ASHA
American Speech-Language-Hearing Association

JSLHR

Research Article

On the Relation Between Leg Motion Rate and Speech Tempo During Submaximal Cycling Exercise

Heather Weston,[a] Wim Pouw,[b] and Susanne Fuchs[c]

[a] Humboldt-Universität zu Berlin, Leibniz-Zentrum Allgemeine Sprachwissenschaft, Germany [b] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands [c] Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin, Germany

ARTICLE INFO

Article History:
Received March 15, 2023
Revision received July 29, 2023
Accepted December 5, 2023

Editor-in-Chief: Cara E. Stepp
Editor: Ben A. M. Maassen

https://doi.org/10.1044/2023_JSLHR-23-00178

ABSTRACT

Purpose: This study investigated whether temporal coupling was present between lower limb motion rate and different speech tempi during different exercise intensities. We hypothesized that increased physical workload would increase cycling rate and that this could account for previous findings of increased speech tempo during exercise. We also investigated whether the choice of speech task (read vs. spontaneous speech) affected results.
Method: Forty-eight women who were ages 18–35 years participated. A within-participant design was used with fixed-order physical workload and counterbalanced speech task conditions. Motion capture and acoustic data were collected during exercise and at rest. Speech tempo was assessed using the amplitude envelope and two derived intrinsic mode functions that approximated syllable-like and footlike oscillations in the speech signal. Analyses were conducted with linear mixed-effects models.
Results: No direct entrainment between leg cycling rate and speech rate was observed. Leg cycling rate significantly increased from low to moderate workload for both speech tasks. All measures of speech tempo decreased when participants changed from rest to either low or moderate workload.
Conclusions: Speech tempo does not show temporal coupling with the rate of self-generated leg motion at group level, which highlights the need to investigate potential faster scale momentary coupling. The unexpected finding that speech tempo decreases with increased physical workload may be explained by multiple mental and physical factors that are more diverse and individual than anticipated. The implication for real-world contexts is that even light physical activity—functionally equivalent to walking—may impact speech tempo.

People commonly speak while performing physical activity, yet it is not fully clear how or why exercise may affect simultaneous speech. Current accounts identify the respiratory system as a key intermediary. For example, changes observed in glottal source and spectral measures have been linked to altered glottal configurations (Doust & Patrick, 1981; Godin & Hansen, 2011, 2015; Patil et al., 2010; Weston et al., 2021; Ziegler et al., 2019) arising from the deeper, faster breathing needed to maintain blood gas homeostasis as physical work increases (Dempsey et al., 2022; Travers et al., 2022). Altered breathing has also been linked to effects in the temporal domain, including increased articulation rate (Baker et al., 2008; Fuchs et al., 2015; Trouvain & Truong, 2015). Since more frequent breath taking increases total pause time, it is hypothesized that speakers produce more linguistic units per utterance to maintain overall information density. However, an alternative explanation for faster speech could lie in another physical by-product of exercise: The rhythmic leg movements needed to locomote. Could this self-generated periodic rhythm influence speech tempo? To date, research has yet to systematically investigate whether such motor action may affect speech during exercise.

Evidence of coupling between temporal patterns in speech and rhythmic limb movements comes from diverse strands of research. A large body of work shows that

Correspondence to Susanne Fuchs: fuchs@leibniz-zas.de. **Publisher Note:** This article is part of the Special Issue: Select Papers From the 8th International Conference on Speech Motor Control. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

humans spontaneously entrain motor actions to external auditory stimuli (e.g., Varlet et al., 2020), likely due to very basic neural entrainment between two oscillation-generating systems. In this regard, numerous published studies have assessed coordination between utterances and finger tapping in adults. For example, early experimental work by Kelso et al. (1983) investigated participants repeatedly tapping a finger while uttering the single syllable /pa/. Not only was it shown that tapping and uttering naturally coordinated in a 1:1 fashion, it was also found that when an inhalation was needed, the tapping anticipated the longer interval for the next utterance and adjusted accordingly to keep in sync (see also Parrell et al., 2014; Zelic et al., 2015). Going beyond tapping studies, there is a rich body of work on gestures showing that the communicative motions of the hand and arm are tightly coupled with prosodic and temporal features of speech, including prominence (e.g., Chu & Hagoort, 2014; Esteve-Girbert & Prieto, 2013; Rochet-Capellan et al., 2008; Shattuck-Hufnagel & Ren, 2018; for an overview, see Wagner et al., 2014).

There is also another type of coupling that may be generated through biomechanics. For example, moving the upper limbs rhythmically with sudden decelerations directly affects vocal amplitude by affecting chest wall kinematics at moments of physical impulse (e.g., Pouw et al., 2020). Based on such research, it has been proposed that vocal and upper limb coupling arises from biomechanical links with respiration (Pouw & Fuchs, 2022), akin to vocal–locomotor–respiratory links observed in other animals. For example, bats coordinate their wingbeats with echolocation pulses—timing the upstroke with the exhalation needed to vocalize is more energetically efficient (Koblitz et al., 2010). However, the mechanism is highly flexible and can be decoupled when either wing rate or vocalization rate changes in response to new situations, such as when maneuvering around an obstacle (Falk et al., 2015) or increasing echolocation pulse rate to more rapidly acquire sensory information (Stidsholt et al., 2021). Importantly, in humans, the limbs used for locomotion are further removed from the structures involved in respiration, so vocal–locomotor–respiratory interactions are generally less pronounced or absent altogether (Banzett et al., 1992; McDermott et al., 2003; O'Halloran et al., 2012; Perségol et al., 1991). Leg–respiratory decoupling may actually be functional, as this allows for free respiratory control during, for example, running (Bramble & Carrier, 1983).

The generalization of the aforementioned research to lower limb–speech coupling may be problematic in some cases. For one, upper limb–speech coupling may be specific to the upper limbs. For another, it has been shown that the physical impulse matters for effects on the respiratory(–vocal) system (Pouw & Fuchs, 2022; Serré et al., 2021), and therefore, mechanically guided movements with low physical impulses, such as cycling, may not be comparable to other actions, such as running, where foot landings lead to sudden decelerations or pulse-like upper limb movements (e.g., Pouw et al., 2021). Additionally, speech is not generally isochronous, so we may wonder whether tempo entrainment between speech and an isochronous weak biomechanical impulse such as leg motion would emerge.

Nevertheless, predictions about changes in speech tempo during physical activity may be drawn from several studies that investigated articulation rate during simultaneous leg-based exercise. These studies hypothesized that faster breathing may affect the temporal organization of speech. Baker et al. (2008) assessed speech read by 12 speakers of American English who cycled on a stationary ergometer at moderate and moderately high physical workloads (quantified as 50% and 75% of maximal oxygen consumption, respectively). Four short texts were read at 3-min intervals at each workload. Articulation rate (in syllables per second, minus pauses) significantly increased for both workloads compared to reading at rest, but there was no significant difference in articulation rate between the two workloads. At the higher workload, articulation rate also significantly increased over time. The study also found an increase in breath frequency, leading the authors to conclude that "[Articulation rate] initially altered to accommodate for the increased respiratory demand" (Baker et al., 2008, p. 1212).

Fuchs et al. (2015) investigated spontaneous speech in 11 German speakers cycling on a stationary ergometer at two fixed power output settings (70 W and 140 W). A list choice task was used to elicit semispontaneous speech. A significant increase in articulation rate was reported from rest to both exercise levels. Since breathing rate also increased, the authors speculated that speakers may try to compensate for more breath pauses by producing more material per utterance to maintain overall information density. In the third study, Trouvain and Truong (2015) used a single high-intensity exercise condition: 23 Dutch speakers (15 women, eight men) ran on a treadmill until volitional exhaustion. A passage was read before and immediately after exercise. The authors reported a postexercise increase in articulation rate of 25% compared to speech at rest, although this was not found to be significant. Here, faster speech was partly attributed to a general increase in arousal due to the physical stress of exercise. Like Baker et al. (2008) and Fuchs et al. (2015), Trouvain and Truong (2015) ventured that speakers may try to counteract the temporal effect of longer, more frequent breath pauses with faster speech.

These findings on articulation rate and their interpretations are somewhat challenged by a study that used high-CO_2 air to examine speech in high respiratory drive (Bailey & Hoit, 2002). Ten speakers of American English read a passage twice while seated, first breathing normal atmospheric air and then breathing air with 7% CO_2. Although breathing rate in the CO_2 condition increased at a comparable rate to the moderate exercise conditions in the studies above, no significant increase was found in articulation rate. This suggests that there is something about the physical activity itself that affects articulation rate. In the three exercise studies, participants ran on a treadmill or cycled on a stationary ergometer, thereby self-generating an isochronous (strictly periodic) rhythm with the motion of their legs. This opens up the possibility of some kind of lower limb–speech interaction.

In summary, our review of the literature shows that speech is faster during physical activity and that temporal patterns in speech may couple with rhythms generated by different movements of the body. However, it is not yet clear under what conditions such coupling would arise or how motion rate and speech tempo would be affected. That this area of study is undeveloped is a pity, as it would deepen our understanding not only of limb–speech interactions but also of the body systems that act on speech during everyday physical activities—two areas that could provide useful insights for clinical applications.

