PREPRINT

Title: The human voice aligns with whole-body kinetics

Authors: Wim Pouw*[†], Raphael Werner[†], Lara Burchardt[† ‡], Luc Selen[†]

*Correspondence should be addressed to wim.pouw@donders.ru.nl.

Affiliations:

† Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Netherlands

‡ Leibniz Institute for General Linguistics ZAS Berlin

Abstract

Humans use their voice concurrently with upper limb movements, known as hand gestures. Recently it has been shown that fluctuations in intensity and the tone of the human voice synchronizes with upper limb movement (including gesticulation). In this research direct evidence is provided that the voice changes with arm movements because it interacts with whole-body muscle activity (measured through surface EMG and postural measurements). We show that certain muscles (e.g., pectoralis major) that are associated with posture and upper limb movement are especially likely to interact with the voice. Adding wrist weights to increase the mass of the moving upper limb segment led to increased coupling between movement and voice. These results show that the voice co-patterns with whole-body kinetics relating to force, rather than kinematics, invoking several implications how the voice is biomechanically modeled, how it should be simulated, and importantly how the human voice must have evolved in relation to the whole-body motor system. We concluded that the human voice is animated by the kinetics of the whole body.

Keywords: Vocalization, Hand Gesture, EMG, posture, Biomechanics

Main Text

In textbooks about the mechanics of respiration, vocalization, and speaking, students learn that the intercostal muscles and diaphragm are the "primary" respiratory muscles that control subtle subglottal pressures needed to phonate at normal levels (1). Yet the "secondary" muscles that only occasionally activate during exceptional conditions of breathing and/or speaking are actually very often recruited during the upper limb movements that humans produce during speaking (2–4). During pulsing hand gestures, muscles activate around the trunk (3, 5), which are known to support forced expiration during more extreme respiratory situations such as coughing (6). Furthermore, muscles around the shoulder girdle are active during upper-limb gesturing that accompanies speech. While these muscles are not activated for speech alone or tidal breathing (1), they are known to be used for respiration in persons with respiratory problems such as COPD (7, 8). Given that these muscles *are* active during arm movements that accompany speech (i.e., gestures) in typical able-bodied humans (2, 3, 9), they might actually regularly enter into functional synergies with respiration for speaking (similar to people with COPD). The contradiction is that when we speak "normally" we tend to activate muscles for whole-body expression that are usually described as more "abnormal" muscle activity recruited for breathing. This contradiction highlights that our best scientific understanding of the respiratory foundations of speaking (10) requires a reconsideration of functional units during *utterances in action*.

So what movements do humans make when speaking? Hand gestures are the most common type of action produced during speaking. Gestures often have a pulsing 'beat-like' quality with sudden stops and accelerations that are temporally integrated with vocal aspects of speech (11). These beating moments of gesture are found to be strategically timed with excursions in voice intensity and fundamental frequency (F0). These acoustic features, among others, contribute to the perception of so-called accented or stressed moments in speech (11), i.e., these contribute to the prosody of speech. This synchronization of pulsing gestural movements with prosodically-relevant vocal inflections (e.g., https://osf.io/29h8z/) is referred to as multimodal prosody (11) and has been of interest to linguistics (12), but generally ignored in kinesiological research.

Recently gesture-speech synchronization has been understood as rooted in biomechanics (13). The available evidence for the biomechanics thesis is from kinematic research and not from

research monitoring muscle activity directly. Namely, more extreme peaks in acceleration of movements with larger (arm) vs smaller (wrist) upper limb segments relate to more chest-circumference changes, which is found to relate to extreme acoustic effects on the intensity of vocal sound (14). Furthermore, acoustic effects of upper limb movements are more extreme when subjects are in a more unstable standing versus sitting position (15). These previous studies assessed continuous voicing (15, 16), mono-syllable utterances (14), singing (17), and fluent speech production (18). The combined evidence supports the gesture-speech biomechanics thesis that a physical impulse *J* (mass x acceleration integrated over some time window) impacts posture (especially when standing) and recruits muscles around the chest, which interact with or recruit respiratory-related muscles that affect rib cage movements, as such modulating respiratory-vocal functioning (such that intensity and secondarily F0 are affected).

These effects of upper limb kinematics on vocal utterances are hypothesized to lie in biomechanics, but no study has been performed on this level of analysis. That biomechanics is the level of interaction lies in the hypothesis that accelerating upper limb segment recruits focal and/or peripheral muscle activations that interact with respiratory-vocal control (2, 3, 19). For example, a chest muscle (pectoralis major) that drives internal rotation of the upper arm (humerus) and stabilizes other arm movements is associated with forced respiratory action (1, 20). Furthermore, upper limb actions are often coordinated with a whole suit of muscle units around the trunk and shoulder girdle, acting as a pre-stressed 'tensegrity' system (21), that react and anticipate the destabilizing effects of movement on body posture (22, 23). Previous research indeed indirectly indicates possible interactions: the erector spinae is implicated as a postural muscle during upper limb movement (9), and also attaches to the rib cage and implicated in "secondary" respiratory control (1). In non-human animals, trunk muscle integration into respiratory action has been well-documented. For example, in flying echo-locating bats (*Pteronotus parnellii*; (24)) muscles activated for flying (pectoralis) are at the same time mechanically driving expiration for echo-vocalizing. Though during passive vocal echo-locating muscle units around the abdomen (rectus abdominis) will drive expiration. This biomechanics research suggests that different muscle synergies are responsible for the expiratory vocal flow of vocalization during flight versus passive activities.

We expect that the upper limbs in humans are similarly integrated with vocalizations. However, no direct evidence has been provided so far to confirm biomechanical factors play a

role in human pectoral limb movement and vocalization, ie., whether a particular type of muscle activity, postural destabilization, and mass (rather than size) of the body segment set in motion, predicts inflections in vocalization. Therefore we must move beyond kinematics and assess signatures of body *kinetics*, the posture-perturbing muscle-generated physical impulses that distribute over the 'tensegritous' musculo-skeletal system (25), affecting respiratory-vocal functioning such making the voice momentarily louder (a key prosodic feature of the voice). With the following study, measuring whole-body muscle activity via surface ElectroMyoGraphy, postural perturbations via tracking of ground-reaction forces, and voice modulations via acoustic analyses, we were able to show that the voice aligns with whole-body kinetics.

