# Differential difficulty in changing speaking rate: Evidence for 'gaits' of speech

Joe Rodd[1,2] Hans Rutger Bosker[1,3] Mirjam Ernestus[2,1] Louis ten Bosch[2,1] Antje S. Meyer[1,3]
[1]*Max Planck Institute for Psycholinguistics,* [2]*Radboud University, Centre for Language Studies,* [3]*Radboud University, Donders Institute for Brain, Cognition and Behaviour*

Speakers can freely vary stylistic features of their speech, such as their speaking rate. How they accomplish this has hardly been studied, let alone implemented in a model of speech production. Much as in walking and running, where qualitatively different gaits are required to cover the range of different speeds, we may predict there to be qualitatively distinct configurations, or 'gaits', in the speech planning system that speakers switch between to alter their speaking rate or style. Alternatively, control might rely on continuous modulation of a single 'gait'. A recent simulation investigation with the EPONA connectionist model of the formulation phase of speech production [1] supported the 'gait' concept, rejecting a model that supposed a linear relationship between speaking rates.

**The present study** We follow up on the simulation study with two aims: (1) to provide independent evidence for the gaits analogy in formulation, and (2) to work out which speaking rates are achieved by a shared gait and which are achieved by an independent gait. In a sample of fast, medium and slow speaking rates, we examine how quickly speakers change speaking rate, as a proxy for how difficult that adjustment is. The assumption was that crossing a gait boundary is harder than changing rate within a gait. If all three speaking rates are achieved by the same gait, we predict no differences in the speed of speaking rate change. If multiple gaits are present, we predict that the speed of change will differ depending on whether the shift was along the medium↔fast 'axis' or along the medium↔slow 'axis'.

**Experiment** Eighteen native speakers of Dutch were trained to name pictures at fast, medium and slow speech rates from a 'clock face' display. A coloured frame around the 'clock face' indicated the required rate. In each trial of the test phase, the colour of the frame indicating the required rate changed (from fast to medium, slow to medium, medium to fast or medium to slow) at an unpredictable moment, instructing the participant to speed up or to slow down.

**Analysis** Word onsets and offset times were identified by the POnSS system [2], which combines forced alignment and human supervision for efficient, accurate word segmentation. Second, for each trial, the local speaking rate (syllable rate in Hz, log transformed) was calculated in windows for the starting rate (before the switch) and target rate (after the switch). Third, we fitted a Bayesian non-linear regression model that predicted the local speaking rate in each trial independently. The model fitted a stepped shape (Figure 1), with a plateau before the switch, a slope after the switch, and a second plateau to capture the target rate. Finally, a Bayesian mixed-effects meta-regression was conducted over the estimates of the single-trial models, predicting the magnitude of the slope by the magnitude of the difference between the local rate in the starting rate and target rate windows, the axis along which the switch occurred, whether the switch involved acceleration or deceleration, and the interaction between these predictors.

**Results** The modelling revealed a clear effect of axis (Cohen's $d = 0.21$, Figure 2), showing that switching from medium to slow or vice versa was harder than switching between medium and fast speaking rates. There was also a clear effect of acceleration (Cohen's $d = 0.52$, Figure 3), showing that speeding up was harder than slowing down, but no interaction.

**Conclusions** The presence of the axis effect is consistent with the gaits analogy, with medium and fast speaking rates achieved by one gait, and slow speech achieved by another, qualitatively distinct gait. The presence of the acceleration effect implies that rate control is also supported by another mechanism working in concert with gait switching. The effect size measurements from this study will support further development of the EPONA model to establish a comprehensive theory of how speaking rate control is achieved.

