Published in: Journal of phonetics vol. 68 (2018), pp. 103-116

Research Article

Lexical frequency co-determines the speed-curvature relation in articulation

Fabian Tomaschek [a,*], Denis Arnold [b], Franziska Bröker [c], R. Harald Baayen [a]

[a] Department of General Linguistics, University of Tübingen, Germany
[b] Institut für Deutsche Sprache, Mannheim, Germany
[c] Gatsby Computational Neuroscience Unit, University College London, United Kingdom

ARTICLE INFO

Article history:
Received 4 August 2016
Received in revised form 8 December 2017
Accepted 22 February 2018

Keywords:
Speed-curvature relation
Lexical frequency
Electromagnetic articulography
Speech production
1/3 power law

ABSTRACT

The relation between speed and curvature provides a characterization of the spatio-temporal orchestration of kinematic movements. For hand movements, this relation has been reported to follow a power law with exponent $-1/3$. The same power law has been claimed to govern articulatory movements. We studied the functional form of speed as predicted by curvature using electromagnetic articulography, focusing on three sensors: the tongue tip, the tongue body, and the lower lip. Of specific interest to us was the question of whether the speed-curvature relation is modified by articulatory practice, gauged with words' frequencies of occurrence. Although analyses imposing linearity a priori indeed supported a power law, relaxation of this linearity assumption revealed that the effect of curvature on speed levels off substantially for lower values of curvature. A modification of the power law is proposed that takes this curvature into account. Furthermore, controlling statistically for number of phones and word duration, we observed that the speed-curvature function was further modulated by an interaction of lexical frequency by curvature, such that for increasing frequency, speed decreased slightly for low curvatures while it increased slightly for high curvatures. The modulation of the balance between speed and curvature by lexical frequency provides further evidence that the skill of articulation improves with practice on a word-to-word basis, and challenges theories of speech production.

1. Introduction

1.1. Hand movements

Moving parts of our bodies is an integral part of the interaction with our environment. The timing of these movements is crucial. Consider, for example, catching a flying ball (Peper, Bootsma, Mestre, & Bakker, 1994). First, the ball's direction and speed have to be assessed in order to predict its movement trajectory. Second, the catcher has to predict *where and when* the ball will fly by in order to catch it. Then, the catching arm's intercept trajectory must be predicted and the appropriate movements executed. In order to move the hand toward the necessary position in space, the three subparts of the arm—upper arm, forearm and hand—have to be coordinated. Often there are multiple constellations between the upper and lower arm to plan and execute movement trajectories for any given task. Across situations, one crucial parameter of the movement is its speed: hand and fingers need to be at the right place and right time to catch a ball.

Kinematic studies observed that the tangential speed of hand movements is inversely proportional to the trajectory's curvature (Viviani & Terzuolo, 1982). This inverse relation is illustrated by what happens when riding a bicycle. On straight tracks, where curvature is small, one can increase speed. In bends, where curvature increases, one has to reduce speed. High curvature values represent *tight curves*, low curvature values represent *broad curves*. The functional relation between speed and curvature has been termed the *1/3 power law*, as speed was found to vary with curvature following a power function with exponent $-1/3$:

$$V(t) = K * C(t)^{-1/3}. \quad (1)$$

In Eq. (1), $V(t)$ represents tangential speed and $C(t)$ represents tangential curvature. K is the speed gain factor and represents a linear scaling of the curvature values depending on movement size and individual speed preferences (Viviani & Schneider, 1991). The power law has been reported for eye

* Corresponding author.
E-mail address: fabian.tomaschek@uni-tuebingen.de (F. Tomaschek).

https://doi.org/10.1016/j.wocn.2018.02.003
0095-4470/

movement (de'Sperati & Viviani, 1997) and movements violating the 1/3 power law are perceived as unnatural (Viviani & Stucchi, 1992).

Several explanations have been put forward for this power law. According to Viviani and Flash (1995), it is a consequence of motor control strategies employed by the central nervous system to minimize the force required to change acceleration (estimated by the third derivative of the movement trajectory, see Todorov & Jordan (1998)), resulting in less jerky and more smoothly executed movements. Lebedev, Tsui, and Van Gelder (2001) describes the power law as reflecting the principle of least action, the motor system striving to minimalize the effort required for movement execution.

Although multiple studies have described the speed-curvature relation for simple hand movements by Eq. (1), several hand movement studies also reported strong deviations from the power exponent of $-1/3$. For example, Saling and Phillips (2005) found the power law explains 60% of the variance of finger movements, but only 30% of wrist movements. Apparently, the speed-curvature relation depends on which subpart of a limb is measured. Furthermore, Viviani and Flash (1995) reported that the power exponent varied with the size and shape of the figure traced by the hand. Similar results were obtained by Hicheur, Pham, Arechavaleta, Laumond, and Berthoz (2007) in a study investigating the speed-curvature relation for walking. Viviani and Schneider (1991) were also able to show that the exponent of the power law increases in magnitude with age: Children at the age of six have a speed-curvature slope of -0.25. By the age of 12, the exponent is -0.3, and starts to approximate the adult value.

1.2. Articulation

The 1/3 power law has also been reported for articulation. Fig. 1, panel (d) visualizes the power law for the tongue tip during the articulation of German [ʃaːpt] '*you scrape, pl.*'. An example of the articulatory trajectory in the midsaggital plane is given in panel (a). Panels (b) and (c) present speed and curvature as a function of time. Note that when speed is high, curvature is low, and vice versa.

Tasko and Westbury (2004) were the first to examine the power law for articulation in the tongue tip, the tongue body, the lower lip, and the mandibular incisor positions. They reported that speed and curvature were well-described by a power function with exponent close to $-1/3$. However, they also reported large deviations from this value depending on sensor position and subject.

Perrier and Fuchs (2008) replicated the study of Tasko and Westburry, and observed that the variation in the exponent increased for decreasing signal duration. They argued for articulation that the power law arises from the biomechanical constraints governing biological kinematic systems. Articulatory trajectories simulated with their model GEPPETO (Perrier, Ma, & Payan, 2005) indicated that simple biomechanical constraints emerging from muscle configurations and muscle stiffness already give rise to the 1/3 power law. Much like Gribble and Ostry (1996)'s simulation of arm movements, the model of Perrier and Fuchs (2008) indicates that the power law is unlikely to reflect motor control strategies within the central nervous system, but rather is a straightforward consequence of the spring-like properties of the articulators (Lebedev et al., 2001; Saltzman & Munhall, 1989; Todorov & Jordan, 1998; Viviani & Flash, 1995).

1.3. The present study

Considering these findings, the aim of the present study is twofold. First, the 1/3 power law may be appropriate for elliptical hand movements, executed with a rigid-body joint-angle system that aims for a single positional target. As illustrated above in the example with the ball catching, timing is crucial. Unlike hand movements, movements of the speech articulators consist of overlapping gestures and potentially competing targets executed by an articulatory apparatus that consists of a rigid jaw and muscular hydrostats which are biomechanically joint and coupled during articulation (Bell-Berti & Harris, 1979; Browman & Goldstein, 1986; Fowler & Saltzman, 1993; Saltzman & Munhall, 1989). More importantly, communication can still be successful when articulatory target positions are reached with a relatively broad margin of error, and therefore allow for a larger variability at the target area (Guenther, 1995). Therefore, the question arises whether the 1/3 power law provides a precise description also for the more complex movement trajectories of articulators.

Second, is the functional relation between speed and curvature modulated by a word's frequency of occurrence, a lexical measure which reflects the amount of practice that speakers have had with individual words? Here, we are interested not so much in a potential main effect of frequency, but rather in whether frequency as a predictor of speed interacts with curvature. A main effect of frequency, although not without interest, might be due to high-frequency words being accessed faster in the mental lexicon. As argued by Bell, Brenier, Gregory, Girand, and Jurafsky (2009) and Jurafsky, Bell, Gregory, and Raymond (2000), faster lexical access may enable faster execution of words' articulatory gestures, resulting in the well-documented shortening of higher-frequency words (see, e.g., Aylett & Turk, 2004; Lindblom, 1990; Gahl, 2008; Zipf, 1935).

However, frequency as a predictor of speed of articulation may interact with curvature, such that with practice higher speed can be maintained for higher curvatures. In other words, where skilled muscle coordination is most essential, in narrow curves, repeated use may make it possible for the speaker to maintain higher speeds compared to broad curves, where precision matters less. In other words, as frequency of use decreases, the effect of curvature on speed is expected to be attenuated, whereas with increasing frequency, the effect of curvature on speed is predicted to be enhanced.

If frequency indeed interacts with curvature, then this would provide evidence that mastering and optimizing the articulation of individual words is a continuous process that unfolds as experience with these words accumulates. The motor skills required for playing tennis, or playing the violin, improve qualitatively with practice, affording not only an increase in speed but also more fluid and more precisely targeted movements. In the same vein, as speakers gain more experience with producing a specific word, the way in which they learn to articulate a word may undergo qualitative changes that are the hallmark of motor learning.