Methods

Informed by a simulation-based power analyses (26) reported in the pre-registration (https://osf.io/jhdq4) we recruited $N = 17$ participants: 7 female, 10 male, M (SD) age = 28.50 (6.50), M (SD) body weight = 72.10 kg (10.20). Note that our digital Supplemental Information with reproducible code is provided on our Github webpage.

In this study (see supporting methods for details) we measured the voice acoustics of participants vocalizing a sustained /a/ vowel (sound examples). Upon 3 seconds of vocalization participants were prompted to make a single movement while keeping their vocalization as steady as possible for another 4 seconds. Participants made different upper limb movements while we measured muscle activity (surface EMG) and concurrent postural perturbation (vertical ground-reaction measurements). See supporting methods for full procedural details.

The study design consisted of a 5-level within subjects Movement condition: no movement as well as 4 actions with the elbow flexed at a 90 degree angle: flexion and extension, internal rotation and external rotation (see Figure 1). Additionally, there was a two-level within subject Weight condition, where participants performed movements with and without a 1kg wrist weight. Finally, there was a 2-level within-subject Vocal condition where participants either expired versus vocalized during the trial. To reduce the number of comparisons, and as stated in our pre-registration, we will only report results for the vocalization condition in this report. Thus we are left with a 5 (Movement condition) x 2 (Weight condition) fully within-subject design. We were able to collect 1425 trials in total (see supporting methods), entailing 2.7 hours of continuous voice data.

Analysis procedure. Our analysis is aimed to relate vocal inflections with postural and muscle activity due to upper limb movement. We therefore identified positive and negative peaks relative to a trend line (see below) in the smoothed amplitude envelope of the vocalization before the peak in wrist speed was reached (see Supplemental Methods for further detail). Thereby we assess peaks in vocalization related to the initiation of the movement rather than the physical impulse produced during deceleration (which will recruit antagonistic muscle units as compared to the initiation phase). Note that there is a general decline in the amplitude envelope of the vocalization during a trial (see Figure 1) as the subglottal pressure falls when the lungs deflate. To quantify deviations from stable vocalizations, we therefore detrend the amplitude envelope time series, so as to assess positive or negative peaks *relative to this trend line*. For the envelope, muscle activity, and the change in center of pressure we will measure the global maxima happening within the analyses window (i.e., within a trial we take a local maximum occuring between movement onset and offset). We will analyze positive and negative peaks within the movement window separately. To ensure that we identify multimodal signal peaks that are associated with onset of a movement we determine the peak speed of a movement and look for multimodal signal peaks (vocal, EMG, ground reactions) before the peak in speed. We replicate the analysis procedure for the no movement condition as a control comparison.

Manipulation checks.

Figure 1 provides an overview of the experiment, an example of time series data collected, and a summary of the muscle activity and postural change patterns obtained for the different movement conditions.