Additionally, the concept of speech tempo invokes complex questions about different linguistic timescales and whether they reflect a global temporal organization of speech. The related concept of speech rhythm remains under fervent debate, and different metrics have been proposed using various timing patterns, including linguistic units such as syllables and stress feet (for an overview, see Fletcher, 2010). An alternative approach is to operationalize speech tempo directly from the speech signal: The smoothed amplitude envelope eschews categorical linguistic units and captures a continuous representation of important rhythmic structure in speech (Tilsen & Arvaniti, 2013). Several signal processing methods exist to calculate an amplitude envelope, which roughly tracks gross changes in the amplitude of the speech signal (for a recent overview, see MacIntyre et al., 2022).

The use of the amplitude envelope has been further enriched with methods to isolate multiple timescales in the signal (Tilsen & Arvaniti, 2013). Empirical mode decomposition is based on another signal-processing transformation, the Hilbert–Huang transform (Huang et al., 1998), which allows the extraction of particular oscillations in a signal based on the shortest timescale present at a particular moment. Applied iteratively, shorter and longer timescale oscillations can be isolated from the signal. Tilsen and Arvaniti (2013) have used this method to approximate "syllable- and stress-driven fluctuations in the envelope" (p. 631). We selected this approach in the current study for two reasons. First, it accounts for phonetic reduction, which is neglected when articulation rates are calculated using phonological syllables. Second, it is possible that slower quasi-rhythmic aspects of speech may relate to leg motion rate, which may be overlooked if we look at coupling only on syllable-level timescales.

Research Questions

This study explores the relationship between lower limb motion and the tempo of simultaneous speech, specifically the following.

1. Does increased physical effort affect leg motion rate?
2. Does increased physical effort affect speech tempo?
3. Does speech tempo show any relation to leg motion rate? Specifically, how well does leg motion rate correlate with (a) the amplitude envelope; (b) a faster, syllable-like speech tempo; and (c) a slower, footlike speech tempo?
4. Is there an effect of speech task? Specifically, do the different cognitive loads associated with read and spontaneous speech affect leg motion rate, speech tempo, or the strength of correlation between them?

Method

Speakers and Speech Tasks

Data were recorded from 48 healthy, nonsmoker female[1] native speakers of standard German. The study was approved by the ethical board of the Linguistic Society of Germany. Participants self-selected by responding to invitations to take part in a study on speech during exercise sent via a participant database to all self-identified women between the ages of 18 and 35 years. All participants were paid €25 for participation in the 2-hr experiment; none reported any speech, hearing, or breathing pathologies or any physical disabilities. The participants were relatively young (x = 23.6 years old, SD = 3.8) and recreationally active: The International Physical Activity Questionnaire showed the weekly level of physical activity to be high for 58% (n = 28), moderate for 38%

[1]While it is acknowledged that sex is not binary, this wording refers to that used in the recruitment text and the way in which the respondents self-identified. Participation was restricted to a single assigned biological sex to obtain statistical power: Considerable sex differences in respiratory physiology (LoMauro & Aliverti, 2018) and breathing during exercise (Horiuchi et al., 2019; Sheel et al., 2004) would require sex be included as an additional factor in statistical analysis.

(n = 18), and low for 4% (n = 2). Mean body mass index was 22.0 (SD = 3.5), which falls within the healthy weight range. Mean resting heart rate was 61.3 beats per minute (bpm; SD = 9.3). For comparison, female college endurance athletes have been reported to have mean resting heart rates between 47 and 57 bpm (Gademan et al., 2012; Uusitalo et al., 1998), and a recent large-scale study of the general population reported a 95% confidence interval of heart rate values of 56–113 bpm for women aged 18–20 years (Avram et al., 2019).

Three speech tasks were recorded in counterbalanced order across participants. The current study analyzes data from two tasks: (a) reading a passage and (b) speaking freely about preferences. The two tasks (described below) differed in terms of cognitive load and spontaneity. In the reading task, participants were expected to produce continuous speech without the hesitation disfluencies (e.g., prolongations) that characterize spontaneous speech. Because the text was repeated across conditions, the speech material was controlled. In the freely spoken task, participants were expected to produce more naturalistic speech but with disfluencies and increased online planning load that could affect speech tempo.

For the reading task, participants read a 126-word passage 3 times per condition. To approximate a spoken register, the passage was generated from unscripted speech as follows: A 24-year-old female German speaker was asked to respond orally to the question: "What does one need to organize the perfect party?" Her response was transcribed, the disfluencies were removed, and punctuation was added. Participants were told that the text stemmed from spoken speech, and they read it aloud once for familiarization prior to the experiment.

Spontaneous speech was elicited in the form of unscripted monologues using nine stimulus questions about general preferences ("What is your favorite season and why?") and opinions ("How do you feel about virtual lectures/meetings?"). Question order was randomized. Participants spoke for roughly 2 min per trial, with three trials per condition. The stimulus questions were selected by conducting an online study with 30 female native German speakers, 18–35 years old, who did not participate in the data-recording experiment. The participants rated 30 potential questions on two criteria, ease of discussion and emotional neutrality, using a 5-point Likert scale following the design in Mitchell et al. (1996, p. 96). Average ratings for each question were calculated, and the nine questions rated easiest to talk about and least emotionally charged were selected for the experiment. The topic for the reading task was selected based on emotional neutrality and the inclusion of the impersonal pronoun "one" to elicit an impersonal rather than first-person text. Speech materials are given in the Appendix.

Exercise Task and Workload Levels

Participants exercised on a low-noise stationary bicycle (daum electronic) at two exercise levels to approximate the physical effort ("workload") respectively required by daily activities (e.g., walking) and certain occupations (e.g., fitness instruction). The speech tasks and exercise task were designed to allow multiple phonetic analyses to be conducted on the data. For this reason, non–weight-bearing exercise was chosen to avoid the sound of footfalls, which would have limited acoustic analyses. It was also deemed safer for the participant to be stationary when reading from the monitor. Cycling was chosen because the posture stabilizes the upper body, which was important because respiratory data were recorded to pursue other research questions.

The workload level for each condition was defined using a common measure of exercise intensity, percentage of heart rate reserve (HRR). HRR equals maximum heart rate (HR_{max}) minus resting heart rate (HR_{rest}). Low workload was defined as 35% of HRR and moderate workload was defined as 65% HRR based on guidance provided by public health organizations (e.g., Centers for Disease Control and Prevention, 2022). These percentages were translated into participant-specific target heart rates (in bpm) using a standard method in sports science, the Karvonen formula, given in Equation 1.

$$\text{target HR} = [(HR_{max} - HR_{rest}) \times \%\text{intensity}] + HR_{rest} \quad (1)$$

This formula version followed Tanaka et al. (2001) to predict HR_{max} using 208 − (0.7 × age).[2] HR_{rest} was estimated by having each participant lie supine for 10 min and taking the average HR of Minute 11 using a wrist-worn HR monitor worn throughout the experiment. Incorporating age and resting pulse accounted for individual physiological differences that affect exercise performance. The physical effort required to attain the individualized target HRs was thus comparable across participants.

The calculated target HRs for each condition were validated by participants using a subjective measure, the Borg Rating of Perceived Exertion (Borg, 1982), a 15-point scale ranging from 6 to 20. After the experiment, participants were asked to rate their level of physical effort for cycling only. The average rating for low

[2]Age-based prediction equations deviate somewhat from an individual's true HR_{max}, but they are widely used when graded exercise testing is not feasible. Tanaka et al.'s (2001) equation has been found to underestimate HR_{max} in young women by 7–8 bpm (Lach et al., 2021). This means that the participants were likely exercising at slightly less than 35% and 65% of HRR.

workload was 9.7 (“very light”) and that for moderate workload was 13.8 (“somewhat hard”).

Recording

The study used a within-participant design with three workload conditions: (a) rest, (b) cycling at 35% of HRR (low workload), and (c) cycling at 65% of HRR (moderate workload). Workload condition order was fixed (rest > low > moderate) because the cardiovascular system may take up to several hours to recover after exercise has ceased (Romero et al., 2017). If workload condition order were randomized, it would thus be difficult to ensure that rest captured baseline speech. After the rest condition, participants cycled at a self-selected pace for approximately 4 min to warm up. Resistance was added as needed to bring each participant to their target heart rate for low workload. After speech tasks were completed at low workload, resistance on the bicycle was increased during another 4-min workload transition to reach the target heart rate for moderate workload. Heart rate was monitored by the experimenters in real time using a tablet readout and was kept at the target by adjusting resistance between speech trials.

Workload was manipulated by increasing resistance, rather than pace, for two reasons. First, the ability to maintain a given pace depends on level of fitness. Since the participants were not uniformly well trained, using a single pace in each condition would have introduced uncontrolled variation in terms of the amount of effort each participant was expending. Second, having to maintain a given pace, for example, by synchronizing with a metronome, would have imposed entrainment to an external stimulus (instead of self-entrainment) as well as an additional cognitive load that could shift attention away from reading and speaking and affect speech tempo.