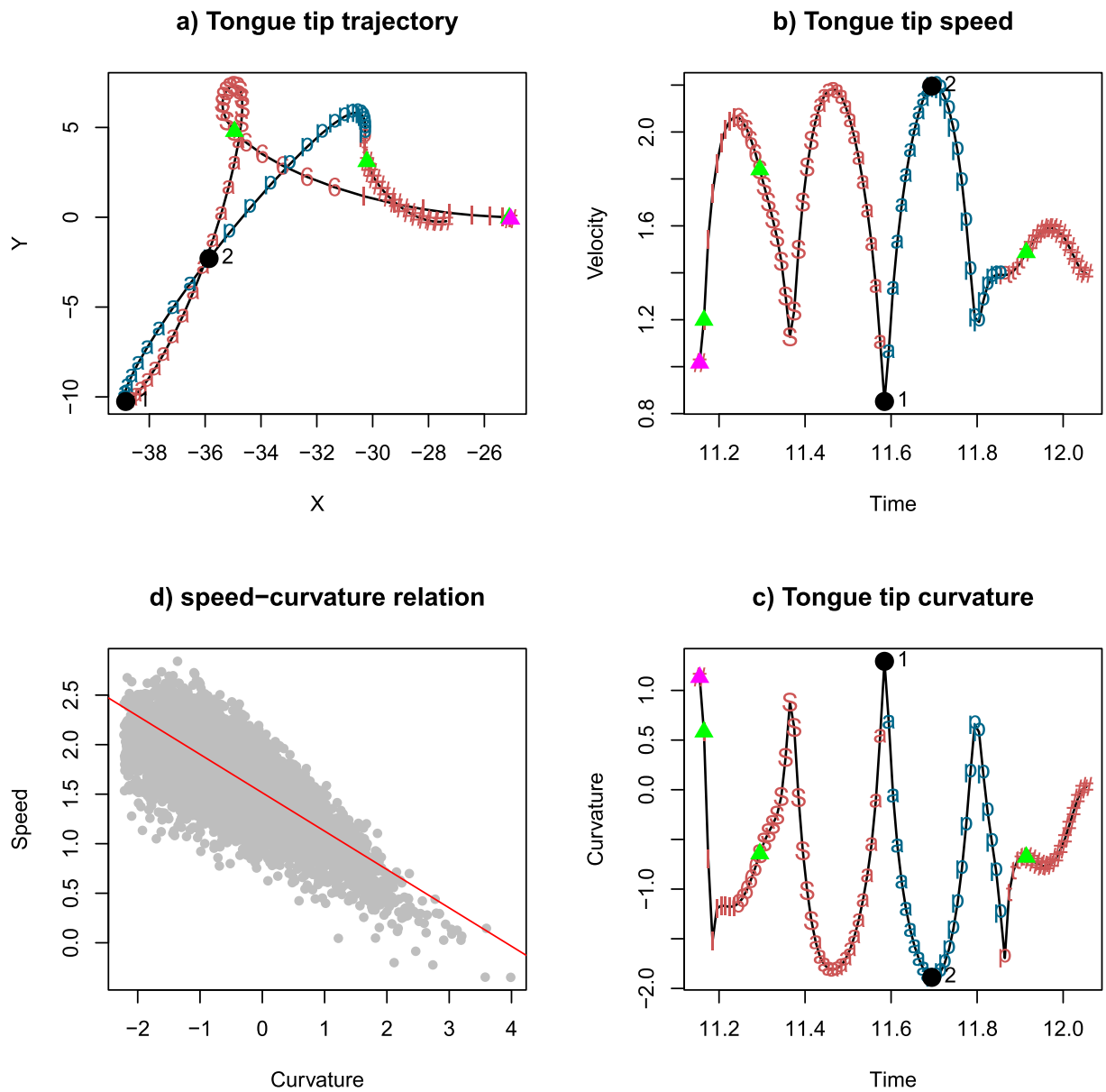

Fig. 1. Movement trajectory (a), speed (b), curvature (d) and the speed-curvature relation (c) for a realization of German [ʃaːpt] '*you scrape, pl.*'. The bullets labeled 1 and 2 highlight the same two points in time across panels a, b and d. The magenta triangle represents the starting point of the investigated trajectory; green triangles the onset of phonation. Partial trajectories labeled with segments in red represent parts of the trajectory during which the articulator was actively involved in production.

The remainder of this paper is structured as follows. We begin with introducing the experimental data, which were collected using electromagnetic articulography. We then examine the 1/3 power law by regressing speed on curvature with a linear model in the bi-logarithmic plane. As a next step, we relax the linearity assumption using thin plate regression spline smoothers as available in the Generalized Additive Model. Analyses addressing the interaction of frequency by curvature are followed by a general discussion of our findings.

2. Methods

We investigated the relation between curvature and speed using data analysed previously with respect to the influence of frequency of use on tongue position (Tomaschek et al., 2014). This study originally recorded 21 native speakers of German, of which 5 had to be excluded from the analysis given substantial missing data due to faulty sensors. The remaining speakers in the analysis (8 females and 8 males) had a mean age of 25.6 (sd = 3.1). All recordings were conducted in a sound attenuated booth at the Department of Linguistics at the University of Tübingen.

All 432 words recorded in this study were selected for further analysis of the speed-curvature relation. The stimuli comprised 157 s person plural forms (e.g. '(ihr) dient', *you serve, pl.*), 197 verbs in the third person plural form (e.g. '(sie) dienen', *they serve*), 39 nouns in the singular (e.g. 'Zunge', *tongue*), and 39 nouns in the plural (e.g. 'Zungen', *tongues*).

Frequency counts for these nouns and verbs were extracted from the SdeWaC corpus of German (approx. 1 billion words, Faaß & Eckart, 2013; Shaoul & Tomaschek, 2013). Frequencies were log transformed to reduce the risk

of overly influential outliers dominating statistical models (henceforth frequency).

The list of 432 words was pseudo-randomized separately for each participant. The words were presented on a computer screen, and participants were requested to utter the word. The words were presented in three blocks, and each block of words was presented at two rates, first a slow rate inducing a slow speaking rate (inter-stimulus-time: 600 ms; presentation-time: 800 ms) and then at a fast rate (inter-stimulus-time: 300 ms; presentation-time: 450 ms) inducing faster speech. These two conditions will be referred to below as the slow and fast speaking rate conditions.

Articulatory movements of the tongue were recorded with an electromagnetic articulograph (EMA, WAVE, Northern Digital Inc.) sampling sensor positions at 100 Hz. Simultaneously, the audio signal was recorded (sampling rate: 22.05 kHz, 16bit). To correct for head movements and to define a local coordinate system, a reference sensor was attached to the subjects' forehead. Before the tongue sensors were attached a recording was made to determine the rotation from the local reference to a standardized coordinate system defined by a bite plate to which three sensors were attached in a triangular configuration. Tongue movements were captured by three sensors: one slightly behind the tongue tip (TT), one at the tongue body (TB) and one sensor between them. The distance between each sensor was roughly 2 cm. Lip movements were captured by three sensors: one on the lower lip (LL), one on the upper lip and one in the right lip corner. Correct attachment of the sensors was checked after every recording block. Recorded positions of the tongue sensors were centered at the midpoint of the bite plate and rotated such that the tongue's front-back direction was aligned to the x-axis with increasing values representing fronting, and to the z-axis with increasing values representing raising. Sensor position was corrected for head movement automatically after every recording. We focused on LL, TT and TB in the current analysis, as these represent the main articulators. All analyses can be found in the Supplementary material, downloadable from https://osf.io/pm2d9/.

2.1. Preprocessing

Audio recordings were synchronized with the articulatory recordings. The audio signal was also automatically aligned with phonetic transcriptions using a Hidden-Markov-Model-based forced aligner for German (Rapp, 1995). Alignments were manually verified and corrected where necessary using PRAAT (Boersma & Weenink, 2015). Positional data was low-pass filtered at 6 Hz using a third order Butterworth filter. Filtering reduces the probability that the $1/3$ power law would emerge merely as a consequence of random noise (Maoz, Portugaly, Flash, & Weiss, 2005). The functional form of the relation between curvature and speed potentially varies with whether an articulator actively executes a gesture or not (Fowler & Saltzman, 1993; Saltzman & Munhall, 1989). We added information about whether or not a sensor was active or inactive on the basis of gestural segmentation. Generally, the onset and offset of gestures are marked at a 20% threshold of an articulator's peak velocity (Bombien, Mooshammer, & Hoole, 2013; Mooshammer & Fuchs, 2002). The 20% velocity threshold is usually assumed for gestures with clear stop phases yielding an absolute minimum in velocity. Inspecting the productions in our data showed that most of the articulations consisted of loops, i.e. continuous movement, even during phases that according to high-level descriptions should have been stationary (such as periods of closure for plosives). As a consequence, the 20% threshold turned out to be highly unstable for our data. We therefore positioned the gestural onset and offset of a segment at the points in time at which the speed of the sensor representing the segment's main articulator reached its minimum value.

Speech production is the result of a joint and coupled system in which especially the height of the tongue and lip is driven by jaw movements. Given the sensors at hand, we considered the lower lip position to represent labial consonants, the tongue tip position coronal and the tongue body velar consonants as well as vowels. In this study, we do not further distinguish between the separate contributions to these sensor movements of the jaw muscles on the one hand, and of the lip and tongue muscles on the other hand.

Panels (a–c) of Fig. 1 present active articulators in red and inactive articulators in blue. Acoustic labels for each time step were recoded so that they represented the segment articulated by the gesture (henceforth *segment gestures*). For example, the German word [ʃaːpt] consists of a TT gesture ([ʃ]), followed by TB ([aː]), LL ([p]), and TT ([t]) gestures. As the tongue tip is active during the articulation of consonants but inactive during the articulation of vowels, tongue tip movements during the initial part of the [aː] were labeled as [ʃ]. For the TT, all labels for coronal phones represent active gestures, while non-coronal sounds represent those phones in which TT was a co-articulator. Each position was coded for 36 segment gestures, whose distribution varied between positions. (For the verbs, which were preceeded by a pronoun, the tongue body movement during the production of the vowels of these pronouns—[ziː] *they* and [iːɒ] *you, pl.*—was classified as inactive.).