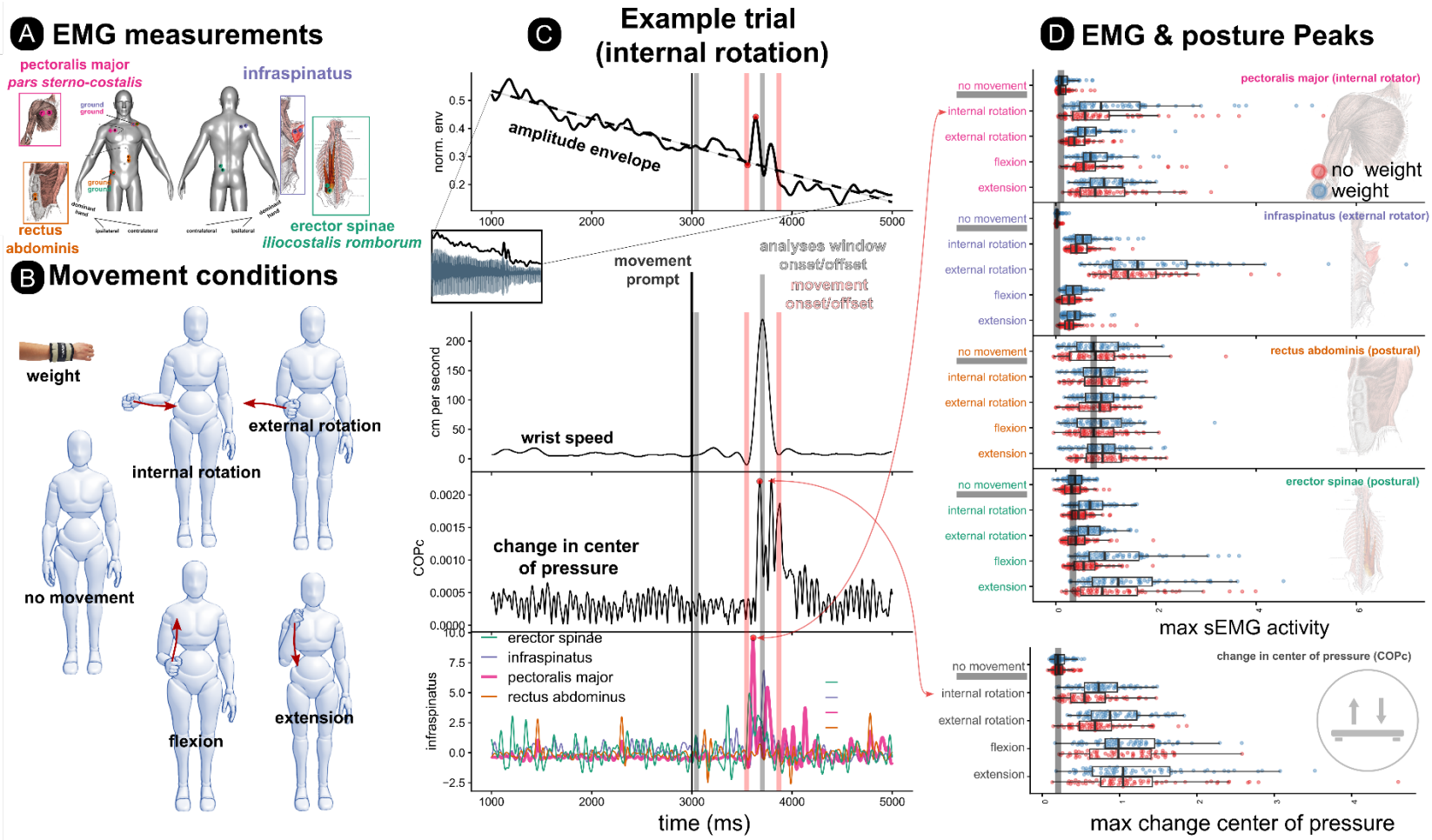

Figure 1. Experiment methods and manipulation checks A) During the experiment we administered surface ElectroMyoGraphy (sEMG) to track muscle activity at four locations: pectoralis major [internal rotator], infraspinatus [external rotator], erector spinae [postural], rectus abdominis [postural]. B) Participants performed vocalizations while wearing a wrist-weight or no weight, and performing several movement conditions. C) For each vocalization trial (example of an internal rotation trial shown here), lasting 7 seconds with a movement prompt after 3 seconds (here 4 seconds is shown), a smoothed amplitude envelope was extracted (z-normalized units, normalized per participant), and a linear trend line was determined so as to express negative and positive vocal peaks relative to this trendline. The vocal acoustics, muscle activity. and postural peaks were determined between 500 milliseconds before movement onset and the moment of peak speed (this ensures to only consider effects of movement onset, rather than the moment of deceleration). It is clear in C) that movement coincides with peaks in muscle (especially pectoralis major), postural, vocal acoustics. D) provides the muscle and postural activity (z-normalized units, normalized by participant).

Results

As compared to vocalizing while not moving, we find that all movement conditions (internal rotation, external rotation, flexion, extension) as compared to the no movement condition lead to more extreme positive peaks in the amplitude envelope of the vocalization (p's $< .001$; see Table S3), but also more extreme negative peaks (p's $< .001$; see Table S5). Thus no matter what type of movement the voice is affected leading to perturbations in the intensity of the voice. Furthermore, in the movement conditions the voice is more affected in the weight condition, for negative peaks ($p = .024$; see Table S6), but especially for positive peaks ($p = .01$; see Table S4). Thus by increasing the mass of the upper limb with 1kg, the movement of said limb affects vocalization more as compared to no weight. Please see the supporting results for a full report of all the analyses (for a computationally reproducible version with code see here). Also see Figure 2 for an overview of the main results.

We further assessed how different general muscle activity relates to positive and negative peaks in the vocalization acoustics in an analysis where we pool all data (disregarding movement condition). We obtain that positive peaks in vocalization was related to activity in pectoralis major ($b = 0.013$, SE = 0.003, $p < .001$), erector spinae ($b = 0.016$, SE = 0.004, $p < .001$), and infraspinatus ($b = 0.006$, SE = 0.002, $p = .005$), but not in rectus abdominis ($p = .711$), as shown in Table S7. In general, we obtained more noisy measurements in the rectus abdominis activity, also having to do with there being more adipose tissue. Only two muscles were associated with negative peaks in vocalization and the relations were weaker than the comparison with positive peaks: pectoralis major ($b = 0.003$, SE = 0.002, $p = .032$) and infraspinatus ($b = 0.003$, SE = 0.001, $p = .02$) related to negative peaks in vocalization, though rectus abdominis ($p = .719$), and erector spinae ($p = .53$) did not, as is visible in Table S8. Finally, we were able to relate the peak in the absolute change in the center of pressure with muscle activity. The erector spinae ($b = 0.118$, SE = 0.012, $p < .001$), pectoralis major ($b = 0.063$, SE = 0.009, $p < .001$), and infraspinatus ($b = 0.028$, SE = 0.007, $p < .001$), as shown in Table S9. The rectus abdominis did not show this relationship in a pooled analysis ($b = 0.025$, SE = 0.023, $p = .268$).

For the previous analyses we lump all movement conditions together to assess relationships between muscle activity, posture and vocalization. However, it is clear that further exploratory analyses yield important context-dependent variability. For example, Pearson correlation matrices for all the different peaks per signal obtained variable relationships per

179 movement condition, for example weakly implicating rectus abdominis in postural control in the
180 extension condition (Pearson $r = .18$, $p = .042$; see Table S12). We obtain very strong
181 associations in the internal rotation condition for the degree to which the body changes center of
182 pressure and the magnitude of the positive peak in vocalization (see Figure 2, panel F & G),
183 (Pearson $r = .51$, $p < .001$; see Table S10 or Figure S9). This shows that posture perturbation can
184 be directly related to vocalization perturbation.