Speech data were recorded at a sampling rate of 22050 Hz using a head-mounted microphone (beyerdynamic) placed 4 cm from the corner of the mouth at a 90° angle. Leg movement rate (cycling cadence) was recorded using an OptiTrack motion capture system (Prime 13 cameras) and its Motive software platform (Version 1.90). A reflective OptiTrack marker was affixed to one of the pedals and tracked by one motion capture camera positioned at pedal level, approximately 2 m from the bicycle, and five motion capture cameras positioned in a half-circle at ceiling height. Respiratory data were recorded for a different research question but are not reported here. Because breath timing is dictated by linguistic structure and utterance length can range from 1 to 12 s, a mean-based measure is not useful to assess possible respiratory–locomotor coupling.

Data Preprocessing

Calculation of Amplitude Envelope, Intrinsic Mode Function 1, and Intrinsic Mode Function 2

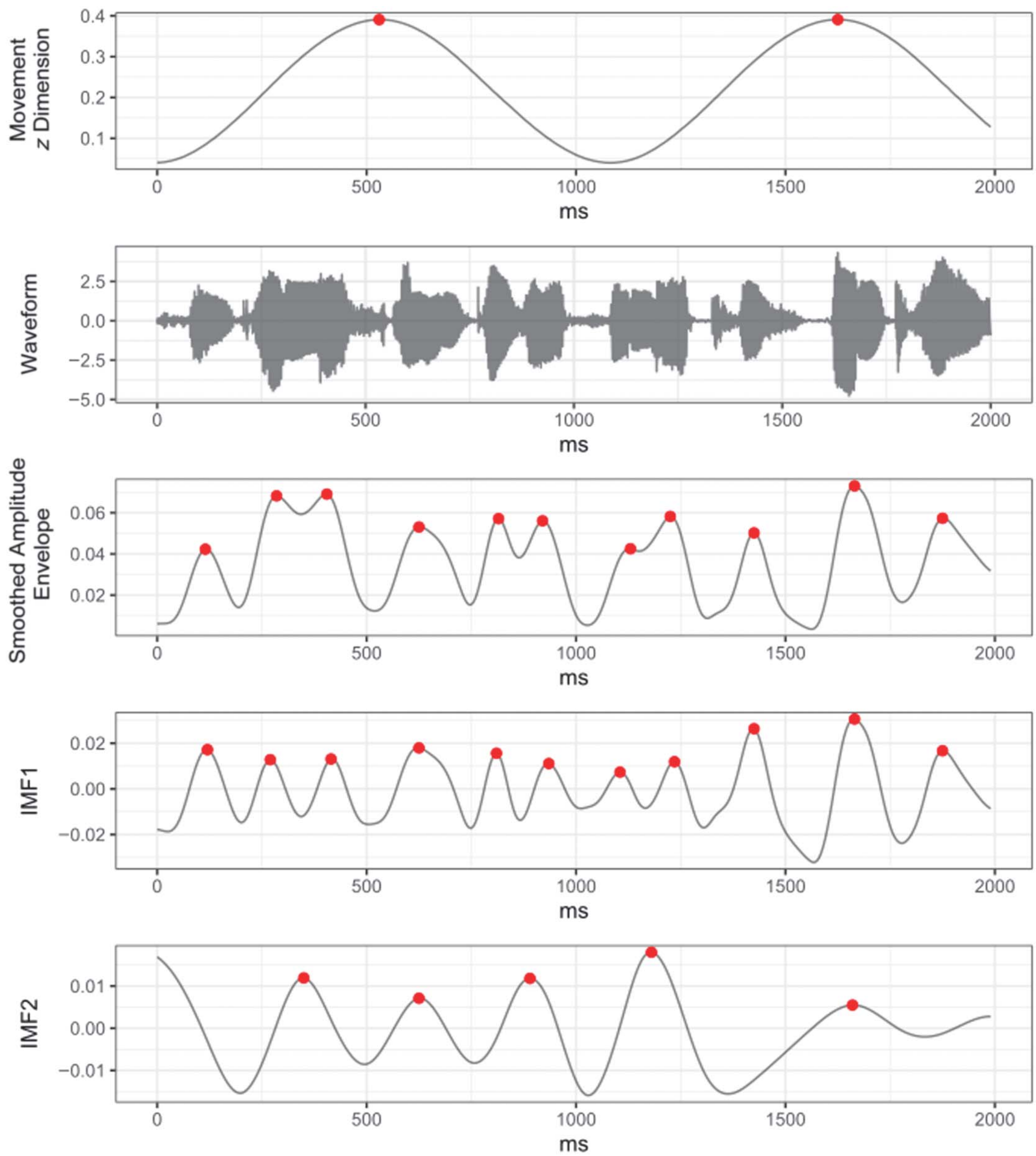
The amplitude envelope was generated by taking the complex modulus from the Hilbert transform of the audio waveform using a custom R script (Pouw & Trujillo, 2021). This yields a one-dimensional signal that was then smoothed with a type of window function: an 8-Hz Hann filter a.k.a. a filter that is based on applying a Hanning window in the frequency domain, which tries to maximize the preservation of the dominant frequency in the signal (we used R-package dlpR). Since speech during exercise differs from neutral speech in ways that affect the amplitude envelope (e.g., greater aspiration noise), we qualitatively checked some trials to determine the best filter. We concluded that 8 Hz would better fit our data than the 10 Hz used by Tilsen and Arvaniti (2013). The smoothed amplitude envelope was then downsampled to 200 Hz.

Further speech tempos were obtained from the amplitude envelope using empirical mode decomposition, a method based on a Hilbert–Huang transform (Huang et al., 1998), which extracts from a complex signal or multiple signals that follow specific intrinsic timescales present in the signal at a moment in time. These intrinsic mode functions (IMFs) are extracted in an iterative way based on specific rules that consider minima, maxima, and zero-crossings (Huang et al., 1998), such that after each iteration, a residue signal is left from which the said summed IMFs are subtracted. This method is optimal for understanding the complex rhythms that are empirically present in the signal over time. We extracted two IMFs, with the first function (IMF1) capturing the faster scale, syllable-like fluctuations in the smoothed amplitude envelope and the second function (IMF2) capturing slower scale, foot-like fluctuations. We determine time intervals based on local maxima (peaks) in the amplitude envelope, IMF1, and IMF2. We then scale the time interval such that it is expressed in intervals/cycles per second (Hz). After creating the different envelope time series, we applied a peak-finding function to determine local maxima using the pracma package (Borchers, 2022) in R. Figure 1 shows labeled peaks in each measure.

Motion Capture Data

The motion capture data consisted of CSV files with the *x*-, *y*-, and *z*-coordinates of the OptiTrack pedal marker over time. Since the pedaling motion is circular, we used only the vertical dimension to identify a full revolution of the pedal, resulting in a movement trace of peaks and valleys (see Figure 1). The OptiTrack cameras are

Figure 1. Labeled peaks for cadence and speech tempi. The upper panel shows vertical displacement of the pedal; the second panel shows the raw audio waveform, and beneath it, the panels show amplitude envelope, IMF1, and IMF2. IMF = intrinsic mode function.



highly sensitive and occasionally record other reflective objects (e.g., athletic shoes) for short periods, which may introduce spurious peaks in the data. To reliably identify the true peaks, the motion signal data were smoothed by removing noise-related jitter using a 10th-order Kolmogorov–Zurbenko filter with a window of 10, performed using the kza package (Close et al., 2020) in R. The peaks were then automatically labeled using the findpeaks function (Borchers, 2022) and subsequently checked manually. Leg cycling rate (cadence) for each trial was calculated by averaging the duration between each peak (corresponding to one revolution of the pedal) and reported in hertz for direct comparison with the rates of the envelope, IMF1, and IMF2.

Statistical Analyses

Statistical analyses were run in R (Version 4.2.2; R Core Team, 2022) using the RStudio environment (Version 2022.07.2 + 576). Linear mixed-effects models for cycling cadence, envelope rate, IMF1, and IMF2 were estimated using the lme4 package (Bates et al., 2015). The degrees of freedom required to obtain p values were obtained using the lmerTest package (Kuznetsova et al., 2017). Independent variables were physical workload condition (rest, cycling at 35% of HRR [low], cycling at 65% of HRR [moderate]), speech task (read speech and spontaneous speech), and trial order per condition (1, 2, 3). Repetition order in each condition ("trial order") was included because Baker et al. (2008) found a small increase in articulation rate across the first three repetitions of their experiment. We started with a model including the interaction between workload and task plus trial and speaker-specific random slopes for the interaction. In some cases, these models did not converge or a high collinearity between the two factors were found using the performance package (Lüdecke et al., 2021) for diagnostics. Speaker-specific random slopes were applied for the fixed factors. All scripts and data for the reproduction of statistical results are provided at https://github.com/sfuchs18/movementspeechtempo_public. Dependent variables were cycling cadence, the amplitude envelope, IMF1, and IMF2 (all in hertz). Prior to analysis, all observations above and below 3 *SD*s from the group mean were excluded. Given that we are testing multiple hypotheses in an exploratory way, we set the significance threshold at $p = .01$.