Following Perrier and Fuchs (2008) and Tasko and Westbury (2004), tangential speed $V(t)$ was calculated using (2), where $\dot{x}$ and $\dot{y}$ were calculated by central difference approximation. Tangential curvature $C(t)$ was calculated using (3), which specifies the rate of change of a trajectory's direction. In these equations, $\dot{x}, \dot{y}$ in (2) and (3) represent the first derivative (speed) and $\ddot{x}$, $\ddot{y}$ the second derivative (acceleration) of the filtered sensor position in the vertical and horizontal axes, respectively.

$$V(t) = \sqrt{\dot{x}^2 + \dot{y}^2} \tag{2}$$

$$C(t) = \left| \frac{\ddot{x}\dot{y} - \ddot{y}\dot{x}}{(\dot{x}^2 + \dot{y}^2)^{3/2}} \right| \tag{3}$$

In (3), we take the absolute value of the tangential curvature, as the speed curvature relation concerns the magnitude of the curvature and abstracts away from its direction (cf. Perrier & Fuchs, 2008; Tasko & Westbury, 2004). The power law for speed and curvature,

$$V(t) = K * C(t)^{\beta}, \tag{4}$$

defines a straight line in the double logarithmic plane

$$\log V(t) = \log K + \beta \log C(t), \tag{5}$$

with intercept $\log K$ and slope β.

Fig. 2 visualizes the distributions of median speed (top panels) and median curvature (bottom panels) using box and whiskers plots for observed (left and center panels) and log10-transformed (right panels) values. In the left panels, outliers are included, illustrating that the distribution in the observed values is strongly skewed. Only after the exclusion of outliers, median values become apparant (center panels). In the center and right panels, boxes represent the interquartile range and the whiskers 1.5 times the interquartile range. Speed of movement is clearly reduced for the lower lip sensor compared to the two tongue sensors. Differences in the interquartile range (lower for LL speed, higher for LL curvature) are almost completely neutralized by the logarithmic transform. For statistical analysis, we therefore analyzed log speed and log curvature.

Furthermore, we log10-transformed curvature and speed (henceforth curvature and speed).

Data points with extreme values for curvature above 3 and below -4 (LL = 0.14%, TT = 0.09%, TB = 0.08% of the data) as well as data points of words with durations shorter than 200 ms and longer than 1 s (1.93% of the data) were excluded. Regression analyses indicated that speed and curvature values of excluded data points were not predictable from the frequencies of occurrence of their carrier words (summary tables are provided in the Supplementary materials).

In all upcoming analyses, numeric predictors were z-transfered in order to locate the intercept in the center of the distribution.

3. Linear mixed-effects regression analysis

We modeled speed as a linear function of curvature using the `mgcv` package (Wood, 2006), Version 1.8-17, in `R` (Version 3.2.2). We note that, different from the `lme4` package (Bates, Maechler, Bolker, & Walker, 2014), the `mgcv` software estimates random effects by imposing ridge penalties. As additional covariate in this analysis, we included speaking rate. Random intercepts were specified for the random-effect factors speaker and word. Separate models were fitted for the lower lip, the tongue tip, and the tongue body sensor.

Analyses revealed that the slope of curvature interacted signficantly with speaking rate (LL: $\beta_{C:SR} = 0.03$, se < 0.001, t = 61.2; TT: $\beta_{C:SR} = 0.03$, se < 0.001, t = 60.0; TB: $\beta_{C:SR} = 0.017$, se < 0.001, t = 34.7), which is why we performed an individual analysis in each of the speaking rate conditions.

Furthermore, intercepts varied by word and speaker. In addition to these random interepts, by-subject random slopes were supported for curvature and speaking rate. Other random slopes did not receive support from our analyses.

To rule out adverse effects of overly influential outliers, we removed data points with absolute standardized residuals exceeding 2.5 standard deviations (2% of the data), and refitted the model. Full details on these models are available in the Supplementary materials. Table 1 summarizes estimates for slope and intercept for the three sensor positions for the slow and fast speaking rate, the range of variation for the slope across subjects (for both, in slow and fast speaking rate), and the amount of variance explained.

The R^2 for our models ranged between 0.69 and 0.73, which compares favorably with the values reported by Tasko and Westbury (2004) ($\bar{R}^2 = 0.68$) and Perrier and Fuchs (2008) ($\bar{R}^2 = 0.538$).

This, however, may be due to differences in the statistical modeling technique used. Whereas Tasko & Westburry and Perrier & Fuchs used linear regression, we made use of linear mixed-effects regression, which typically accounts for a larger proportion of the variance thanks to the inclusion of both subject and word as random-effect factors. In line with the findings of Saling and Phillips (2005), the amount of variance explained was very similar across sensors. The proportion of variance explained was highly similar across the three sensors.

Population estimates of the slopes ranged from -0.37 to -0.3. As observed also by Tasko and Westbury (2004) and

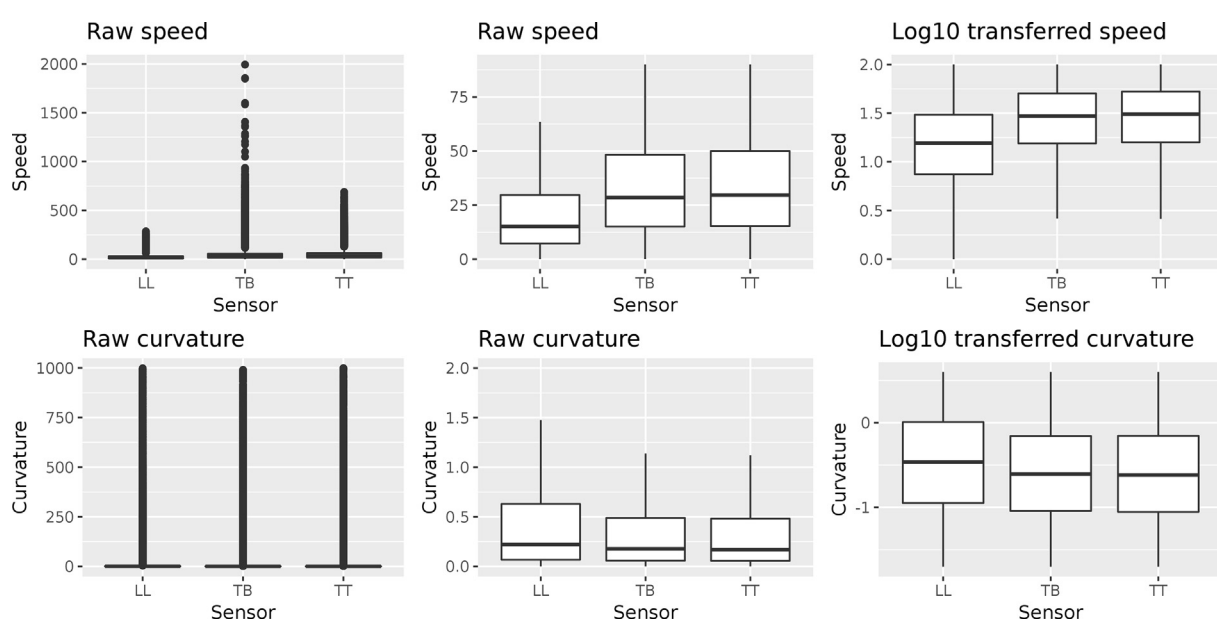

Fig. 2. Box and whisker plots for speed (left with outliers, center without outliers) and curvature (right without outliers) in each of the three sensor positions.

Table 1
Summaries of linear mixed models predicting speed from curvature for three sensors, and the two speaking rate conditions. Min/max: range of β across subjects, obtained by adding the by-subject random slopes to the population slope.

Sensor	Condition	Intercept	Slope (β)	Min	Max	Adj. R^2
LL	Fast	1.22	−0.34	−0.38	−0.32	0.74
TT	Fast	1.54	−0.30	−0.32	−0.27	0.72
TB	Fast	1.51	−0.30	−0.35	−0.26	0.69
LL	Slow	1.15	−0.37	−0.41	−0.33	0.73
TT	Slow	1.48	−0.33	−0.35	−0.31	0.70
TB	Slow	1.45	−0.32	−0.35	−0.29	0.68

Perrier and Fuchs (2008), all positions revealed slopes that are more negative than predicted by the power law. Furthermore, by-subject slopes revealed medium to strong variation (Table 1, columns Min/Max), with the theoretical value of −1/3 being reached only by one subject, and only for the TB position, whereas another speakers revealed a slope of no less than −0.40 for the LL position.

The initial model indicated that speed of articulation increased in the fast speaking rate compared to the slow speaking rate (LL: β_{fast} = 0.06, se = 0.027, t = 2.2; TT: β_{fast} = 0.06, se = 0.028, t = 2.3; TB: β_{fast} = 0.06, se = 0.021, t = 2.7),

Apparently, changes in speaking rate not only affect the intercept *K*, but also the slope, contradicting the assumption in the previous literature that it is only the gain factor *K* (the intercept) that accounts for differences in velocity.

4. Relaxing the linearity assumption

In the preceeding analyses, we have assumed that the functional relation between speed and curvature is a linear one. The spline smooths of the Generalized Additive Mixed Model (GAMM, Hastie & Tibshirani, 1990; Wood, 2006, 2011, 2013) makes it possible to relax this linearity assumption, and to consider instead whether the functional relation between speed and curvature is non-linear. Importantly, GAMMs do not impose non-linearity, as the model incorporates penalties for wiggliness. Thus, if a predictor has a linear effect, a spline smooth will produce a linear regression line, instead of a curvilinear pattern. For univariate smooths, we made use of thin plate regression splines. For multivariate smooths, we used tensor product smooths. Models were fit with the `bam` function of the `mgcv` package (version 1.8-5), and the `itsadug` package (Version 2.2, van Rij, Wieling, Baayen, & van Rijn, 2015) was used for model comparison (using the `compareML` function) as well as for visualization. A short conceptual introduction to the generalized additive model is provided by Baayen, Vasishth, Bates, and Kliegl (2017).

Separate GAMMs were fitted to each of the three sensor positions, with speed as response variable. The predictors of primary interest were curvature and frequency of occurrence. As control covariates, we included number of phones (in the dictionary pronunciation of the word), word duration, and the fixed-effect factors working mode (with levels `active` and `inactive`) and speaking rate (with levels `fast` and `slow`). These fixed-effect factors were included in the model using treatment dummy coding.