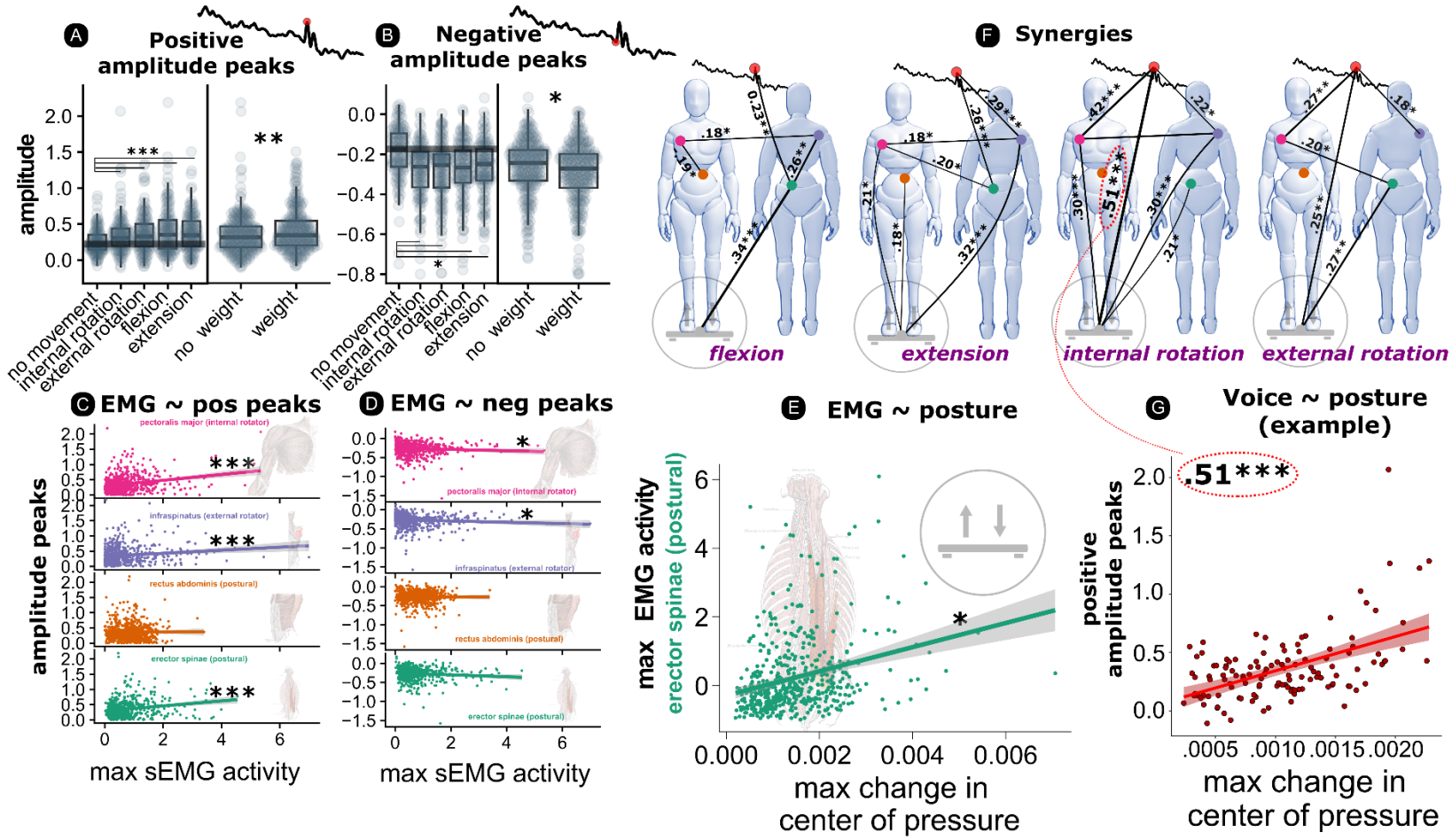

Figure 2. Overview Main results A) shows the magnitude positive peaks (relative to trendline) for the different movement conditions. B) shows this for the negative peaks (note A and B have different y-axis ranges). Plots C and D show the relationship between peak muscle activity and positive (C) and negative (D) vocal amplitude envelope peaks. E) shows the relationship with change in center of pressure and the peak activity in the postural muscle erector spinae. All axes are given in z-normalized units (normalized within participants). F) provides the statistically reliable Pearson correlation coefficients (line thickness also represents correlation strength) between all the measurements, presented separately for movement condition. G) highlights the highly reliable correlation obtained in the internal rotation condition (in F) between change in the center of pressure and the magnitude of the positive peak in vocal amplitude envelope (i.e., posture directly relates to voice). * $p < 0.05$; ** $p < 0.01$; *** $p < .001$

Discussion

The current report demonstrates that the human voice resonates to forces produced by upper limb movements that mechanically propagate through the musculo-skeletal system - in other words, the voice aligns with whole-body kinetics. This research provides the first direct evidence for whole-body kinetics affecting the voice as we obtain that we can statistically predict acoustic fluctuations that are different in movement versus passive conditions by muscle activity patterns, mass of the upper limb segment, and by postural destabilization. Adding 1kg weights to the wrist increases the intensity peaks of the voice that results from movement. Our confirmatory and exploratory analyses show that it is primarily positive peaks in the intensity of the voice that result from upper limb movement, suggesting that different kind of arm movements primarily increase subglottal pressures. There is a special role for posture in the way arm movements affect the voice, as posture-related muscle activity is implicated in voice modulations, and we even find highly reliable direct alignment of degree of center of pressure change and increases in voice intensity. These analyses confirm that the voice interacts with muscle chains of the whole body.

While we have focused on generalized patterns in the data, further research should focus on context-conditioned variability, that we highlight in Figure 2, panel E. Do particular combinations of muscle activity within a single type of movement predict the shape of vocal trajectory? For example, if pectoralis major activity is stabilizing an external rotation that is produced primarily by the infraspinatus, does this then modulate the perturbing effects on vocalization? Further research, of which the current open data are a resource, will provide more detailed context-dependent associations between gesture and vocalization. In a more general outlook of this line of research there is an open challenge to assess how in more natural speech conditions whole-body kinetics is aligned with prosodic targets.

There are important implications to the current results. Firstly, any shortcomings with speech synthesis systems (e.g., models, robots) to reach believably human feature characteristics, need to be aware that said disembodied systems are abstracting out how body posture and whole-body muscle chains can modulate vocal acoustics through interacting with a pre-stressed musculo-skeletal system that allows to distribute local perturbations more globally, affecting the respiratory-vocal system. Furthermore, our understanding of how effective communication may be compromised by pathologies should be weighed against how humans voice operates in action. It is known that people with Parkinson's have trouble with not only expressing emotions (27), but

also with speech prosody (28, 29), overall appearing monotonic in expression. Obviously whole-body postural dynamics are frustrated in Parkinson's, which according to the current research might be part of the spectrum of attenuated vocal features.

It has been shown that people who hear another's vocalization without seeing the vocalizer can glean information about body motions (16, 30). This research indicates that persons are attuning to how the voice can inform about whole-body kinetics, where upper limb physical impulses perturb postural stability, invoking whole-body muscle chains that constrain the respiratory-vocal system. James Gibson (31) in his ecological approach to perception emphasized that the visual system is not limited to the eyes alone but is rather a dynamic system that involves the entire body and its interaction with the environment. Our conclusion is similar when it comes to the voice. The human voice is not contained by the vocal cords. It lies within the tube-like upper vocal tract, lengthened by moving cavities that change resonant properties of the sound. It is powered and modulated by respiration, which is situated in a thoracic muscle complex. We show that at times this thoracic muscle complex becomes part of flexibly assembled muscle synergies that maintain the postural integrity during moments when the body organizes into gesticulation. We should indeed start to wonder whether the expressiveness of gesture lies in part in its embodied connection with the voice. Body movements directly animate the voice.

Supporting information[1]

Supporting methods

This study has been approved by the Ethics Committee Social Sciences (ECSS) of the Radboud University (reference nr.: 22N.002642). This study has been pre-registered before data collection (see preregistration).

Experimental design

This study concerned a two-level wrist-weight manipulation (no weight vs. weight), a two-level within-subject vocalization condition (expire vs.vocalize), and a five-level within-subject movement condition ('no movement', 'extension', 'flexion', 'external rotation', 'internal rotation'). With 4 trial repetitions over the experiment, we yield 80 (2 weights x 2 vocalizations x 5 movements x 4 repetitions) trials per participant. Trials were blocked by weight condition and vocalization condition (so that weight and task conditions did not switch constantly from trial to trial). Within blocks all movement conditions were randomized.