Leg Cycling Rate (Cadence)

We first ran a model including the interaction between workload and task as fixed factors, trial (as additive), and speaker-specific random slopes for the interaction. The model showed high collinearity of the two fixed factors. For this reason, we chose a model without interaction (motion with no speech as a reference level) and found a main effect of workload only ($\beta = .064$, $t = 4.146$, $p < .001$; see Figure 2). Group means and standard deviations are reported in Table 1.

Amplitude Envelope, IMF1, and IMF2

Group means and standard deviations for all measures are reported in Table 2; data are plotted in Figure 3. Figures for trial comparisons are given in the Appendix. All coefficients were significant ($|t| > 3.5$, $p < .001$; see the Appendix). Under both workload conditions, speakers decreased their amplitude envelope rate, even more under moderate workload. Envelope rate also differed when comparing spontaneous with read speech with the former being slower than the latter. Additionally, an interaction between workload and task was found. In the reading task, speaking without motion (rest) and speaking under low effort differed, whereas in spontaneous speech, they did not.

For IMF1, we found significant effects for all predictors ($|t| > 7.2$, $p < .001$) apart from trial. The direction of the effects is similar to the envelope with a decrease in IMF1 with workload and a slower rate for spontaneous than read speech.

IMF2, the slowest rhythm, shows a comparable outcome, again with significant main effects but no interaction for all predictors ($|t| > 5.17$, $p < .001$) except the comparison between Trials 1 and 3 ($\beta = -.036$, $t = 3.231$, $p = .001$). No matter which speech rate we considered (envelope, IMF1, IMF2), both workload and speech task had a strong effect.

Figure 2. Box plots with violin plots for leg cycling rate (in Hz) across different speech conditions and workloads.

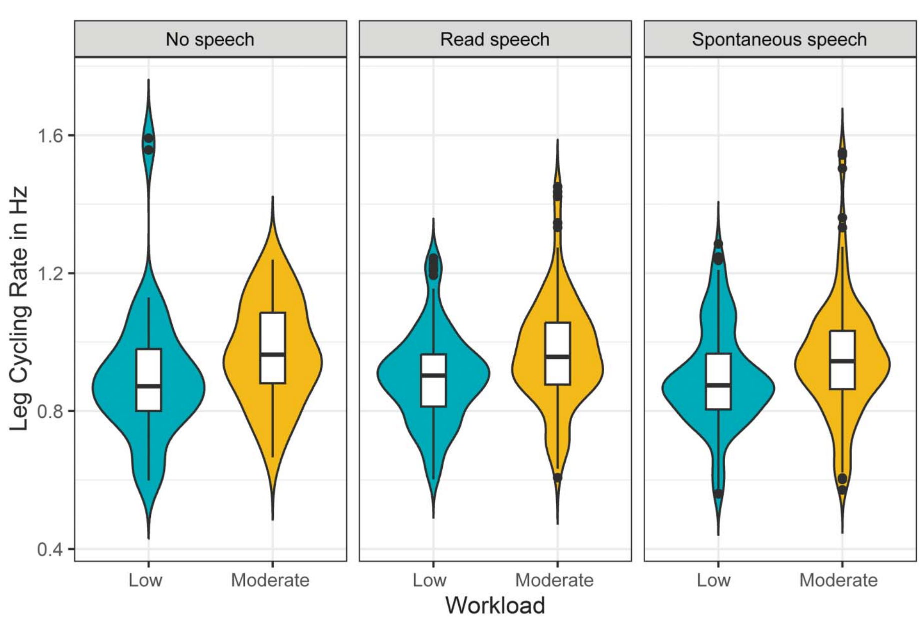


Table 1. Group means (standard deviations) for cycling rate in hertz and revolutions per minute.

Condition	Revolutions per second (Hz)			Revolutions per minute		
	Cycling only	Read text	Speech	Cycling only	Read text	Speech
Low	0.90 (0.20)	0.90 (0.13)	0.90 (0.14)	54.1 (11.8)	53.9 (7.8)	53.9 (8.6)
Moderate	0.97 (0.15)	0.97 (0.16)	0.95 (0.17)	58.0 (8.8)	58.1 (9.4)	57.1 (10.2)

Correlation Between Cadence and Speech Tempi

To investigate possible relations between leg cycling rate and speech tempo, three linear-mixed effects models were calculated with cadence as a predictor for speech outcomes (amplitude envelope, IMF1, IMF2). The model was as follows: lmer(y ~ scale(cadence) + workload + speech task + trial order + (1 + cadence + speech task | speaker)). Cadence was scaled (i.e., z scored) in the model, but not in Figure 4, in order to allow for a better understanding of the raw values. None of the regression models turned out to be significant for cadence as a predictor (envelope: $\beta = -.044$, $t = -1.744$, $p = .1$; IMF1: $\beta = -.049$, $t = -1.124$, $p = .232$; IMF2: $\beta = -.0157$, $t = -0.823$, $p = .443$). Thus, our findings do not provide evidence of a relation between cadence and speech tempo at the group level. A large variance is visible overall in the data, which suggests a more complex interplay of factors we may not have considered here.

Discussion

Leg Cycling Rate Increased With Physical Workload

Leg cycling rate (cadence) was compared between workloads. A significant increase in cadence was found for both speech tasks from low workload to moderate workload. Expressed in revolutions per minute (rpm), the mean difference between the workload conditions is evidently small (3–4.8 rpm). For comparison, studies that assess performance response to different cycling cadences typically look at differences of 15–20 rpm (e.g., Argentin et al., 2006; Foss & Hallén, 2005).

A likely explanation for the observed increase in leg cycling rate at the higher workload is that faster cadences improve pedaling efficiency (Dantas et al., 2009). Pedaling faster decreases the force of each pedal stroke—the leg muscles contract more often but less powerfully, which uses metabolic energy more efficiently and delays muscle fatigue. For these reasons, cadence has also been found to increase as power output increases (Brisswalter et al., 2000; Coast & Welch, 1985). In the current study, power output (resistance) was increased to increase physical workload. A possible explanation for the model-predicted increase in cadence for Trial 3 is thus the early onset of muscle fatigue.

At the same time, a relatively large range of cycling cadences was observed (0.56–1.54 Hz; 33.6–92.4 rpm). Work in sports science has demonstrated that cadence is affected by diverse body characteristics, including leg muscle strength (Hansen et al., 2007) and fitness level (Bieuzen et al., 2007), as well as factors specific to the bout of exercise, such as resistance/grade (Lucía et al., 2001) and saddle position (Lee & Park, 2021). Self-selected cadences will thus be highly individual, which may explain the variability observed in our data.

Speech Rates Slowed With Increased Physical Workload

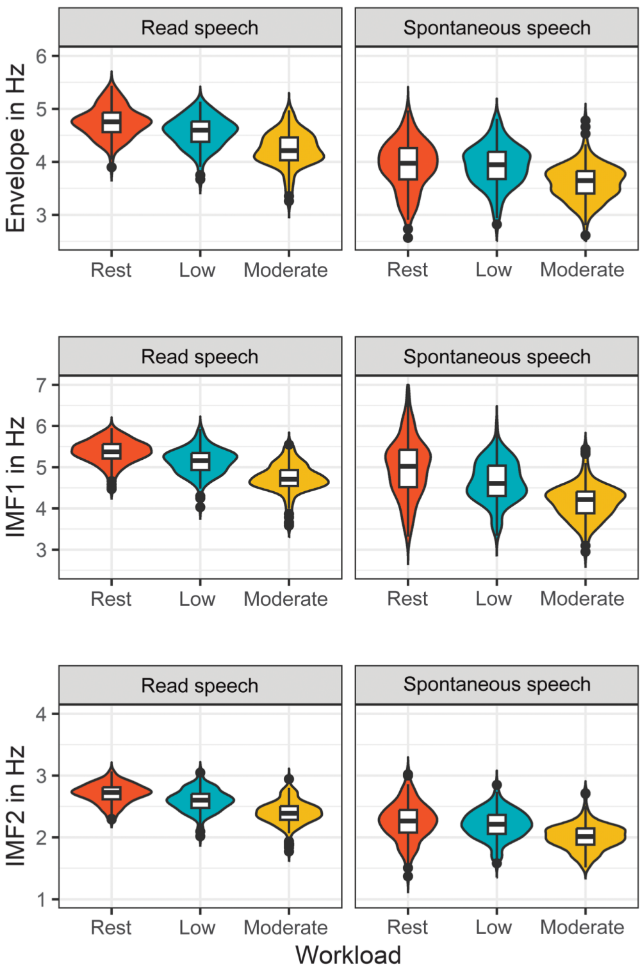
Unexpectedly, we found that speech rates decreased as physical workload increased. This finding is contrary to those reported in previous studies (Baker et al., 2008; Fuchs et al., 2015; Trouvain & Truong, 2015). A possible explanation may lie in methodological differences. In our study, participants exercised for roughly 45 min in total, about twice the time in the other studies. Hence, it is possible that our participants' speech may have slowed due to fatigue over time, although this would not explain the

Table 2. Group means (standard deviations) for each speech tempo per condition and speech task.