Number of phones and word duration are both measures of word length, and were included as length of articulatory movement has been reported to affect the speed-curvature relation in articulation (Perrier & Fuchs, 2008; Tasko & Westbury, 2004). Spearman rank correlations among frequency, word duration and number of phones were low ($r_{frq,nphon}$ = −0.25; $r_{frq,wd}$ = −0.13; $r_{wd,nphon}$ = 0.33), suggesting informally that their joint presence in the model is unlikely to give rise to strong suppression or enhancement (Chatterjee & Hadi, 2012; Farrar & Glauber, 1967; Friedman & Wall, 2005; Tomaschek & Hendrix, 2017). All covariates were z-transformed for analysis.

Four non-linear interactions emerged, all of which involved curvature: curvature by `number of phones` (modeled with a tensor product smooth), curvature by word duration (likewise modeled with a tensor smooth), curvature by working mode (modeled with two thin plate regression splines, one for each level of working mode), and the theoretically expected interaction of curvature by frequency (modeled with a tensor product smooth). These interactions were modeled using a decompositional approach with main effect smooths for all covariates involved, and tensor product smooths for specifically the interactions riding on top of these main effects (using the `ti` function of `mgcv`). For ease of interpretation, the number of basis functions for spline smooths was restricted to three.[1]

Five random-effect factors were taken into account: speaker, word, and the place of articulation of the preceeding segment, of the target segment, and of the following segment. For word, random intercepts sufficed. For Subject, interactions with `Number of phones`, curvature, word duration, frequency, and working mode received strong support. Factor smooths (the nonlinear counterpart of the combination of random intercepts and random slopes in the linear mixed model) were included for curvature by speaker, place of articulation of the preceeding segment, the target segment, and the following segment.

Fig. 3 illustrates the random factor smooths for preceeding segment, the target segment, and the following segment for each of the places of articulation.

4.1. Results for speed of articulation

Table 2 provides a summary overview of the effects of the factorial control predictors, for each of the three positions. Across all positions, speed of articulation increased for the fast speaking rate. Lower lip and tongue tip were articulated significantly faster in the active working mode. In the tongue body position no significant difference between active and inactive was found.

[1] working mode also entered into a significant two-way interaction with curvature at the lower lip and the tongue body sensors, and a significant three-way interaction with curvature and frequency at the lower lip position. Although these additional interactions improved the model fit, they did not affect the main pattern of results nor associated conclusions. As these interactions were not well interpretable, we have not included them in the model reported here. Details on models including these interactions are available in the Supplementary materials.

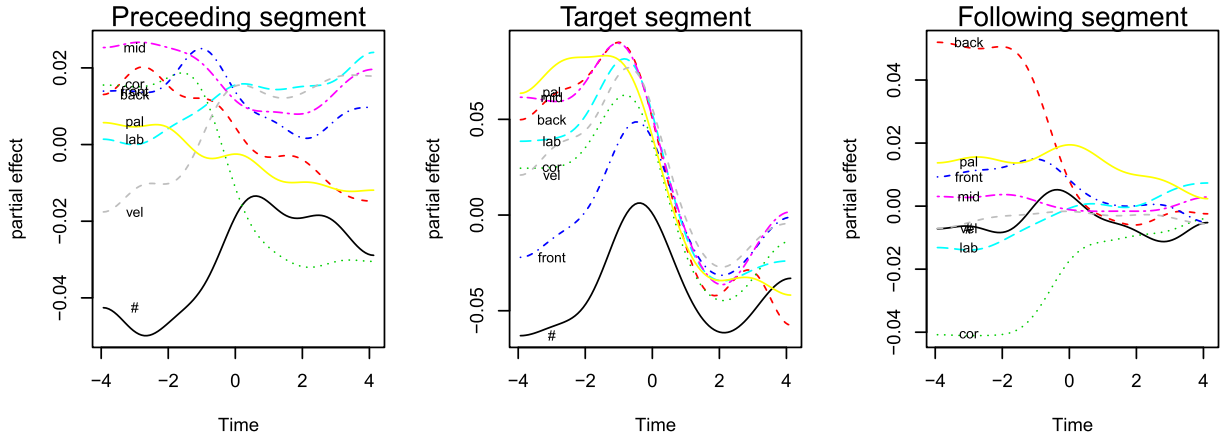

Fig. 3. Partial effect of the by-POA factor smooths for curvature, for preceeding segment (left), target segment (center) and following segment (right) in the tongue tip. Each curve represents a place of articulation: # = word boundary, front/mid/back vowels, cor = coronal, lab = labial, pal = palatal, vel = velar.

Table 2
Summary table of the partial linear effects in the GAMM analyses for lower lip (LL), tongue tip (TT) and tongue body (TB).

Partial effect	Estimate	Std.Error	t-value	p-value
(LL) (Intercept)	0.972	0.026	36.741	<0.001
(LL) Condition: fast	0.068	0.001	58.865	<0.001
(LL) Modus: active	0.077	0.005	16.452	<0.001
(TT) (Intercept)	1.253	0.027	46.818	<0.001
(TT) Condition: fast	0.069	0.001	60.137	<0.001
(TT) Modus: active	0.041	0.002	17.575	<0.001
(TB) (Intercept)	1.271	0.026	49.283	<0.001
(TB) Condition: fast	0.072	0.001	65.105	<0.001
(TB) Modus: active	−0.003	0.002	−1.511	0.131

Figs. 4–6 present those partial effects of the nonlinear predictors that are of theoretical interest. All partial effects are centered around zero. The horizontal axis (at $Y = 0$) in the graphs can be thought of as representing the group mean for a given combination of factorial predictors. In other words, the partial effects clarify how, given the covariate on the horizontal axis, the response departs from this group mean. Departures from the group mean are significant wherever the confidence interval of the smooth does not include the zero line. Summaries of tests of significance for the smooth term are available in the Appendix.

The top panels of Fig. 4 present the non-linear functional relation between speed (vertical axis) and curvature (horizontal axis) for active articulators. Across sensors, we find the same nonlinear functional relation. For low curvatures, the effect appears to be at ceiling. For (scaled) curvatures around -1, a downward trend sets in that becomes nearly linear for (scaled) curvatures exceeding $+1$. Across all three sensors, the $1/3$ power law provides a characterization of the relation between curvature and speech that is too simple. In the general discussion, we show how the power law can be adjusted to bring it closer in line with our data.

Unsurprisingly, the effect size of curvature is large, ranging from nearly $+0.5$ to around -1.0. Much smaller effect sizes characterize the effects of frequency, number of phones, and word duration.

The partial effect of frequency on the speed of articulation (second row of panels in Fig. 4) shows that across all three sensors, speed decreased with increasing frequency of occurrence. The effect of frequency is not confounded with measures of word length, as both number of phones and word duration are included in the model as control covariates. As can be seen in the bottom row of panels in Fig. 4, speed of articulation decreases as word duration increases: In longer words, speakers appear to take more time. Furthermore, the third row of panels of Fig. 4 shows that number of phones has a small and just significant effect only for the tongue body sensor. Given these control predictors, the partial effect of frequency, which holds when other predictors (including duration and number of phones) are held constant, is statistically well-supported.

There are two, not mutually exclusive, explanations for the negative correlation of frequency and speed. First, several studies report that more frequently executed gestures come with faster movement velocities (Platz, Brown, & Marsden, 1998; Raeder, Fernandez-Fernandez, & Ferrauti, 2015; Sosnik, Hauptmann, Karni, & Flash, 2004; Tiede, Mooshammer, Goldstein, Shattuck-Hufnagel, & Perkell, 2011). Since higher-frequency words tend to have higher-frequency phone transitions (Baayen, 2001; Nusbaum, 1985), and since these phone transitions tend to be phonotactically simpler, it is possible that the simplicity of the phone transitions in higher frequencies words drive the present effect. Second, higher-frequency words are read faster Kliegl et al. (2010). If one assumes that fast lexical access in reading also boosts the speed of articulation (Bell et al., 2009; Jurafsky et al., 2000), the negative slope of the frequency effect follows immediately.

As mentioned above, number of phones, word duration, and frequency entered into interactions with curvature. These interactions are visualized in Fig. 5. In this figure, the horizontal axis