Participants

For the current pre-registered confirmatory experiment, as planned and supported by a power analysis (see preregistration), we collected N = (17) participants: 7 female, 10 male, *M* (*SD*) age = 28.50 (6.50), *M* (*SD*) body weight = 72.10kg (10.20), *M* (*SD*) body height = 175.10 cm (8.50), *M*(*SD*) BMI = 23.40 (2.20), *M* (*SD*) triceps skinfold = 19.10mm (4.30).

Exclusions and deviations from pre-registration

We also performed the experiment with one other participant, but due to an issue with LSL streaming, this dataset could not be synchronized and was lost. Furthermore, due to running over time one participant had to terminate the study earlier about half way through. Note further, that in our pre-registration we wanted to admit participants with a BMI lower than 25, but since participants were difficult to recruit we accepted three participants with slightly higher BMIs too (max BMI of the current dataset: 27.10). Participants were all able-bodied and did not have any constraints in performing the task. Finally, as stated in the pre-registration we will only report results on the vocalization trials (and not the expiration-only trials). This means that for this

[1] Further information on the supporting methods and results and a computationally reproducible manuscript can be accessed via https://wimpouw.github.io/kineticsvoice/.

report in total we have N = 17 participants, with 636 analyzable trials, with balanced conditions: weight = 50.20%, no movement = 19.70%, internal rotation = 20.00 %, external rotation = 20.30%, flexion = 20.10 %, extension = 20.00 %.

Measurements and equipment

Body measurements. To enable future analyses of possible modulating individual-specific body properties we collect some basic information about body properties. Namely, weight, under arm length, upper arm length, triceps skinfold, and upper arm circumference.

Experiment protocol. The experiment was coded in Python using functions from PsychoPy. The experiment was controlled via a Brainvision Button Box (Brain Products GmbH, Munich, Germany), which was also streaming its output to the data collection PC unit.

Wrist weights. To manipulate the mass set in motion, we apply a wrist weight. We use a TurnTuri sports wrist weight of 1 kg.

Video and kinematics. The participants are recorded via a video camera (Logitech StreamCam), sampling at 60 frames per second. We used Mediapipe (32) to track the skeleton and facial movements, which is implemented in Masked-piper which we also use for masking the videos (33). The motion-tracked skeleton, specifically the wrist of the dominant hand, is used to estimate movement initiation, peak speed, and the end of the movement. The motion tracking is however only used for determining movement windows, and is not of central concern.

Muscle activity (surface ElectroMyography: EMG). We measured surface EMG using a wired BrainAmp ExG system (Brain Products GmbH, Munich, Germany). Disposable surface electrodes (Kendall 24mm Arbo H124SG) were used, and for each of the four muscle targets we had 3 (active, reference, ground) electrodes (12 electrodes total). The sEMG system sampled at 2500 Hz (for post-processing filters see below).

For an overview of the electrode attachments see Figure S1. We prepare the skin surface for EMG application with a scrub gel (NuPrep) followed by cotton ball swipe with alcohol (Podior 70%). Active and reference electrodes were attached with a 15mm distance center to center.

We attached electrodes for focal muscles which directly participate in the internal (pectoralis major) and external rotation (infraspinatus) of the humerus. Electrodes were applied for focal muscles ipsilaterally (relative to the dominant hand). We attached electrodes to the

muscle belly of the clavicular head of the pectoralis major, with a ground electrode on the clavicle on the opposite side.

We also attached electrodes for postural muscles which will likely anticipate and react to postural perturbations due to upper limb movements. Since these muscles should act in the opposite direction of the postural perturbation of the dominant hand, we applied electrodes contralaterally to the dominant hand. We attach electrodes to the rectus abdominis, with a ground electrode on the iliac crest on the opposite side. We also attached electrodes to the erector spinae muscle group (specifically, the iliocostalis lumborum).

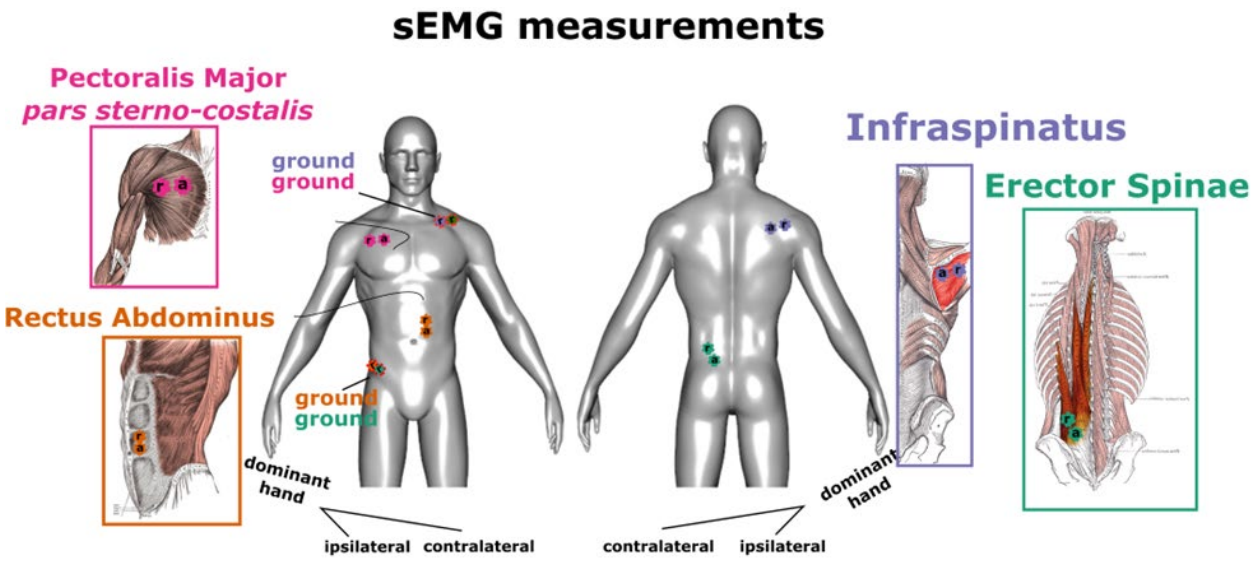

Figure S1. Overview sEMG target muscles. Active (a) and reference (r), and ground (g) sEMG electrodes, for each muscle target.

Ground reaction measurements. We used an inhouse-built 1m² balance board with vertical pressure sensors. The sensors were derived and remodified from the Wii-Balance board sensors. The sampling rate was 400 Hz. The system was time-locked within millisecond accuracy, and has a spatial accuracy of several sub-millimeters. A National Instruments card, USB-62221 performed the A/D conversion and was connected via USB to the PC.