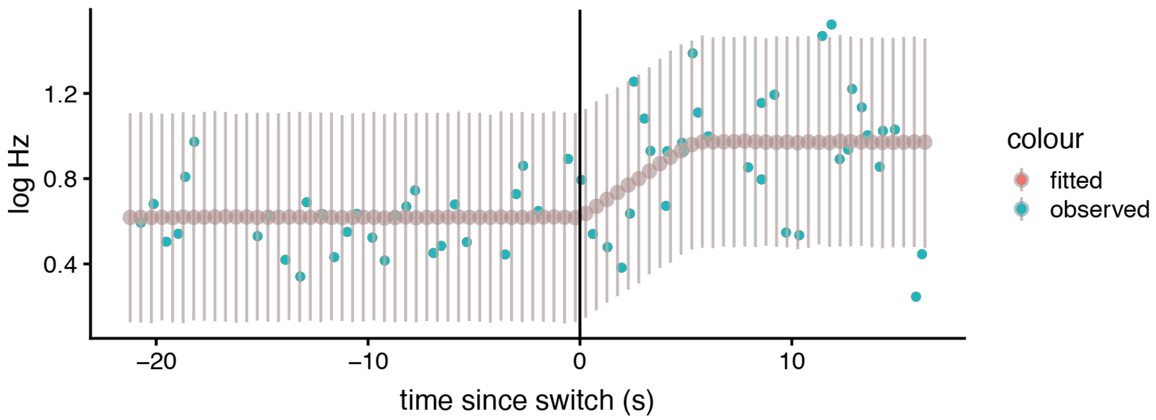


*Figure 1. An example single trial model for the switch between slow and medium. Blue dots indicate measured rates, red dots with CIs indicate model fits. The plateau before 0 ms captures the starting rate. After the switch (0 ms) the slope of the model is allowed to vary. An additional parameter determines when the slope ends and the plateau for the target rate begins.*

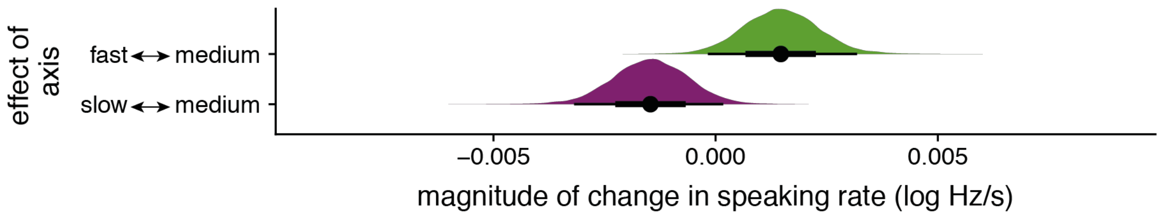


*Figure 2. Posterior sampling distributions for the effect (coefficient) of axis. Lower values indicate shallower slopes, associated with greater difficulty in rate switching.*

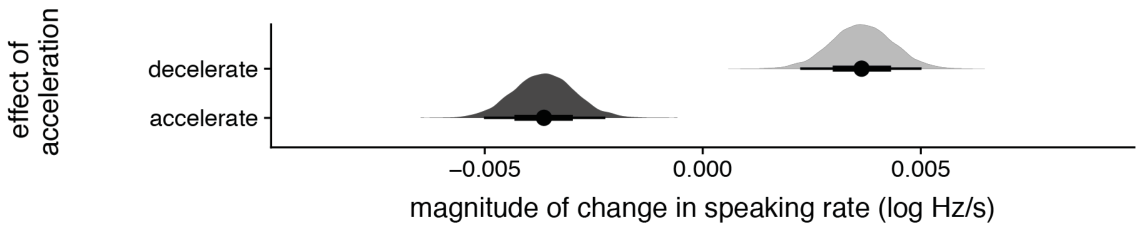


*Figure 3: Posterior sampling distributions for the effect (coefficient) of acceleration. Lower values indicate shallower slopes, associated with greater difficulty in rate switching.*

[1] Rodd, J., Bosker, H. R., Ernestus, M., Alday, P. M., Meyer, A. S., & ten Bosch, L. (2019). Control of speaking rate is achieved by switching between qualitatively distinct cognitive 'gaits': Evidence from simulation. *Psychological Review*. In press. https://doi.org/10.1037/rev0000172

[2] Rodd, J., Decuyper, C., & ten Bosch, L. (in prep). Efficient, reliable semi-manual annotation of speech materials with POnSS.