Condition	Envelope		IMF1		IMF2	
	Read text	Speech	Read text	Speech	Read text	Speech
Rest	4.75 (0.29)	3.93 (0.47)	5.37 (0.29)	5.0 (0.72)	2.71 (0.16)	2.24 (0.28)
Low	4.56 (0.29)	3.93 (0.39)	5.14 (0.32)	4.65 (0.50)	2.59 (0.19)	2.21 (0.23)
Moderate	4.23 (0.33)	3.62 (0.35)	4.70 (0.37)	4.16 (0.43)	2.39 (0.21)	2.01 (0.21)

Note. IMF = intrinsic mode function.

Figure 3. Box plots with violin plots for speech tempo (in Hz) as measured by amplitude envelope, IMF1 and IMF2 for each speech task and workload. IMF = intrinsic mode function.



decrease in the earlier (low workload) condition, the duration of which was comparable to the other studies. Another methodological difference is that the aforementioned studies quantified articulation rate using the phonological syllable, whereas we used something akin to the phonetic syllable (IMF1). Phonological syllables do not capture phonetic reduction, so if phonetic reduction increased with physical workload, articulation rate would also increase. However, given the opposite direction of effects, it seems unlikely that phonetic reduction would increase that much.

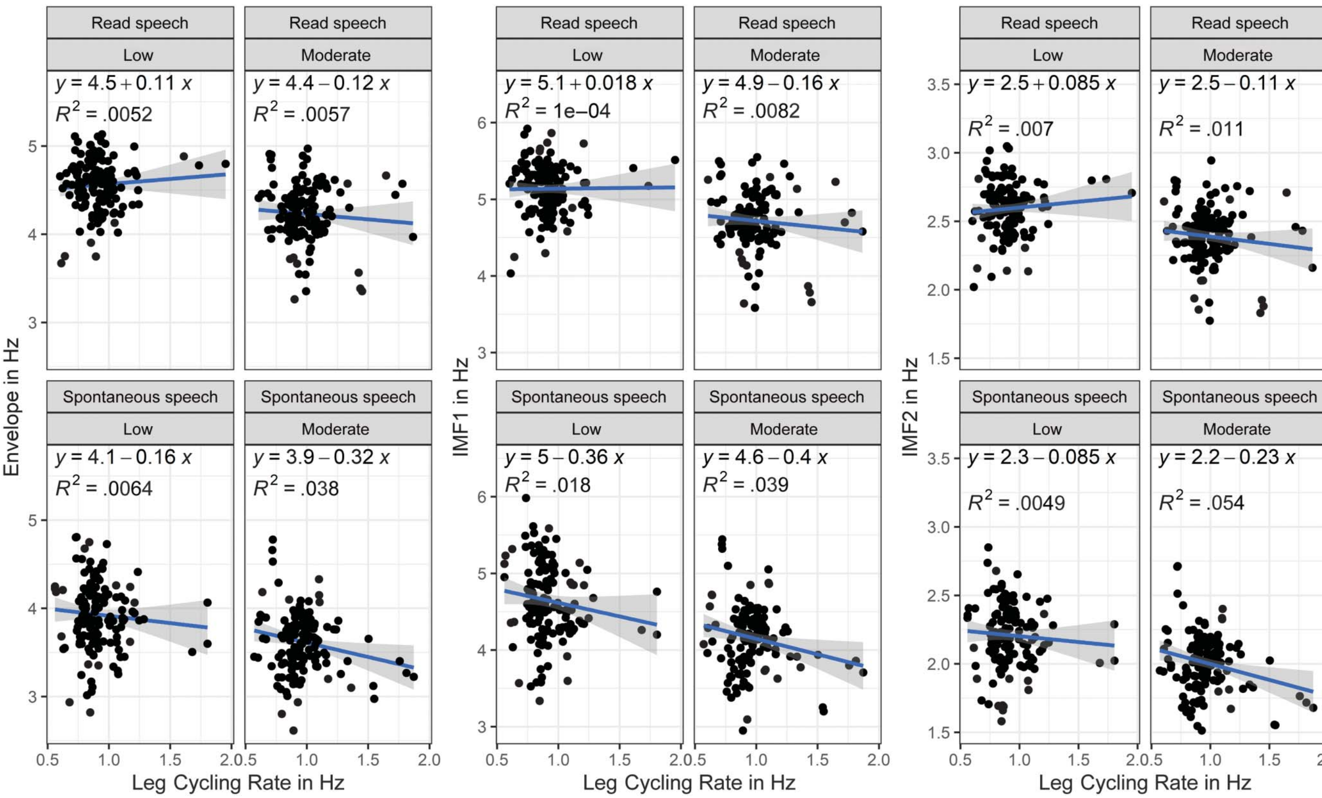
An alternative approach would be to view the experiment as a dual-task paradigm—doing two things at once. For example, walking-while-talking studies typically show decreased performance in one or both tasks (e.g., Lamberg & Muratori, 2012; Plummer-D'Amato et al., 2011; Raffegeau et al., 2018) due to inadequate attentional resources (Kemper et al., 2003). Factors inherent to the exercise task could also be distracting. For example, in a study design similar to ours, Ziegler et al. (2019) hypothesized that "the increased concentration (mental effort) to breathe at higher intensities of respiratory drive (Hoit et al., 2007) may reflect a vulnerability of behavioral vocal motor control to the needs of automatic respiratory control" (p. 2585). This presents an intriguing possibility in the current study: Did factors incidental to the exercise task distract speakers? The motion capture and speech data analyzed here were taken from a larger corpus that included qualitative data in the form of written free-form questionnaires completed postexperiment. Participants were asked (a) how difficult they found each speech task at each workload, (b) whether cycling affected their speech, and (c) whether they felt uncomfortable because they were out of breath. Pursuing the hypothesis that breathing-related sensations may distract speakers, we searched the questionnaires for statements pertaining to concentration and dyspnea. The results are summarized here; comments are cited using participant codes (three-letter strings).

Almost half (22) of participants mentioned difficulty concentrating on the speech tasks. Illustrative comments are given in Table 3. Some described general strain on attentional resources, whereas others reported task-specific difficulties. For example, 11 speakers reported that breath planning made reading more difficult during exercise. During the spontaneous speech task, eight participants felt their responses were less creative or less coherent. However, five speakers perceived a positive effect of exercise. These comments indicate that participants experienced the different speech tasks differently. Furthermore, they suggest that perceived speech difficulty does not always correlate with physical workload. These findings are in line with those obtained by Rotstein et al. (2004), in a study that examined the relation between individual perceptions of speech difficulty and physiological response variables.

Almost three quarters (34) of participants reported some other kind of difficulty related to exercise. For example, 13 participants cited concern with their reading performance, six said they experienced dyspnea, and six reported feeling self-conscious. In summary, most (85%) participants described at least one exercise-related difficulty—typically dyspnea, speech-performance concerns, or uncomfortable sensations—and 40% reported experiencing more than one.

Finally, it bears noting that three participants mentioned slower speech. One cited a distraction: "My legs were kind of burning and I couldn't block it out, this affected speaking (slower)" (speaker rno). Two participants reported being physically unable to speak faster, though no details were given. Speaker fpk wrote: "Speech became slower because of the exercise, but the thoughts came just as quickly." For reading, speaker php wrote: "The longer

Figure 4. Correlations between leg cycling rate and speech tempi for both workloads and speech tasks.



the biking lasted, the harder it got to read, because I knew the text and wanted to read faster than was possible."

Taken together, these insights into participant experience led us to three conclusions that may explain both the decrease and speaker variability in speech tempi with exercise. First, it seems that the cycling task gave rise to various distractions that could have divided participants' attention and thus slowed speech tempo. Second, participants perceiving multiple distractions may have seen a greater effect on speech tempo. Third, exercise seems to have had a negative impact on spontaneous speech performance for some participants but a positive impact for others. The latter point also suggests that perceived speech difficulty is not always aligned with level of physical workload.

Table 3. Illustrative comments describing difficulties and distractions related to exercise.