Acoustics. To ensure proper acoustic intensity measurements we used a headset microphone; MicroMic C520 (AKG, Inc.) headset condenser cardioid microphone sampling at 16 kHz. The gain levels of the condenser power source were set by the hardware (and could not be changed).

Recording setup and synchronization. We use LabStreamLayer that provides a uniform interface for streaming different signals along a network, where a common time stamp for each signal ensures sub-millisecond synchronization. We used a Linux system to record and stream the microphone recordings. Additionally a second PC collected video, and streamed ground reaction forces, and EMG. A data collection PC collected the audio, ground reaction force, and EMG streams and stored the output in XDF format for efficient storing of multiple time-varying signals. The video data was synchronized by streaming the frame number to the LSL recorder, allowing us to match up individual frames with the other signals (even when a frame is dropped).

Procedure

Participants are admitted to the study based on exclusion criteria and sign an informed consent. We ask participants to take off their shoes and we proceed with the body measurements, while instructing the participant about the nature of the study. After body measurements, we apply the surface EMG. We prepared the muscle site with rubbing gel and alcohol, and active/reference electrodes were placed with a distance of 15mm from each other's center. See Figure S1 for the sEMG electrode locations. The procedures up to the start of the experiment take about 20 minutes or less in total.

Upon the start of the experiment participants take a standing position on the force platform. The experiment commences with calibration and practice trials. First, 10 seconds of silent breathing with body movements are recorded. Then participants are asked to take a maximum inspiration followed by a maximum expiration to measure signal conditions under respiratory boundary conditions. Then, for the practice trials, each movement was practiced with expiring and vocalization while performing the movement conditions, and the participant is introduced to wearing the wrist weight of 1 kg. After practice trials, participants performed 80 blocked trials.

For each (practice) trial participants are closely guided by the information on the monitor. Firstly, participants are shown the movement to be performed for the trial, and have to prompt the experimenter that they are ready to continue. Then participants are instructed to adopt the start position of the movement, which is a 90 degree elbow flexion, with either an externally rotated humerus (start position for internal rotation), or a non-rotated humerus with the wrist in front of the body (rest position for the other movement conditions). For the no movement condition participants are asked to rest their arms alongside their body. Upon trial start,

participants are asked to inhale deeply with a timer counting down from 4 seconds. Then, participants are asked to 'vocalize' or 'expire', with a screen appearing after 3 seconds to perform the movement with visual guidance to where the movement end position is so that participants are reminded of the movement. After an additional 4 seconds the trial ends, which allows more than enough time to perform the movement and stabilize vocalization after the perturbation. In the no movement condition, a prompt is given to maintain one's posture.

Preprocessing of the data streams

EMG. To reduce heart rate artifacts we apply a common method (34) of high-pass filtering the signal at 30 Hz using a zero-phase 4th order Butterworth filter. We then full-wave rectified the EMG signal and applied a zero-phase low-pass 4th order Butterworth filter at 20 Hz. When filtering any signal we pad the signals to avoid edge effects. We normalized the EMG signals within participants before submitting to analyses.

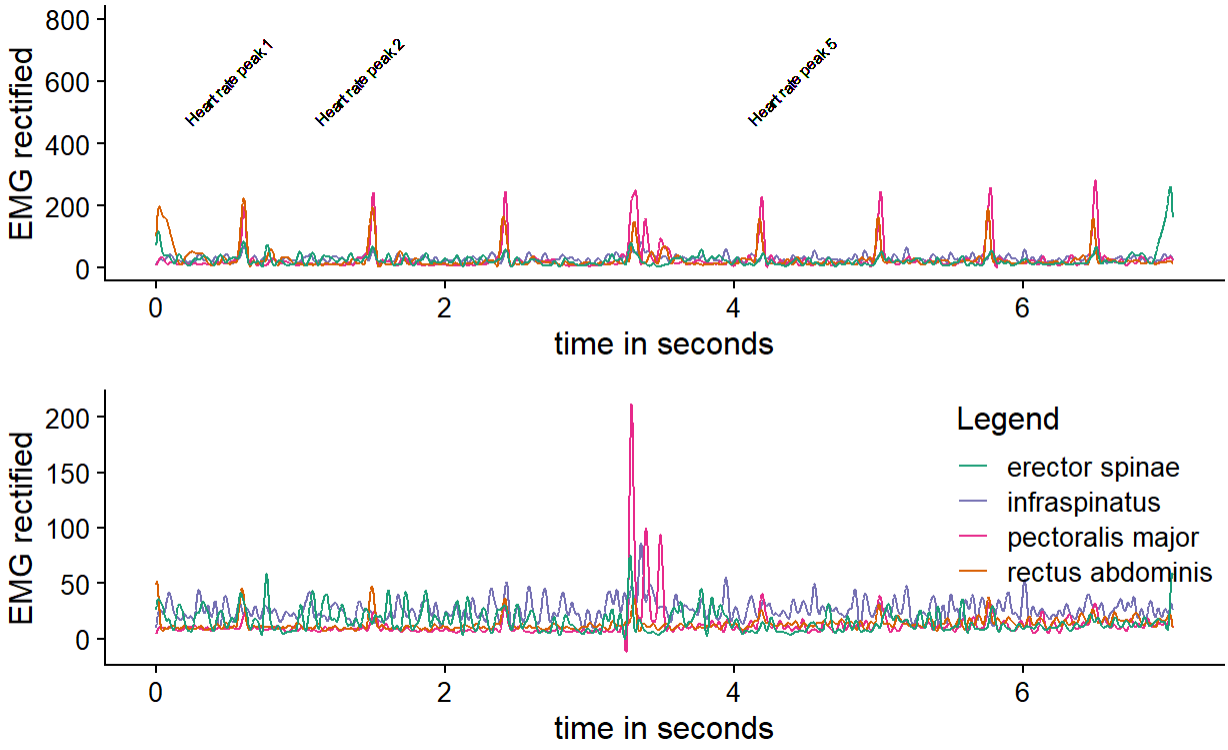

Figure S2. Example of smoothing settings for EMG signals. The upper panel shows the raw rectified high-pass filtered EMG, and the lower panel shows the low-pass filtered data to reduce artifacts of heart rate. This example shows an internal rotation trial, where we successfully retrieve the peak in pectoralis major that internally rotates the arm.

