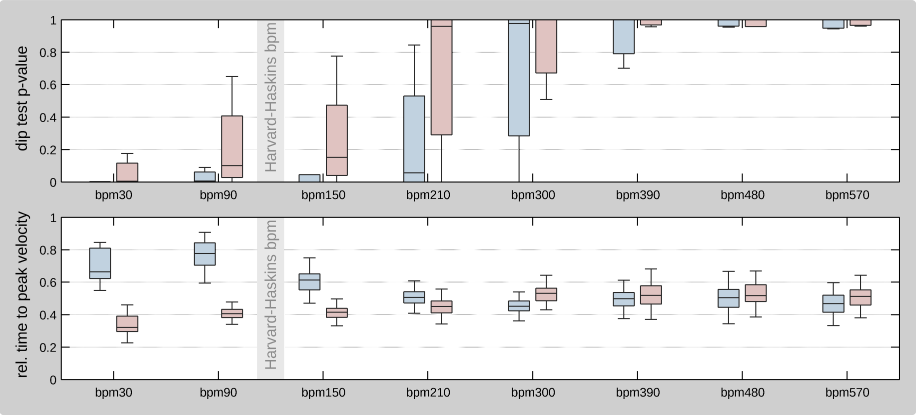
- all these systems are **second order systems**; a **two-dimensional phase space** is conjectured to be sufficient to render both discrete and cyclic movements (Vatikiotis-Bateson and Kelso, 1990)
- some of these models admit limit cycle dynamics (which have not been applied to speech yet)
- **fixed point** and **limit cycle attractors** differ with respect to their topology (Strogatz, 2014):



- fixed point dynamics characterizes **targeted, discrete movements**, whereas limit cycle dynamics characterizes **stationary, periodic movements**
- systems containing multiple attractors show the phenomenon of **bifurcation**, that is a **qualitative change in model behaviour** on change of some model parameter
- there are hints that (speech) rate is such a bifurcation parameter (Tuller and Kelso, 1991; Goldstein et al., 2007; Huys et al., 2008)

Potsdam KORSA pilot

- **repetitive speech** paradigm using systematically **controlled speech rate** (by metronome)
- pilot EMA data of tongue and jaw movements in **repetitions of CV** and CVC sequences
- native English speaker, 4 trials, each 15-30 syllables; here: **tongue tip** in /tata.../:

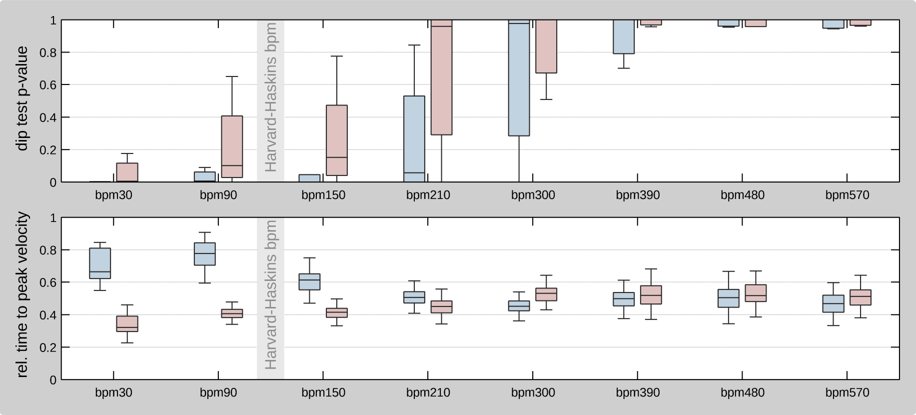


- properties show a **qualitative change** at 210 bpm:

multimodal	⟶	unimodal velocity profiles
two classes	⟶	single class of movements

Potsdam KORSA pilot

- **repetitive speech** paradigm using systematically **controlled speech rate** (by metronome)
- pilot EMA data of tongue and jaw movements in **repetitions of CV** and CVC sequences
- native English speaker, 4 trials, each 15-30 syllables; here: **tongue tip** in /tata.../:

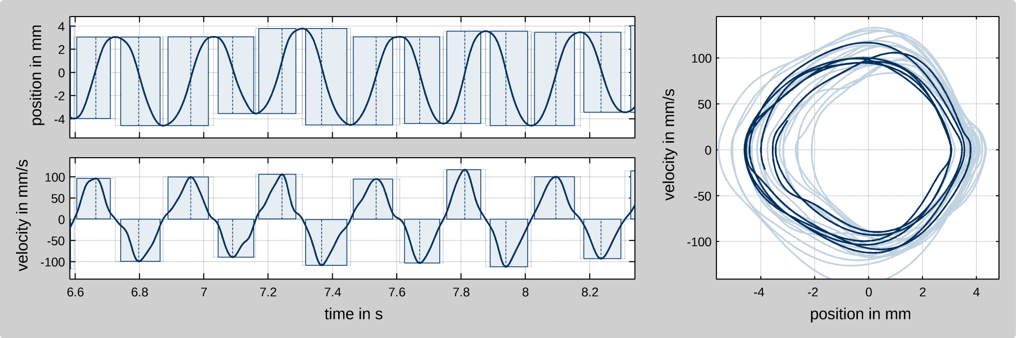


- properties show a **qualitative change** at 210 bpm:

multimodal	⟶	**unimodal velocity profiles**
two classes	⟶	**single class of movements**

Conclusion: evidence for limit cycle

- example of tongue tip trajectory at the critical rate of 210 bpm



- evidence for two distinct timing mechanisms:
  1. dynamics at **low rates** (below critical) conform to **fixed point dynamics**
  2. dynamics at **high rates** (above critical) are **cyclic** and not governed by fixed points
- the **qualitative change** can be identified as bifurcation with **speech rate** as bifurcation parameter
- further investigations have to verify the topological structure (limit cycle)

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), pp. 1067–1076. DOI: 10.1044/1092-4388(2010/10-0131).

Goldstein, Louis M., Marianne Pouplier, Larissa Chen, Elliot L. Saltzman, and Dani Byrd (2007). "Dynamic action units slip in speech production errors". In: *Cognition* 103.3, pp. 386–412. DOI: 10.1016/j.cognition.2006.05.010.

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). DOI: 10.1371/journal.pcbi.1000061.

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), pp. 35–51. DOI: 10.3200/JMBR.37.1.35-51.

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), pp. 1878–1889. DOI: 10.1121/1.413374.

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, pp. 405–408.

Saltzman, Elliot L. and Kevin G. Munhall (1989). "A dynamical approach to gestural patterning in speech production". In: *Ecological Psychology* 1 (4), pp. 333–382. DOI: 10.1207/s15326969eco0104_2.

Schöner, Gregor (1990). "A dynamic theory of coordination of discrete movement". In: *Biological Cybernetics* 63 (4), pp. 257–270. DOI: 10.1121/1.413374.

Sorensen, Tanner and Adamantios I. Gafos (2016). "The gesture as an autonomous nonlinear dynamical system". In: *Ecological Psychology* 28 (4), pp. 188–215. DOI: 10.1080/10407413.2016.1230368.

Strogatz, Steven H. (2014). *Nonlinear Dynamics and Chaos. With Applications to Physics, Biology, Chemistry, and Engineering*. Westview Press.

Tuller, Betty and J. A. Scott Kelso (1991). "The production and perception of syllable structure". In: *Journal of Speech, Language, and Hearing Research* 34.3, pp. 501–508. DOI: 10.1044/jshr.3403.501.

Vatikiotis-Bateson, Eric and J. A. Scott Kelso (1990). "Linguistic structure and articulatory dynamics: a cross language study". In: *Haskins Laboratories Status Report in Speech Research* SR-103/104, pp. 67–94.

Thank you very much!