Read speech	
Distraction	**Illustrative comment**
Divided attention	"With increasing effort, more effortful to speak because you had to concentrate on talking + thinking + biking + breathing" (speaker kgj)
Breath planning	"Suddenly I had to think about when I pause for breath, which I don't normally do" (speaker npn)
Embarrassment	"I was embarrassed that I was breathing so loudly at the end" (speaker ilu)
Sound of voice	"My voice was breaking, like when you're about to cry" (speaker hct);
Performance	"When I took deeper breaths, I could no longer read very smoothly" (speaker gjs)
Spontaneous speech	
Perceived effect	**Illustrative comment**
Less creative	"Creativity declined as effort increased" (speaker hsx)
Less coherent	"It was more difficult to form a complete/logical sentence" (speaker yvc)
More free	"The harder I worked, the easier it became for me to talk freely" (speaker qlk)
Distraction from discomfort	"With topics I liked I almost forgot about the biking and had no problem speaking and breathing" (speaker csm)

No Correlations Between Leg Cycling Rate and Speech Tempi

We found no evidence of a correlation between leg cycling rate and any of the speech tempi. One explanation could lie in our data. It may be difficult to identify consistent rhythmicity when speech stretches vary greatly in duration: short stretches (< 1 s) may not contain enough information; long stretches (> 3 s) may contain sections of different tempos (Tilsen & Arvaniti, 2013). Our data showed considerable variability in speech stretch duration, particularly in the spontaneous speech task. Further research should therefore focus on more moment-to-moment coupling between leg and speech cycles.

Another explanation could simply be that the temporal variability of continuous speech is too great to allow consistent temporal coupling with an external rhythm. Numerous studies have shown that speech timing patterns are highly dynamic and affected by factors at different linguistic levels, from prosodic phrasing (e.g., Stehwien & Meyer, 2022) to stylistic differences (e.g., Volín, 2023). In light of this, some researchers have asked whether speech may actually be arhythmic (Nolan & Jeon, 2014). Related to this is the differing motor nature of the speaking and cycling tasks. Articulatory movements must be actively and continuously planned, while the rhythmic leg motion of pedaling is a largely automatic, non–weight-bearing exercise, with more stable balance and less neuromuscular regulation than walking. Such automatic motions may be controlled in a different way than movements that require continuous cognitive control (Schaal et al., 2004). A comparison of the effects of load-bearing and non–load-bearing exercise on speech would therefore be important to pursue.

Limitations

Exercise was performed on a stationary bicycle, and resistance was increased to indirectly manipulate leg cycling rate. An alternative approach would have been to control leg cycling rate with an external stimulus like a metronome. However, we decided against this because it would have introduced uncontrolled variation in physical effort across participants as well as potentially requiring additional attention to entrain pedaling to the beat. Moreover, the entrainment would have been to an external auditory stimulus than the intrinsic rate entrainment we are interested in. Another potential limitation of cycling exercise is that leg accelerations are kept relatively stable to create a continuous cycling output. Other locomotor actions, such as walking, that generate biomechanical impulses through the force of heel-strikes (Daley et al., 2013) could make it more likely for rhythmic stabilities to arise between leg motion and speech rates. A direct comparison of walking and cycling would thus be an interesting future research direction.

The speech tasks may also have influenced articulation rates. The long sentences in the reading task reflected spoken language, but some speakers found them challenging to read, and this may have increased cognitive load. We are also aware of the drawbacks of using monologues to elicit spontaneous speech. Although this method is common in other subfields (e.g., language documentation), speakers typically self-select topics. In our qualitative data, some speakers reported unequal interest in the topics or initial acclimation to the task, which likely accounts for some variability, particularly in the first condition (speech at rest).

Finally, the condition order was fixed to prevent exercise-related physiological changes from affecting speech at rest. A study over several days would allow for a counterbalanced condition order and shorter bouts of exercise, which would perhaps separate the potential effects of fatigue from workload level.

Conclusions

The temporal communicative patterns that humans can generate within any bodily system do not emerge in a vacuum. Those patterns are anchored within other subsystems that interact on several levels, among them cognition, biomechanics, and the communicative context. We find that temporal patterns in speech are likely affected by the leg locomotion system on a cognitive level during simultaneous cycling and speaking. Our study shows that speech rates decreased with increased physical workload, with little to no direct entrainment between leg cycling rate and speech rate. Several open questions about potential speech–locomotor coupling now become emphasized through this research: Is cycling different from running in interactions with speech? How do we disentangle cognitive split attention due to increased respiratory load from other types of interactions with the speech system, such as biomechanics? Could there be faster scale momentary coupling between leg motion and speech, and how would we measure that? The current study thereby furthers our inquiry into the dynamic openness of the speech system.

Data Availability Statement

The full data set will be made available in online supplemental materials (https://github.com/sfuchs18/movementspeechtempo_public).

Acknowledgments

This work was supported by a Deutsche Forschungsgemeinschaft grant (FU791/8-1) awarded to the last author. The second author is supported by a VENI grant (VI.Veni.201G.047: PI: Wim Pouw) awarded by the Dutch Research Council (NWO). The authors would like to thank the reviewers and the study participants as well as Jörg Dreyer for technical support and Theres Weißgerber for assistance with data annotation.

References

Argentin, S., Hausswirth, C., Bernard, T., Bieuzen, F., Leveque, J., Couturier, A., & Lepers, R. (2006). Relation between preferred and optimal cadences during two hours of cycling in triathletes. *British Journal of Sports Medicine, 40*(4), 293–298. https://doi.org/10.1136/bjsm.2005.020487

Avram, R., Tison, G. H., Aschbacher, K., Kuhar, P., Vittinghoff, E., Butzner, M., Runge, R., Wu, N., Pletcher, M. J., Marcus, G. M., & Olgin, J. (2019). Real-world heart rate norms in the Health eHeart study. *NPJ Digital Medicine, 2*(1), Article 58. https://doi.org/10.1038/s41746-019-0134-9

Bailey, E. F., & Hoit, J. D. (2002). Speaking and breathing in high respiratory drive. *Journal of Speech, Language, and Hearing Research, 45*(1), 89–99. https://doi.org/10.1044/1092-4388(2002/007)

Baker, S. E., Hipp, J., & Alessio, H. (2008). Ventilation and speech characteristics during submaximal aerobic exercise. *Journal of Speech, Language, and Hearing Research, 51*(5), 1203–1214. https://doi.org/10.1044/1092-4388(2008/06-0223)

Banzett, R. B., Mead, J., Reid, M. B., & Topulos, G. P. (1992). Locomotion in men has no appreciable mechanical effect on breathing. *Journal of Applied Physiology, 72*(5), 1922–1926. https://doi.org/10.1152/jappl.1992.72.5.1922

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

Bieuzen, F., Vercruyssen, F., Hausswirth, C., & Brisswalter, J. (2007). Relationship between strength level and pedal rate. *International Journal of Sports Medicine, 28*(7), 585–589. https://doi.org/10.1055/s-2007-964859

Borchers, H. W. (2022). *Pracma* (2.4.2) [Computer software]. https://cran.r-project.org/web/packages/pracma/pracma.pdf

Borg, G. A. V. (1982). Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise, 14*(5), 377–381. https://doi.org/10.1249/00005768-198205000-00012

Bramble, D. M., & Carrier, D. R. (1983). Running and breathing in mammals. *Science, 219*(4582), 251–256. https://doi.org/10.1126/science.6849136

Brisswalter, J., Hausswirth, C., Smith, D., Vercruyssen, F., & Vallier, J. M. (2000). Energetically optimal cadence vs. freely-chosen cadence during cycling: Effect of exercise duration. *International Journal of Sports Medicine, 21*(1), 60–64. https://doi.org/10.1055/s-2000-8857

Centers for Disease Control and Prevention. (2022). *Target heart rate and estimated maximum heart rate*. https://www.cdc.gov/physicalactivity/basics/measuring/heartrate.htm

Chu, M., & Hagoort, P. (2014). Synchronization of speech and gesture: Evidence for interaction in action. *Journal of Experimental Psychology: General, 143*(4), 1726–1741. https://doi.org/10.1037/a0036281

Close, B., Zurbenko, I., & Sun, M. (2020). *Kza* (4.1.0.1) [Computer software]. https://cran.r-project.org/web/packages/kza/kza.pdf

Coast, J. R., & Welch, H. G. (1985). Linear increase in optimal pedal rate with increased power output in cycle ergometry. *European Journal of Applied Physiology and Occupational Physiology, 53*(4), 339–342. https://doi.org/10.1007/BF00422850

Daley, M. A., Bramble, D. M., & Carrier, D. R. (2013). Impact loading and locomotor-respiratory coordination significantly influence breathing dynamics in running humans. *PLOS ONE, 8*(8), Article e70752. https://doi.org/10.1371/journal.pone.0070752

Dantas, J. L., Smirmaul, B., Altimari, L., Okano, A., Bodnariuc Fontes, E., Camata, T., & Moraes, A. (2009). The efficiency of pedaling and the muscular recruitment are improved with increase of the cadence in cyclists and non-cyclists. *Electromyography and Clinical Neurophysiology, 49*(6-7), 311–319.