Ground reaction forces. We upsampled the balance board from 400 Hz to 2,500 Hz. We then applied a zero-phase low-pass 20 Hz 2nd order Butterworth filter to the padded signals. As a key measure for postural perturbation we computed the change in 2D magnitude (L2 norm of the center of pressure x and y) in center of pressure (hereafter COPc).

Acoustics. For acoustics we extract the smoothed amplitude envelope (hereafter envelope). For the envelope we apply a Hilbert transform to the waveform signal, then take the complex modulus to create a 1D time series, which is then resampled at 2,500 Hz, and smoothed with a Hann filter based on a Hanning Window of 12 Hz. We normalize the amplitude envelope signals within participants before submitting to analyses.

Data aggregation. All signals were sampled at, or upsampled to, 2500 Hz. Then we aggregated the data by aligning time series in a combined by-trial format to increase ease of combined analyses. We linearly interpolated signals when sample times did not align perfectly.

Data sharing & Privacy. Video data is deidentified using the masked-piper tool to mask faces and body while maintaining kinematic information (31).

Overview data analyses

Note that there is a general decline in the amplitude of the vocalization during a trial (see Fig. S3, top panel). This is to be expected, as the subglottal pressure falls when the lungs deflate. To quantify deviations from stable vocalizations, we therefore detrend the amplitude envelope time series, so as to assess positive or negative peaks relative to this trend line. For the envelope, muscle activity, and the change in center of pressure we will measure the global maxima happening within the analyses window (i.e., within a trial we take a local maximum occurring between movement onset and offset). We will analyze positive and negative peaks within the movement window separately.

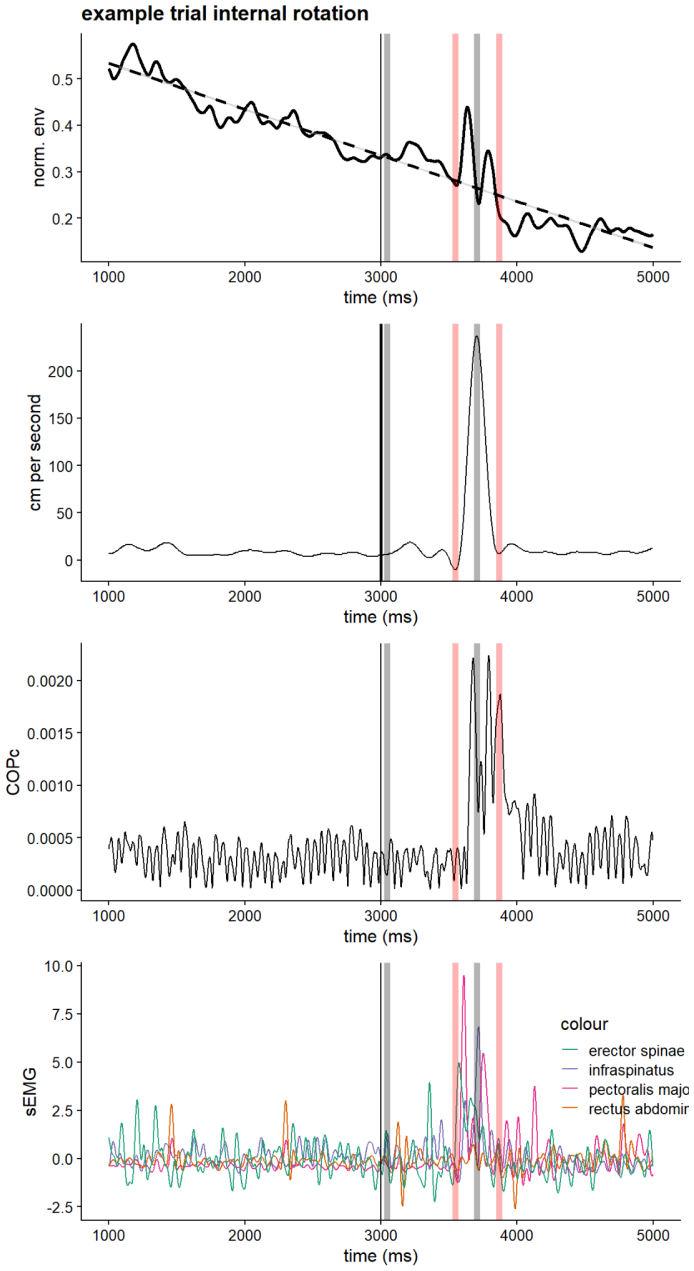

Figure S3: Combined time series. One example trial and the associated signals are shown, for the internal rotation movement condition. At time = 0 s, the prompt is given to the participant to vocalize. We determine a detrending line using linear regression for the 1 to 5 s after the vocalization prompt. Note, at 3 s (3000 ms), there is a movement prompt. However, we determine our window where we assess peaks in signals at 500 ms before and after the movement onset/offset (using peakfinding function on the 2D speed time series of the wrist). In these trials, the analysis window is given in gray dashed bars, which is 500 ms after and before movement onset/offset.

Supporting results

Descriptive results

Figure S4 shows an overview of the muscle activity patterns for each movement condition, split over weight conditions. Table S1 provides the numerical information. Table S2 provides normalized peak muscle activity by weight. All in all, these descriptive results pattern in sensible ways. Firstly, we can confirm that indeed the focal muscles powering internal (pectoralis major) or external (infraspinatus) rotation are peaking in activity during these movement conditions.