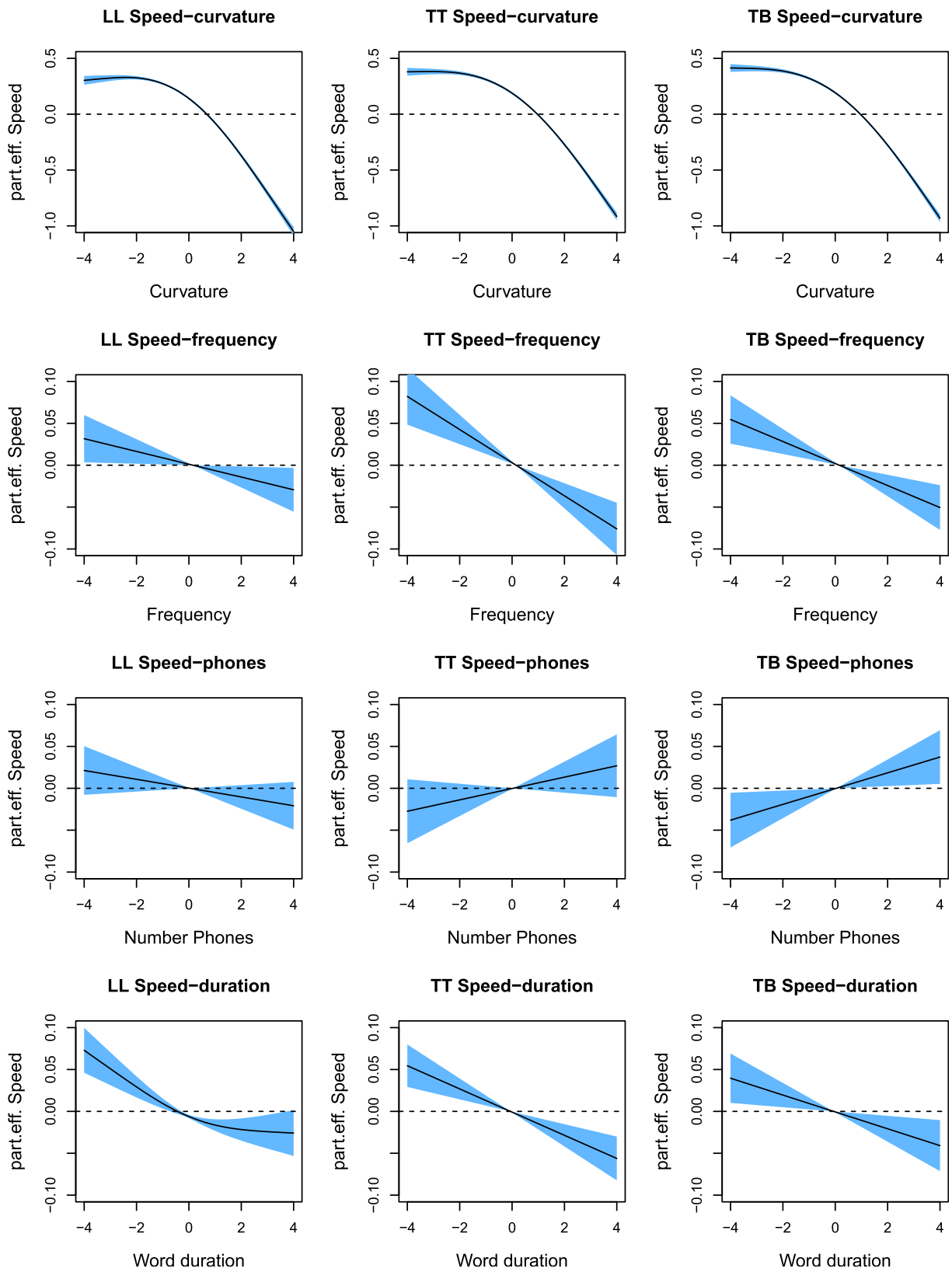

Fig. 4. Partial effects of speed for the lower lip (left column), tongue tip (mid column) and tongue body positions (right column).

always represents curvature. Each panel visualizes a regression surface (modeled with a tensor product smooth) by means of three curves at high (1.5 SD above the mean), median, and low (1.5 SD below the mean) predictor values for number of phones (top panels), word duration (center panels), and frequency (bottom panels). Note that the effect size of these interactions is substantially smaller than the corresponding partial 'main' effects depicted Fig. 4.

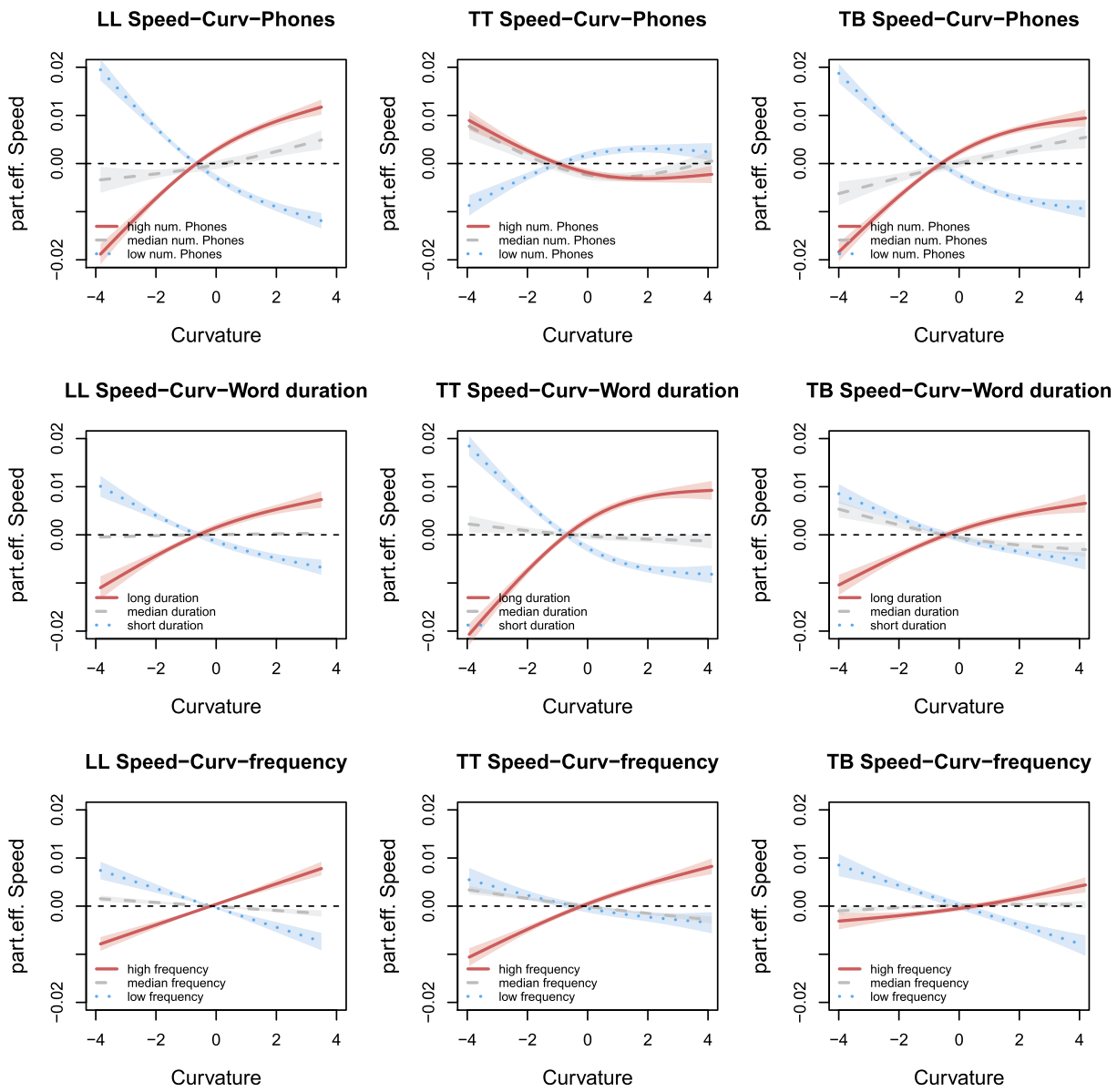

Fig. 5. Partial effects of speed for the interactions between curvature and number of phones, word duration and frequency, in the lower lip (left column), tongue tip (mid column) and tongue body positions (right column).

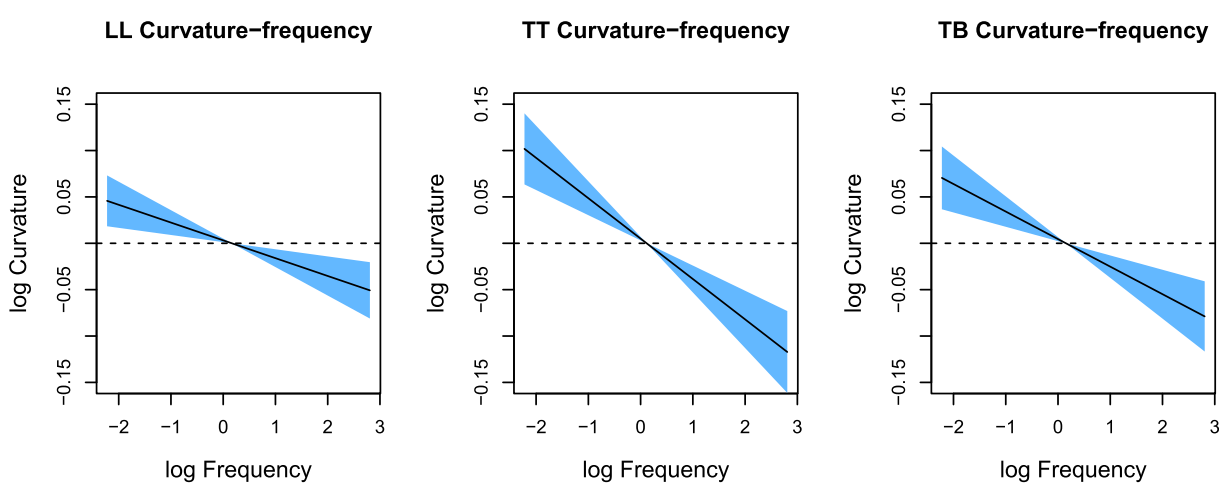

Fig. 6. Partial effects for word frequency as a predictor of curvature, for the LL, TT and TB sensors.

As can be seen in the top left panel of Fig. 5, words with many phones (red) and few phones (blue) had opposite effects on the speed-curvature function. For long words, the curve in the upper left panel of Fig. 4 is modulated downward for low curvatures and moved upward for high curvatures. For short words, the shifts are in the opposite directions. A similar pattern is seen at the TB sensor. The pattern reverses completely, however, at the TT sensor, together with a reduced effect size. Since this reversal is exceptional—all other panels in Fig. 5 consistently show red lines with a positive slope and blue lines with a negative slope—we refrain from interpreting the interaction of curvature by number of phones for the TT sensor.

The interactions in Fig. 5 indicate that speed of articulation increases with curvature for longer words, both when length is measured in number of phones, and when it is assessed by duration in ms, as well as for more frequent words. Conversely, as curvature increases, speed decreases for shorter and lower-frequency words. We note that according to our model, the effects for these three predictors are independent of each other, but we hasten to add that our model specification may have given rise to a simplification of the true complexity of the relation between speed and curvature. However, without more theoretical guidance, it does not make sense to consider higher-order nonlinear interactions.

One would expect, given studies such as Gay (1977), that other things being equal, a larger number of phones should go hand in hand with a higher speed of articulation. The present data show that much depends on curvature. Recall that in Fig. 4, a main effect of Number of phones did not receive solid support. However, at higher curvatures, longer words show (slightly) higher speed, whereas at low curvatures, shorter words show (slightly) higher speed. As demonstrated by the second row of panels of Fig. 5, this characterization of the pattern of results extends to word length assessed with acoustic duration. Possibly, when words are long, speakers are better situated to plan complex articulatory movements, and thus to execute tight curves with greater speed.