Dempsey, J. A., Neder, J. A., Phillips, D. B., & O'Donnell, D. E. (2022). The physiology and pathophysiology of exercise hyperpnea. In R. Chan & P. G. Guyenet (Eds.), *Handbook of clinical neurology* (Vol. 188, pp. 201–232). Elsevier. https://doi.org/10.1016/B978-0-323-91534-2.00001-1

Doust, J. H., & Patrick, J. M. (1981). The limitation of exercise ventilation during speech. *Respiration Physiology, 46*(2), 137–147. https://doi.org/10.1016/0034-5687(81)90096-7

Esteve-Gibert, N., & Prieto, P. (2013). Prosodic structure shapes the temporal realization of intonation and manual gesture movements. *Journal of Speech, Language, and Hearing Research, 56*(3), 850–864. https://doi.org/10.1044/1092-4388(2012/12-0049)

Falk, B., Kasnadi, J., & Moss, C. F. (2015). Tight coordination of aerial flight maneuvers and sonar call production in insectivorous bats. *Journal of Experimental Biology, 218*(22), 3678–3688. https://doi.org/10.1242/jeb.122283

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 521–602). Wiley. https://doi.org/10.1002/9781444317251.ch15

Foss, Ø., & Hallén, J. (2005). Cadence and performance in elite cyclists. *European Journal of Applied Physiology, 93*(4), 453–462. https://doi.org/10.1007/s00421-004-1226-y

Fuchs, S., Reichel, U., & Rochet-Capellan, A. (2015). Changes in speech and breathing rate while speaking and biking. *Proceedings of 18th International Congress of Phonetic Sciences,* 1005–1009.

Gademan, M. G. J., Uberoi, A., Le, V.-V., Mandic, S., van Oort, E. R., Myers, J., & Froelicher, V. F. (2012). The effect of sport on computerized electrocardiogram measurements in college athletes. *European Journal of Preventive Cardiology, 19*(1), 126–138. https://doi.org/10.1177/1741826710392669

Godin, K. W., & Hansen, J. H. L. (2011). Analysis of the effects of physical task stress on the speech signal. *The Journal of the Acoustical Society of America, 130*(6), 3992–3998. https://doi.org/10.1121/1.3647301

Godin, K. W., & Hansen, J. H. L. (2015). Physical task stress and speaker variability in voice quality. *EURASIP Journal on Audio, Speech, and Music Processing, 2015*(1). https://doi.org/10.1186/s13636-015-0072-7

Hansen, E. A., Raastad, T., & Hallén, J. (2007). Strength training reduces freely chosen pedal rate during submaximal cycling. *European Journal of Applied Physiology, 101*(4), 419–426. https://doi.org/10.1007/s00421-007-0515-7

Hoit, J. D., Lansing, R. W., & Perona, K. E. (2007). Speaking-related dyspnea in healthy adults. *Journal of Speech,* Language, and Hearing Research, 50*(2), 361–374. https://doi.org/10.1044/1092-4388(2007/026)

Horiuchi, M., Kirihara, Y., Fukuoka, Y., & Pontzer, H. (2019). Sex differences in respiratory and circulatory cost during hypoxic walking: Potential impact on oxygen saturation. *Scientific Reports, 9*(1), Article 9550. https://doi.org/10.1038/s41598-019-44844-6

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 454*(1971), 903–995. https://doi.org/10.1098/rspa.1998.0193

Kelso, J. A. S., Tuller, B., & Harris, K. S. (1983). A "dynamic pattern" perspective on the control and coordination of movement. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 137–173). Springer. https://doi.org/10.1007/978-1-4613-8202-7_7

Kemper, S., Herman, R. E., & Lian, C. H. T. (2003). The costs of doing two things at once for young and older adults: Talking while walking, finger tapping, and ignoring speech of noise. *Psychology and Aging, 18*(2), 181–192. https://doi.org/10.1037/0882-7974.18.2.181

Koblitz, J. C., Stilz, P., & Schnitzler, H.-U. (2010). Source levels of echolocation signals vary in correlation with wingbeat cycle in landing big brown bats (*Eptesicus fuscus*). *Journal of Experimental Biology, 213*(19), 3263–3268. https://doi.org/10.1242/jeb.045450

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lach, J., Wiecha, S., Śliż, D., Price, S., Zaborski, M., Cieśliński, I., Postuła, M., Knechtle, B., & Mamcarz, A. (2021). HR max prediction based on age, body composition, fitness level, testing modality and sex in physically active population. *Frontiers in Physiology, 12*. https://doi.org/10.3389/fphys.2021.695950

Lamberg, E. M., & Muratori, L. M. (2012). Cell phones change the way we walk. *Gait & Posture, 35*(4), 688–690. https://doi.org/10.1016/j.gaitpost.2011.12.005

Lee, J., & Park, K. (2021). Modeling cycling performance: Effects of saddle position and cadence on cycle pedaling efficiency. *Science Progress, 104*(4). https://doi.org/10.1177/00368504211041495

LoMauro, A., & Aliverti, A. (2018). Sex differences in respiratory function. *Breathe, 14*(2), 131–140. https://doi.org/10.1183/20734735.000318

Lucía, A., Hoyos, J., & Chicharro, J. L. (2001). Preferred pedalling cadence in professional cycling. *Medicine & Science in Sports & Exercise, 33*(8), 1361–1366. https://doi.org/10.1097/00005768-200108000-00018

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), Article 3139. https://doi.org/10.21105/joss.03139

MacIntyre, A. D., Cai, C. Q., & Scott, S. K. (2022). Pushing the envelope: Evaluating speech rhythm with different envelope extraction techniques. *The Journal of the Acoustical Society of America, 151*(3), 2002–2026. https://doi.org/10.1121/10.0009844

McDermott, W. J., Van Emmerik, R. E. A., & Hamill, J. (2003). Running training and adaptive strategies of locomotor-respiratory coordination. *European Journal of Applied Physiology, 89*(5), 435–444. https://doi.org/10.1007/s00421-003-0831-5

Mitchell, H. L., Hoit, J. D., & Watson, P. J. (1996). Cognitive-linguistic demands and speech breathing. *Journal of Speech and Hearing Research, 39*(1), 93–104. https://doi.org/10.1044/jshr.3901.93

Nolan, F., & Jeon, H.-S. (2014). Speech rhythm: A metaphor? *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658), Article 20130396. https://doi.org/10.1098/rstb.2013.0396

O'Halloran, J., Hamill, J., McDermott, W. J., Remelius, J. G., & Van Emmerik, R. E. A. (2012). Locomotor-respiratory coupling patterns and oxygen consumption during walking above and below preferred stride frequency. *European Journal of Applied Physiology, 112*(3), 929–940. https://doi.org/10.1007/s00421-011-2040-y

Parrell, B., Goldstein, L., Lee, S., & Byrd, D. (2014). Spatiotemporal coupling between speech and manual motor actions. *Journal of Phonetics, 42*, 1–11. https://doi.org/10.1016/j.wocn.2013.11.002

Patil, S., Sangwan, A., & Hansen, J. H. L. (2010). Speech under physical stress: A production-based framework. *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 5146–5149. https://doi.org/10.1109/ICASSP.2010.5495016

Perségol, L., Jordan, M., & Viala, D. (1991). Evidence for the entrainment of breathing by locomotor pattern in human. *Journal de Physiologie, 85*(1), 38–43.

Plummer-D'Amato, P., Altmann, L. J. P., & Reilly, K. (2011). Dual-task effects of spontaneous speech and executive function on gait in aging: Exaggerated effects in slow walkers. *Gait & Posture, 33*(2), 233–237. https://doi.org/10.1016/j.gaitpost.2010.11.011

Pouw, W., & Fuchs, S. (2022). Origins of vocal-entangled gesture. *Neuroscience & Biobehavioral Reviews, 141*. https://doi.org/10.1016/j.neubiorev.2022.104836, 104836

Pouw, W., Jonge-Hoekstra, L., Harrison, S. J., Paxton, A., & Dixon, J. A. (2021). Gesture–speech physics in fluent speech and rhythmic upper limb movements. *Annals of the New York Academy of Sciences, 1491*(1), 89–105. https://doi.org/10.1111/nyas.14532

Pouw, W., Paxton, A., Harrison, S. J., & Dixon, J. A. (2020). Acoustic information about upper limb movement in voicing. *Proceedings of the National Academy of Sciences, 117*(21), 11364–11367. https://doi.org/10.1073/pnas.2004163117

Pouw, W., & Trujillo, J. (2021). *Extracting a smoothed amplitude envelope from audio*. https://wimpouw.github.io/EnvisionBootcamp2021/extract_AmplitudeEnvelope_Rmarkdown.html

Raffegeau, T. E., Haddad, J. M., Huber, J. E., & Rietdyk, S. (2018). Walking while talking: Young adults flexibly allocate resources between speech and gait. *Gait & Posture, 64*, 59–62. https://doi.org/10.1016/j.gaitpost.2018.05.029

R Core Team (2022). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. https://www.R-project.org/

Rochet-Capellan, A., Laboissière, R., Galván, A., & Schwartz, J.-L. (2008). The speech focus position effect on jaw–finger coordination in a pointing task. *Journal of Speech, Language, and Hearing Research, 51*(6), 1507–1521. https://doi.org/10.1044/1092-4388(2008/07-0173)

Romero, S. A., Minson, C. T., & Halliwill, J. R. (2017). The cardiovascular system after exercise. *Journal of Applied Physiology, 122*(4), 925–932. https://doi.org/10.1152/japplphysiol.00802.2016