The interaction of curvature by frequency mirrors the interactions of length by curvature. However, whereas length can be viewed as a post-lexical factor governing phonetic spell-out, frequency is a lexical factor that apparently also co-determines the fine details of articulation. Again, we observe cross-over interactions for all three sensors, with a reversal in the general slope of the regression lines as we move from rare to frequent words. With speed being already high in low curvature trajectories, articulatory speed is increased in rare words and decreased in frequent words. As expected, higher speeds can be reached for high curvature when word frequencies are higher, that is, when speakers have had substantial practice with articulating these words.

Thus far, we have modeled speed as a function of curvature. Of course, curvature can also be modeled as a function of speed. In this context, the role of word frequency is again of special interest.

It has been shown that repetition of complex trajectories consisting of multiple targets reduces the number of speed peaks and increases the smoothness of movement trajectories (Sosnik et al., 2004; Tiede et al., 2011). It follows that a higher frequency of occurrence should go hand in hand with a reduction in curvature. We therefore fitted a GAMM to the data, with curvature as response and speed as main predictor, while retaining frequency and all other control predictors in the model specification. Here, we only report the main effect of frequency, illustrated in Fig. 6. Just as speed decreases with frequency, curvature likewise decreases with frequency. Again, the effect size is quite small: The largest effect (for the TT) is around 0.3 whereas the range of curvature values is [-4, 4] (see Fig. 4 and 5). Here too, the highly-practiced phone transitions that characterize high-frequency words are at issue. These transitions likely allow for smoother trajectories with reduced curvature (Sosnik et al., 2004; Tomaschek et al., 2014).

5. General discussion

Studies of simple hand movements report that the functional relation between speed and curvature is governed by a power law with -0.33 as exponent (Viviani & Terzuolo, 1982; Viviani & Schneider, 1991). Perrier and Fuchs (2008) and Tasko and Westbury (2004) studied the 1/3 power law for articulatory movements, and although they replicated previous findings for hand movements, they also observed considerable variation depending on sensor, subject and the length of the speech signal. Nevertheless, Perrier and Fuchs (2008) maintained that the speed-curvature relation follows from the spring-like properties of the articulators.

We studied the functional relation between speed and curvature for German verbs and nouns, using electromagnetic articulography, focusing on the lower lip, the tongue tip, and the tongue body. An analysis imposing linearity on this functional relation revealed exponents with a slightly increased absolute magnitude (-0.37 and -0.39, rather than -0.33). We also replicated the results of Perrier and Fuchs (2008) and Tasko and Westbury (2004), in that we observed considerable by subject variation (with individual exponents ranging from -0.33 to -0.44).

However, when the assumption of a linear functional relation between speed and curvature is relaxed, a substantially modified pattern of results emerges. A Generalized Additive Mixed Model (Wood, 2006) using thin plate regression splines provided strong support for a non-linear functional relation between speed and curvature. For lower values of curvature, speed is at ceiling, and as a consequence, for the lowest range of curvatures, speed hardly decreases with curvature. It is only in the second quartile of curvature that speed starts to decrease as curvature is increased, and it is only for the last quartile of curvature that speed decreases approximately linearly with curvature. In other words, for our data, the power law is not valid.

The insight gained with the thin plate regression spline into the true functional relation between speed and curvature indicates that a modification of the power law is required. Following Mandelbrot's emendation of Zipf's rank-frequency law (Mandelbrot, 1953, 1959), we propose to enrich the model with a second parameter, z,

$$V(t) = \frac{K}{(C(t) + z)^{\beta}} \quad (\beta > 0) \tag{6}$$

to ensure that the effect of curvature on speed levels off for low values of curvature. Fig. 7 illustrates that a Zipf-Mandelbrot model for the speed-curvature relation fits the estimated rela-

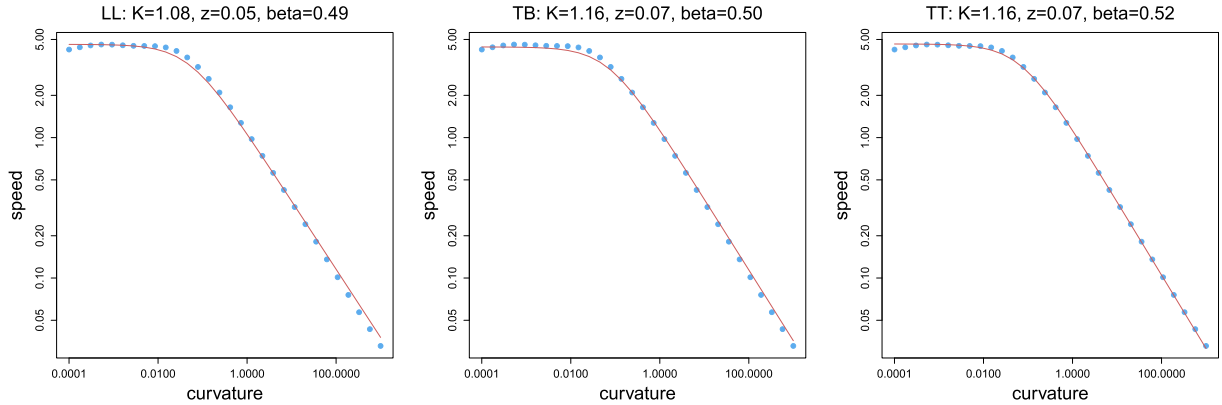

Fig. 7. The Zipf-Mandelbrot model applied to the speed curvature relation for LL sensor (left), tongue body sensor (center), and tongue tip sensor (right). Estimated observed curve in blue, and the Zipf-Mandelbrot approximation in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tion fairly well: Values on a grid of 30 equally-spaced curvatures, predicted by a GAMM, are shown in blue, and the Zipf-Mandelbrot approximation is shown in red. The parameter values shown for each sensor above its panel were obtained with a simple grid search. In model (6), the β parameter regulates the balance between speed and curvature, K defines the maximum speed that can be reached under minimal curvature, and z enforces a constraint that this maximum speed is closely approximated at curvatures that exceed the minimum curvature by a substantial margin.

Speed depends not only on curvature. It varies not only with word length in number of phones, or word length assessed with acoustic duration, but it also varies with lexical frequency. We observed a medium-sized effect of lexical frequency, such that the speed of articulation decreased with increasing frequency. Slower articulatory movements with in words with higher frequency are indeed surprising, as usually more repetitions of a kinematic task result in faster movements, both in hand movements and articulation (Platz et al., 1998; Raeder et al., 2015; Sosnik et al., 2004; Tiede et al., 2011). Given that we controlled for effects of word duration and number of phones in the canonical form, and therefore rule out confounding of our results by these predictors, we argue that a probable source of slower articulation is reading latency.

This negative correlation may have arisen as a consequence of two mutually not exclusive effects. The first is faster lexical access in reading affording more rapid production (cf. Bell et al., 2009; Jurafsky et al., 2000; Kliegl et al., 2010). Having adapted to the time frame in our experiment, faster access provided speakers with a longer time window to articulate the word, reducing articulatory speed.

The second is that the phonotactics of higher frequency words tend to be simpler than the phonotactics of lower-frequency words (see, e.g., Baayen, 2001; Nusbaum, 1985, for a discussion of the relation between frequency and transitional probabilities). As a consequence, the distances that the articulators have to travel for higher frequency words are reduced, and hence, the speed of articulation can be relaxed. In other words, thanks to simpler phone transitions, the articulatory effort required for higher frequency words is reduced (Zipf, 1949).

Instead of modeling speed as a function of curvature, one can model curvature as a function of speed. A brief exploration

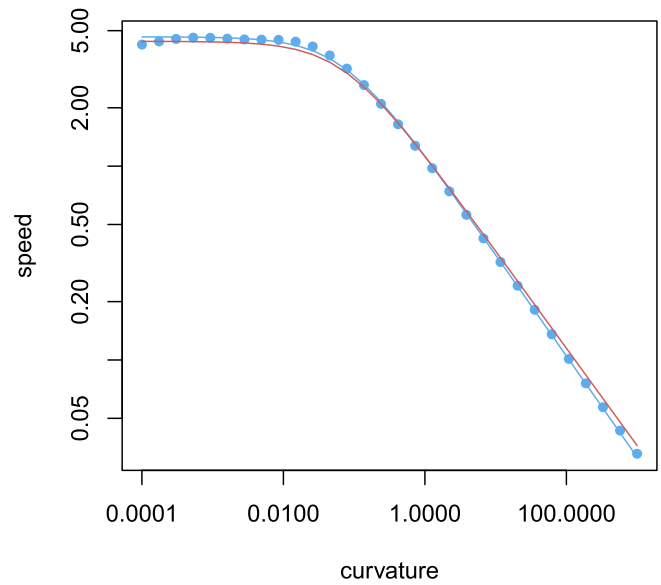

Fig. 8. Slightly decreasing β in the Zipf-Mandelbrot model results in the interactions document in Fig. 5 of curvature by length, by duration, and by frequency. Observed values in blue, in red a curve obtained by increasing β. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of curvature as response variable showed that, as expected, curvature decreases with speed. Of interest is that we also observed lower curvature for words of higher frequency, indicating smoother articulatory trajectories. This mirrors the findings of other kinematic studies. Practice leads to smoother transitions across sequences of gestures, be they in hand movement (Sosnik et al., 2004) or articulation (Tiede et al., 2011). Of further theoretical interest is that in the model regressing speed on curvature, an interaction of curvature by frequency is present in addition to the independently estimated main effects of these predictors.[2]

Fig. 8 illustrates that this interaction (as well as the other interactions with `curvature` summarized in Fig. 5) can be modeled by modifying β in (6). By slightly decreasing β, speed

[2] In fact, in the model regressing curvature on speed, an interaction of curvature by frequency is also present, with a functional form that is the mirror image of that in the model regressing speed on curvature.

decreases for low curvatures and increases for high curvatures. The reverse pattern (not shown) is obtained when β is slightly increased (resulting in the interaction for `Length in phones` for the tongue tip sensor). Fig. 5 shows that this modulation of β is required for both word length (in phones or in ms) and for word frequency. The effect of frequency, a lexical effect, shows that higher speed can be reached for high curvatures for words the articulation of which is better practiced. This increase in speed at high curvatures goes hand in hand with a decrease in speed at lower curvatures, indicating that effort is re-distributed across curvature.

The effects of kinematic practice in experiments, either within one session or within the time span of a few weeks (Platz et al., 1998; Raeder et al., 2015; Sosnik et al., 2004; Tiede et al., 2011), are similar to the effects we observe for word frequency, which reflect the experience with a word accumulated over a participant's lifetime. The way in which lexical frequency shifts the trade-off of speed and curvature (as reflected by changes in β) suggests an improvement in the skill of articulation. Other studies on the role of lexical frequency in articulation also suggest that articulatory practice makes perfect (Tomaschek, Wieling, Arnold, & Baayen, 2013; Tomaschek et al., 2014; Tomaschek et al., 2017). Articulatory practice thus emerges as a new explanatory factor for theories addressing the relation between frequency and speed and duration of articulation, complementing the factors of audience design and information load (Aylett & Turk, 2004, 2006; Lindblom, 1990; Priva, 2015).

The effects of word-specific experience on the fine detail of speed and curvature in articulatory trajectories challenges current theories of speech production. According to the WEAVER model (Levelt, Roelofs, & Meyer, 1999), lexical frequency drives the activation of phonological word forms, which in turn activate phones that are bundled together in syllables. Effects of number of phones and acoustic duration, as observed in the present study, can be accomodated as these effects can be understood as post-lexical. However, a lexical effect is at odds with the modular conception of the speech production process embodied in the WEAVER model. Possibly, a lexical frequency effect can be accomodated within articulatory phonology (Browman & Goldstein, 1986, 1992), if one assumes that each word has its own unique set of vocal tract variables (as defined in Saltzman & Munhall, 1989). However, such a move seems at odds with the conceptualization of gestures as "primitives of phonological contrasts" (Browman & Goldstein, 1992, p. 24).

More in general, high-level theories tend to profit from a modular approaches, and provide a framework within which many phenomena can be understood in an insightful way. However, when it comes to the finer details, Sapir's aphorism that "all grammars leak" (Sapir, 1921, p. 39) becomes relevant, and strictly modular approaches are stretched beyond their limits.

In morphology, for instance, whole word frequency effects have recently been found to be predictive for lexical processing, rather than lemma frequency effects, even in a highly agglutinative language such as Estonian Loo, Jaervikivi, Tomaschek, Tucker, and Baayen (2017). In syntax, frequency effects for word n-grams (see, e.g., Arnon & Snider, 2010; Bannard & Matthews, 2008; Tremblay & Tucker, 2011) bear witness to the brain's unexpectedly high degree of sensitivity to co-occurrence probabilities. From this general perspective, the present findings that lexical frequency modulates articulation is perhaps unsurprising. The challenge for linguistic theory is how to explain these findings in a theoretically coherent and insightful way. Meeting this challenge is beyond the scope of the present paper. All that we can say here is that repeated exposure and use appears to be part of a general never-ending process of local optimization in human cognition (Ramscar, Sun, Hendrix, Baayen, & Rapp, 2017).

Acknowledgements

This research was funded in part by the Alexander von Humboldt professorship awarded to R. H. Baayen (grant 1141527), and in part by a collaborative grant from the Deutsche Forschungsgemeinschaft (BA 3080/3-1). We are indebted to Ryan Podlubny, Tine Mooshammer and two anonymous reviewers for their helpful comments and suggestions.

Appendix A

See Tables 3–5.

Table 3
Summary table of the partial non-linear effects in the GAMM analysis for **lower lip**, containing the estimated degrees of freedom (edf), the referential degrees of freedom (Ref. edf), the F-value (F) and the associated p-value. P-values below 0.05 indicate significant non-linearity. The upper half represents the fixed effects, the lower half represents the random effects. s() represents a smooth; ti() represents a partial tensor. POA = Place of articulation.

Partial Effect	edf	Ref. edf	F	p-value
s(Curvature)	1.999	1.999	1592.212	< 0.001
s(Frequency)	1.001	1.001	5.037	0.025
ti(Curvature,Frequency)	1.888	2.078	65.46	< 0.001
s(Number of Phones)	1.001	1.001	2.144	0.143
ti(Curvature,Number of Phones)	3.822	3.986	94.946	< 0.001
s(Word duration)	1.968	1.998	22.907	< 0.001
ti(Curvature,Word duration)	1.893	1.989	55.804	< 0.001
s(WordGestures)	405.472	429	31.175	< 0.001
s(Curvature,Speaker)	138.225	152	1997.507	< 0.001
s(Curvature,POA.Segment)	64.404	71	737.62	< 0.001
s(Curvature,POA following segment)	55.669	71	224.299	< 0.001
s(Curvature,POA preceeding segment)	51.468	71	47.912	< 0.001
s(Number of Phones,Speaker)	14.641	16	56.273	< 0.001
s(Word duration,Speaker)	15.307	16	82.942	< 0.001
s(Frequency,Speaker)	14.296	16	21.738	< 0.001
s(Active,Speaker)	15.212	32	9.366	< 0.001

Table 4
Summary table of the partial non-linear effects in the GAMM analysis for **tongue tip**, containing the estimated degrees of freedom (edf), the referential degrees of freedom (Ref. edf), the F-value (F) and the associated p-value. P-values below 0.05 indicate significant non-linearity. The upper half represents the fixed effects, the lower half represents the random effects. s() represents a smooth; ti() represents a partial tensor. POA = Place of articulation.

Partial Effect	edf	Ref. edf	F	p-value
s(Curvature)	1.999	1.999	1392.894	< 0.001
s(Frequency)	1	1	23.746	< 0.001
ti(Curvature,Frequency)	3.375	3.798	43.728	< 0.001
s(Number of Phones)	1	1	2.06	0.151
ti(Curvature,Number of Phones)	3.879	3.994	25.313	< 0.001
s(Word duration)	1.009	1.018	18.27	< 0.001
ti(Curvature,Word duration)	3.763	3.974	107.795	< 0.001
s(WordGestures)	412.132	429	52.498	< 0.001
s(Curvature,Speaker)	135.595	152	2618.011	< 0.001
s(Curvature,POA.Segment)	48.634	71	102.162	< 0.001
s(Curvature,POA following segment)	63.725	71	230.762	< 0.001
s(Curvature,POA preceeding segment)	51.718	71	95.11	< 0.001
s(Number of Phones,Speaker)	15.39	16	148.714	< 0.001
s(Word duration,Speaker)	15.326	16	256.991	< 0.001
s(Frequency,Speaker)	14.455	16	36.183	< 0.001
s(Active,Speaker)	15.047	32	7.988	< 0.001

Table 5
Summary table of the partial non-linear effects in the GAMM analysis for **tongue body**, containing the estimated degrees of freedom (edf), the referential degrees of freedom (Ref. edf), the F-value (F) and the associated p-value. P-values below 0.05 indicate significant non-linearity. The upper half represents the fixed effects, the lower half represents the random effects. s() represents a smooth; ti() represents a partial tensor. POA = Place of articulation.

Partial Effect	edf	Ref. edf	F	p-value
s(Curvature)	1.999	1.999	1483.509	< 0.001
s(Frequency)	1	1	14.33	< 0.001
ti(Curvature,Frequency)	3.227	3.719	24.572	< 0.001
s(Number of Phones)	1.001	1.001	5.43	0.02
ti(Curvature,Number of Phones)	3.892	3.995	103.449	< 0.001
s(Word duration)	1.001	1.002	7.2	0.007
ti(Curvature,Word duration)	3.75	3.968	51.955	< 0.001
s(Curvature,Speaker)	132.196	153	1740.594	< 0.001
s(Curvature,POA.Segment)	65.548	72	210.968	< 0.001
s(Curvature,POA following segment)	55.575	72	273.051	< 0.001
s(Curvature,POA preceeding segment)	54.383	72	157.002	< 0.001
s(Number of Phones,Speaker)	15.279	16	130.822	< 0.001
s(Word duration,Speaker)	15.541	16	270.773	< 0.001
s(Frequency,Speaker)	14.419	16	34.61	< 0.001
s(Active,Speaker)	15.357	33	11.205	< 0.001
s(WordGestures)	407.823	430	38.081	< 0.001

References

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*(1), 67–82.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech, 47*(1), 31–56.

Aylett, M., & Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *Journal of the Acoustical Society of America, 119*(5), 3048ff.

Baayen, R. H. (2001). *Word frequency distributions*. Dordrecht: Kluwer Academic Publishers.

Baayen, R. H., Vasishth, S., Bates, D., & Kliegl, R. (2017). The cave of Shadows. Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language, 94*, 206–234.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science, 19*, 241–248.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. URL: < http://CRAN.R-project.org/package=lme4>.

Bell-Berti, F., & Harris, K. (1979). Anticipatory coarticulation: Some implications from a study of lip rounding. *Journal of the Acoustical Society of America, 65*(5), 1268–1270.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language, 60*(1), 92–111. ISSN 0749-596X.

Boersma, P., & Weenink, D. (2015). Praat: Doing phonetics by computer [Computer program], Version 5.3.41. Retrieved from < http://www.praat.org/>.

Bombien, L., Mooshammer, C., & Hoole, P. (2013). Articulatory coordination in word-initial clusters of German. *Journal of Phonetics, 41*(6), 546–561. ISSN 0095-4470.

Browman, C., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology, 3*, 219–252.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*(3–4), 155–180. ISSN 0031-8388.