Rotstein, A., Meckel, Y., & Inbar, O. (2004). Perceived speech difficulty during exercise and its relation to exercise intensity and physiological responses. *European Journal of Applied Physiology, 92*(4-5), 431–436. https://doi.org/10.1007/s00421-004-1160-z

Schaal, S., Sternad, D., Osu, R., & Kawato, M. (2004). Rhythmic arm movement is not discrete. *Nature Neuroscience, 7*(10), 1136–1143. https://doi.org/10.1038/nn1322

Serré, H., Dohen, M., Fuchs, S., Gerber, S., & Rochet-Capellan, A. (2021). Speech breathing: Variable but individual over time and according to limb movements. *Annals of the New York Academy of Sciences, 1505*(1), 142–155. https://doi.org/10.1111/nyas.14672

Shattuck-Hufnagel, S., & Ren, A. (2018). The prosodic characteristics of non-referential co-speech gestures in a sample of academic-lecture-style speech. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.01514

Sheel, A. W., Richards, J. C., Foster, G. E., & Guenette, J. A. (2004). Sex differences in respiratory exercise physiology. *Sports Medicine, 34*(9), 567–579. https://doi.org/10.2165/00007256-200434090-00002

Stehwien, S., & Meyer, L. (2022). Short-term periodicity of prosodic phrasing: Corpus-based evidence. *Proceedings of Speech Prosody, 2022,* 693–698. https://doi.org/10.21437/SpeechProsody.2022-141

Stidsholt, L., Johnson, M., Goerlitz, H. R., & Madsen, P. T. (2021). Wild bats briefly decouple sound production from wingbeats to increase sensory flow during prey captures. *iScience, 24*(8). https://doi.org/10.1016/j.isci.2021.102896

Tanaka, H., Monahan, K. D., & Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology, 37*(1), 153–156. https://doi.org/10.1016/S0735-1097(00)01054-8

Tilsen, S., & Arvaniti, A. (2013). Speech rhythm analysis with decomposition of the amplitude envelope: Characterizing rhythmic patterns within and across languages. *The Journal of the Acoustical Society of America, 134*(1), 628–639. https://doi.org/10.1121/1.4807565

Travers, G., Kippelen, P., Trangmar, S. J., & González-Alonso, J. (2022). Physiological function during exercise and environmental stress in humans—An integrative view of body systems and homeostasis. *Cell, 11*(3), Article 383. https://doi.org/10.3390/cells11030383

Trouvain, J., & Truong, K. P. (2015). Prosodic characteristics of read speech before and after treadmill running. *Interspeech 2015*, 3700–3704. https://doi.org/10.21437/Interspeech.2015-734

Uusitalo, A. L. T., Uusitalo, A. J., & Rusko, H. K. (1998). Exhaustive endurance training for 6–9 weeks did not induce changes in intrinsic heart rate and cardiac autonomic modulation in female athletes. *International Journal of Sports Medicine, 19*(08), 532–540. https://doi.org/10.1055/s-2007-971956

Varlet, M., Williams, R., & Keller, P. E. (2020). Effects of pitch and tempo of auditory rhythms on spontaneous movement entrainment and stabilisation. *Psychological Research, 84*(3), 568–584. https://doi.org/10.1007/s00426-018-1074-8

Volín, J. (2023). Variation in speech tempo and its relationship to prosodic boundary occurrence in two speech genres. *AUC Philologica, 2022*(1), 65–81. https://doi.org/10.14712/24646830.2022.29

Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication, 57,* 209–232. https://doi.org/10.1016/j.specom.2013.09.008

Weston, H., Koenig, L. L., & Fuchs, S. (2021). Changes in glottal source parameter values with light to moderate physical load. *Interspeech 2021*, 3350–3354. https://doi.org/10.21437/Interspeech.2021-1881

Zelic, G., Kim, J., & Davis, C. (2015). Articulatory constraints on spontaneous entrainment between speech and manual gesture. *Human Movement Science, 42,* 232–245. https://doi.org/10.1016/j.humov.2015.05.009

Ziegler, A., VanSwearingen, J., Jakicic, J. M., & Verdolini Abbott, K. (2019). Phonation demonstrates goal dependence under unique vocal intensity and aerobic workload conditions. *Journal of Speech, Language, and Hearing Research, 62*(8), 2584–2600. https://doi.org/10.1044/2019_JSLHR-S-18-0126

Appendix (p. 1 of 2)

Speech Materials

Read passage

Um ein perfektes Fest zu organisieren, braucht man eigentlich gar nicht so viel. Vor allem braucht man eigentlich gar nicht so viel Zeit, man kann das eigentlich auch relativ spontan machen, so lange Läden noch offen haben. Wichtig ist einfach, dass man gute Laune hat, dass auf dem Fest gute Stimmung herrscht, deshalb ist die Gästeliste etwas sehr Wichtiges. Die Leute, die man einlädt, sollten keinen Streit untereinander haben, sie sollten sich alle untereinander ganz gut verstehen, vielleicht auch teilweise schon kennen, jeder sollte irgendeine Person haben, die er als Ansprechpartner haben kann und sich nicht einsam fühlt. Und außerdem sollte natürlich für das leibliche Wohl gesorgt sein, also sollte man sich vorher Gedanken machen was man zu Essen anbietet, das vorher einkaufen und zubereiten.

"To organize a perfect party you actually don't need that much. And you actually don't even need that much time, you can actually do it pretty spontaneously as long as the shops are still open. The important thing is just that people are in a good mood, that there's a good vibe at the party, so the guest list is something very important. The people you invite shouldn't have any problems with each other, they should all get along well, maybe some already know each other, everyone should have a person they can talk to and not feel alone. And of course food and drinks should be provided, so you should think about what to eat beforehand, buy it before the party and prepare it."

Note: The English translation uses "you" for the impersonal pronoun *man* "one" to better capture the register of the original.

Monologue prompts

Questions to prompt unscripted monologues with mean ratings for ease of discussion and emotional connotation using a 5-point Likert scale (1 = *not at all*; 3 = *it depends*; 5 = *extremely*). The reading prompt is shown in gray.

	Mean rating	
Questions	**Ease**	**Emotional**
What kind of vacations do you like? What about your friends/family?	4.0	3.3
How has your daily routine changed since the pandemic?	3.7	3.0
Which mobile phone app do you use the most and why?	4.0	2.5
What is your favorite season and why?	4.3	3.0
Do you think technology makes life easier or more complicated?	3.7	2.6
How do you feel about virtual lectures/meetings?	3.7	2.6
What do you do to relax after a completing a difficult task?	3.5	3.0
Describe what you like about Berlin. What's missing in Berlin?	4.3	3.5
What is your idea of a nice weekend?	4.2	3.5
What does one need to organize the perfect party?	3.2	2.7

Appendix (p. 2 of 2)

Speech Materials

Model coefficient outputs

	Estimate	*SE*	*df*	*t* value	Pr(>|*t*|)
			Cadence		
(Intercept)	0.902	0.022	45.705	40.783	**< .001**
workload[T.mod.]	0.064	0.015	46.088	4.146	**< .001**
task[T.read speech]	−0.007	0.017	48.435	−0.438	.663
task[T.spont. speech]	−0.015	0.017	49.247	−0.862	.393
trial[T.2]	0.000	0.005	446.814	0.070	.944
trial[T.3]	0.013	0.005	447.224	2.482	.013
			Envelope		
(Intercept)	4.754	0.039	46.999	121.107	**< .001**
workload[T.low]	−0.197	0.034	41.661	−5.798	**< .001**
workload [T.mod.]	−0.533	0.044	44.529	−12.233	**< .001**
task[T.spont.speech]	−0.828	0.066	46.971	−12.627	**< .001**
workloadlow:taskspontaneous	0.197	0.056	46.090	3.499	**.001**
workloadmoderate:taskspontaneous	0.223	0.062	44.110	3.575	**< .001**
			IMF1		
(Intercept)	5.406	0.047	48.109	113.826	**< .001**
workload [T.low]	−0.293	0.041	46.837	−7.220	**< .001**
workload [T.mod.]	−0.755	0.055	46.768	−13.819	**< .001**
task[T.spont. speech]	−0.467	0.060	45.680	−7.789	**< .001**
trial[T.2]	0.001	0.028	206.861	0.052	.958
trial[T.3]	0.036	0.030	76.821	1.232	.222
			IMF2		
(Intercept)	2.697	0.021	56.801	125.485	**< .001**
workload [T.low]	−0.119	0.023	40.631	−5.170	**< .001**
workload [T.mod.]	−0.326	0.027	44.555	−12.130	**< .001**
task[T.spont. speech]	−0.457	0.038	46.582	−12.069	**< .001**
trial[T.2]	0.001	0.011	591.875	0.115	.908
trial[T.3]	0.036	0.011	592.532	3.231	**.001**
workloadlow:taskspontaneous	0.074	0.037	47.031	1.996	.052
workloadmoderate:taskspontaneous	0.077	0.041	45.417	1.870	.068

Note. Bolded Pr(>|*t*|) values indicate statistically significant results. *SE* = standard error.