Chatterjee, S., & Hadi, A. S. (2012). *Regression analysis by example* (5th ed.). Wiley.

de'Sperati, C., & Viviani, P. (1997). The relationship between curvature and velocity in two-dimensional smooth pursuit eye movements. *The Journal of Neuroscience, 17* (10), 3932–3945.

Faa, G., & Eckart, K. (2013). SdeWaC – A corpus of parsable sentences from the web. In I. Gurevych, C. Biemann, & T. Zesch (Eds.), *Language processing and knowledge in the web, lecture notes in computer science* (pp. 61–68). Berlin Heidelberg: Springer.

Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics, 49*(1), 92–107. URL: < http://www.jstor.org/stable/1937887>.

Fowler, C., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech, 36*(2–3), 171–195. https://doi.org/10.1177/002383099303600304.

Friedman, L., & Wall, M. (2005). Graphical views of suppression and multicollinearity in multiple regression. *The American Statistician, 59*, 127–136.

Gahl, S. (2008). Thyme and Time are not homophones. Word durations in spontaneous speech. *Language, 84*(3), 474–496.

Gay, T. (1977). Articulatory movements in VCV sequences. *The Journal of the Acoustical Society of America, 62*(1), 183–193. https://doi.org/10.1121/1.381480.

Gribble, P. L., & Ostry, D. J. (1996). Origins of the power law relation between movement velocity and curvature: Modeling the effects of muscle mechanics and limb dynamics. *Journal of Neurophysiology, 76*(5), 2853–2860. ISSN 0022-3077. URL: < http://jn.physiology.org/content/76/5/2853>.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Biological Cybernetics, 72*(1), 43–53.

Hastie, T., & Tibshirani, R. (1990). *Generalized additive models*. London: Chapman & Hall.

Hicheur, H., Pham, Q.-C., Arechavaleta, G., Laumond, J.-P., & Berthoz, A. (2007). The formation of trajectories during goal-oriented locomotion in humans. I.A stereotyped behaviour. *European Journal of Neuroscience, 26*(8), 2376–2390. https://doi.org/10.1111/j.1460-9568.2007.05836.x. ISSN 1460-9568.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2000). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.

Kliegl, R, Wei, P., Dambacher, M., Yan, M., & Zhou, X. Experimental effects and individual differences in linear mixed models: Estimating the relation of spatial, object, and attraction effects in visual attention. *Frontiers in Psychology 1*(238). ISSN 1664-1078. doi: https://doi.org/10.3389/fpsyg.2010.00238. URL: < http://www.frontiersin.org/quantitativepsychologyandmeasurement/10.3389/fpsyg.2010.00238/abstract>.

Lebedev, S., Tsui, W. H., & Van Gelder, P. (2001). Drawing movements as an outcome of the principle of least action. *Journal of Mathematical Psychology, 45*, 43–52.

Levelt, W. J., Roelofs, A., & Meyer, A. S. A theory of lexical access in speech production. *The Behavioral and Brain Sciences 22*(1).

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (Vol. 55, pp. 403–439). Netherlands: Springer. ISBN 978-94-010-7414-8.

Loo, K., Jaervikivi, J., Tomaschek, F., Tucker, B., & Baayen, R. (2017). Production of Estonian case-inflected nouns shows whole-word frequency and paradigmatic effects. *Morphology*. accepted for publication.

Mandelbrot, B. (1953). An information theory of the statistical structure of language. In W. E. Jackson (Ed.), *Communication theory* (pp. 503–512). New York: Academic Press.

Mandelbrot, B. (1959). A note on a class of skew distribution functions: Analysis and critique of a paper by H.A. Simon. *Information and Control, 2*, 90–99.

Maoz, U., Portugaly, E., Flash, T., & Weiss, Y. (2005). Noise and the two-thirds power law. In P. J. Weiss & B. Scholkopf (Eds.), *Advances in Neural Information Processing Systems 18: Proceedings of the 2005 Conference* (pp. 851–858). Cambridge, MA: MIT Press.

Mooshammer, C., & Fuchs, S. (2002). Stress distinction in German: Simulating kinematic parameters of tongue-tip gestures. *Journal of Phonetics, 30*(3), 337–355. ISSN 0095-4470.

Nusbaum, H. C. (1985). *A stochastic account of the relationship between lexical density and word frequency* (Tech. Rep.). Indiana University (research on Speech Perception, Progress Report #11).

Peper, L., Bootsma, R. J., Mestre, D. R., & Bakker, F. C. (1994). Catching balls: How to get the hand to the right place at the right time. *Journal of Experimental Psychology Human Perception & Performance, 20*(3), 591–612.

Perrier, P., & Fuchs, S. (2008). Speed-curvature relations in speech production challenge the 1/3 power law. *Journal of Neurophysiology, 100*(3), 1171–1183. ISSN 0022-3077.

Perrier, P., Ma, L., & Payan, Y. (2005). Modeling the production of VCV sequences via the inversion of a biomechanical model of the tongue. In *Proceedings of the ICPHS IX, Paris, France*. .

Platz, T., Brown, R., & Marsden, C. (1998). Training improves the speed of aimed movements in Parkinson's disease. *Brain, 121*, 505–513.

Priva, U. C. (2015). Informativity affects consonant duration and deletion rates. *Laboratory Phonology, 6*(2), 243–278.

Raeder, C., Fernandez-Fernandez, J., & Ferrauti, A. (2015). Effects of six weeks of medicine ball training on throwing velocity, throwing precision, and isokinetic strength of shoulder rotators in female handball players. *The Journal of Strength & Conditioning Research, 29*(7), 1904–1914.

Ramscar, M., Sun, C. C., Hendrix, P., & Baayen, R. H. The mismeasurement of mind: Life-span changes in paired-associate-learning scores reflect the cost of learning, not cognitive decline. *Psychological Science*. doi:https://doi.org/10.1177/0956797617706393.

Rapp, S. (1995). Automatic phonemic transcription and linguistic annotation from known text with Hidden Markov Models/An Aligner for German. In *Proceedings of ELSNET goes east and IMACS Workshop, Moscow*. .

Saling, L. L., & Phillips, J. G. (2005). Variations in the relationship between radius of curvature and velocity as a function of joint motion. *Human Movement Science, 24* (5–6), 731–743. https://doi.org/10.1016/j.humov.2005.09.001. ISSN 0167-9457. URL: < http://www.sciencedirect.com/science/article/pii/S0167945705000667>, neural, Cognitive and Dynamic Perspectives of Motor Control.

Saltzman, E., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology, 1*(4), 333–382.

Sapir, E. (1921). *Language: An introduction to the study of speech*. Brace, New York: Harcourt.

Shaoul, C., & Tomaschek, F. (2013). A phonological database based on CELEX and N-gram frequencies from the SDEWAC corpus. URL: < https://fabiantomaschek.files.wordpress.com/2016/07/tomaschekcorpusreadme.pdf>.

Sosnik, R., Hauptmann, B., Karni, A., & Flash, T. (2004). When practice leads to co-articulation: The evolution of geometrically defined movement primitives. *Experimental Brain Research, 156*, 422–438.

Tasko, S. M., & Westbury, J. R. (2004). Speed-curvature relations for speech-related articulatory movement. *Journal of Phonetics, 32*, 65–80.

Tiede, M., Mooshammer, C., Goldstein, L., Shattuck-Hufnagel, S., & Perkell, J. (2011). Motor learning of articulator trajectories in production of novel utterances. In *Proceedings of the ICPHS XVII, Hong Kong*. .

Todorov, E., & Jordan, M. I. (1998). Smoothness maximization along a predefined path accurately predicts the speed profiles of complex arm movements. *Journal of Neurophysiology, 80*(2), 696–714. ISSN 0022-3077.

Tomaschek, F., & Hendrix, P. Coping with collinearity in linguistic regression analysis. *Journal of Phonetics*.

Tomaschek, F., Tucker, B. V., & Baayen, R. H. The consequences of lexical proficiency for articulation. Linguistic Vanguard.

Tomaschek, F., Tucker, B. V., Wieling, M., & Baayen, R. H. (2014). Vowel articulation affected by word frequency. In *Proceedings of the 10th ISSP, Cologne* (pp. 425–428).

Tomaschek, F., Wieling, M., Arnold, D., & Baayen, R. H. (2013). Word frequency, vowel length and vowel quality in speech production: An EMA study of the importance of experience. In *Proceedings of the Interspeech, Lyon*. .

Tremblay, A., & Tucker, B. V. (2011). The effects of N-gram probabilistic measures on the recognition and production of four-word sequences. *The Mental Lexicon, 6*(2), 302–324.

van Rij, J., Wieling, M., Baayen, R.H., & van Rijn, H. (2015). itsadug: Interpreting Time Series, Autocorrelated Data Using GAMMs, r package version 0.8.

Viviani, P., & Flash, T. (1995). Minimum-jerk, two-thirds power law, and isochrony: Converging approaches to movement planning. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 32–53.

Viviani, P., & Schneider, R. (1991). A developmental study of the relationship between geometry and kinematics in drawing movements. *Journal of Experimental Psychology Human Perception & Performance, 17*(1), 198–218.

Viviani, P., & Stucchi, N. (1992). Biological movements look uniform: Evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 603–623.

Viviani, P., & Terzuolo, C. (1982). Trajectory determines movement dynamics. *Neuroscience, 7*(2), 431–437. ISSN 0306-4522.

Wood, S. N. (2006). *Generalized additive models: An introduction with R*. Boca Raton, Florida, U.S.A: Chapman and Hall/CRC.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B), 73*, 3–36.

Wood, S. N. (2013). On p-values for smooth components of an extended generalized additive model. *Biometrika, 100*, 221–228.

Zipf, G. (1935). *The psycho-biology of language*. An Introduction to Dynamic Philology: MIT Press, Cambridge.

Zipf, G. (1949). Addison-Wesley Press. Cambridge.