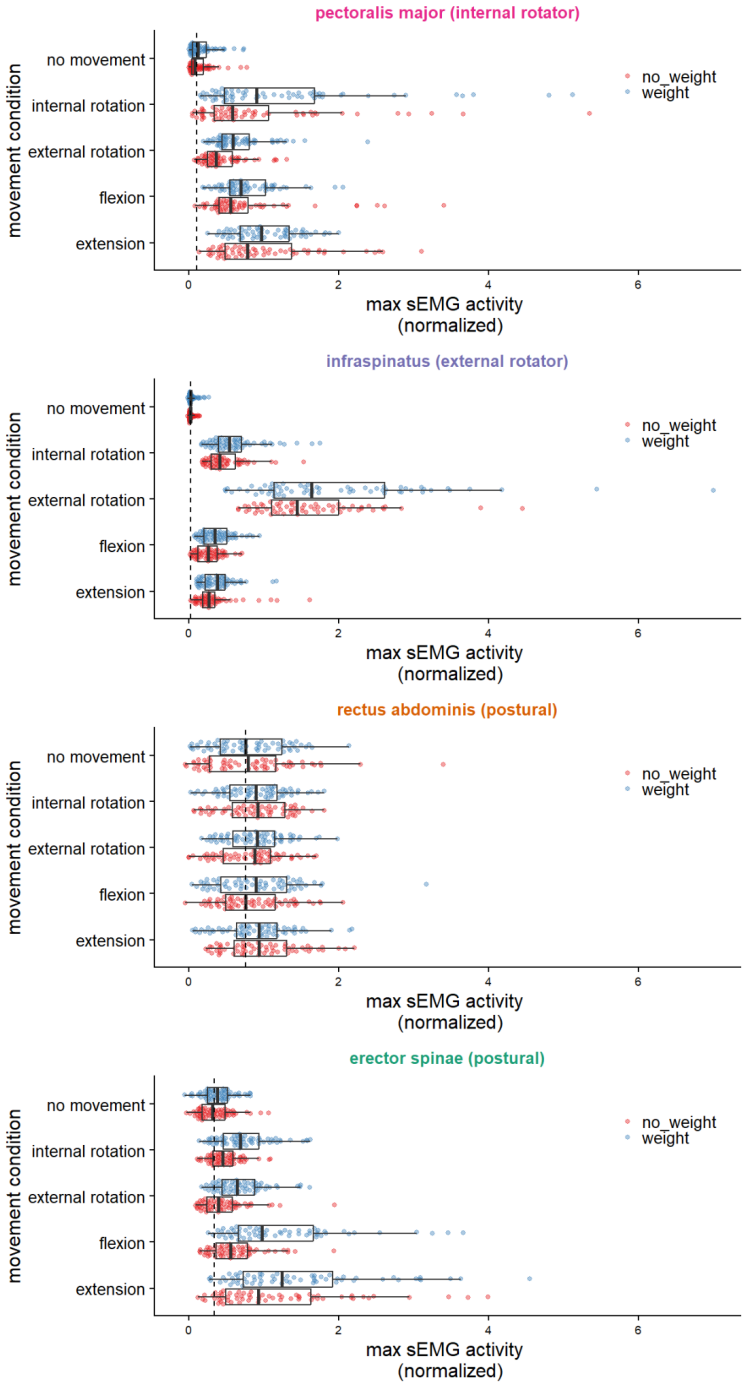

Figure S4: Manipulation check movement condition and peak muscle activity. The peak muscle activity, normalized for each muscle, is shown per movement and weight condition. A) highlights the fact that the pectoralis major, is - as to be expected - most active for the internal rotation movement condition. B) Also to be expected, the infraspinatus is most active during the external rotation. Note further, that both postural muscles seem more active during extension (and secondarily flexion).

Muscle	Movement condition	Mean	SD	Low 95%CI	High 95%CI
pectoralis major	no movement	0.16	0.01	0.13	0.19
pectoralis major	internal rotation	1.09	0.10	0.90	1.28
pectoralis major	external rotation	0.56	0.03	0.50	0.62
pectoralis major	flexion	0.77	0.05	0.68	0.86
pectoralis major	extension	1.01	0.05	0.91	1.11
infraspinatus	no movement	0.04	0.00	0.03	0.05
infraspinatus	internal rotation	0.55	0.03	0.49	0.60
infraspinatus	external rotation	1.78	0.09	1.61	1.96
infraspinatus	flexion	0.32	0.02	0.29	0.35
infraspinatus	extension	0.35	0.02	0.31	0.40
rectus abdominis	no movement	0.81	0.05	0.71	0.91
rectus abdominis	internal rotation	0.89	0.04	0.81	0.96
rectus abdominis	external rotation	0.85	0.04	0.78	0.93
rectus abdominis	flexion	0.88	0.05	0.79	0.97
rectus abdominis	extension	0.94	0.04	0.86	1.03
erector spinae	no movement	0.37	0.02	0.33	0.41
erector spinae	internal rotation	0.61	0.03	0.56	0.67
erector spinae	external rotation	0.57	0.03	0.51	0.63
erector spinae	flexion	0.92	0.06	0.80	1.04
erector spinae	extension	1.32	0.08	1.16	1.48

Table S1. Normalized peak muscle activity for the different movement conditions. These are the numerical results associated with Figure S4. The values indicate the peak EMG activity normalized for each muscle.

Further, postural muscles such as the rectus abdominis are especially active for extension movements (and secondarily flexion). Confirming their combined postural role, the muscle activity of the rectus abdominis and erector spinae are reliably correlated. Confirming their antagonistic role in rotating the humerus, the pectoralis and infraspinatus muscle activity are reliably correlated, indicative of their joint agonist/antagonistic control of posture and upper limb rotation, respectively. Lastly, from Table S2 it is clear that adding a wrist weight increases the peak muscle activity for all muscles (with no or barely overlapping 95 % confidence intervals in the weight vs. no weight condition).

Muscle	Weight condition	Mean	SD	Low 95% CI	High 95% CI
pectoralis major	no weight	0.65	0.04	0.58	0.73
pectoralis major	weight	0.79	0.04	0.71	0.86

infraspinatus	no weight	0.54	0.04	0.47	0.61
infraspinatus	weight	0.69	0.05	0.59	0.79
rectus abdominis	no weight	0.87	0.03	0.81	0.92
rectus abdominis	weight	0.88	0.03	0.83	0.93
erector spinae	no weight	0.62	0.03	0.56	0.68
erector spinae	weight	0.90	0.04	0.83	0.98

Table S2: Normalized peak muscle activity per weight condition. The values indicate the peak EMG activity (normalized per muscle) per weight condition.

Exploratory visualization

Fig. S5 shows the amplitude envelope and EMG activity of a 1-second interval around the movement onset. This plot is informative about the rationale of the confirmatory analyses. Firstly, it can be seen that in general there is a positive peak during a movement onset in the amplitude envelope (relative to the trend line for that vocalization), that may or may not be *preceded* by a less extreme negative peak. Especially, for the internal rotation we observe such a negative peak. We will therefore assess whether particular movement conditions predict higher magnitude negative and positive peaks in the amplitude envelope (analysis 1). We further assess whether particular muscles predict the magnitude of positive or negative peaks (analysis 2). Finally, we will assess whether particular muscles are related to postural stability, to confirm the posture-mediating role of the muscles (analysis 3). While we focus in this pre-registration and pilot analyses report on the peak analyses (in part to support a more straightforward power analysis), in the confirmatory study we may perform continuous trajectory analysis using generalized additive modeling similar to previous work (14).

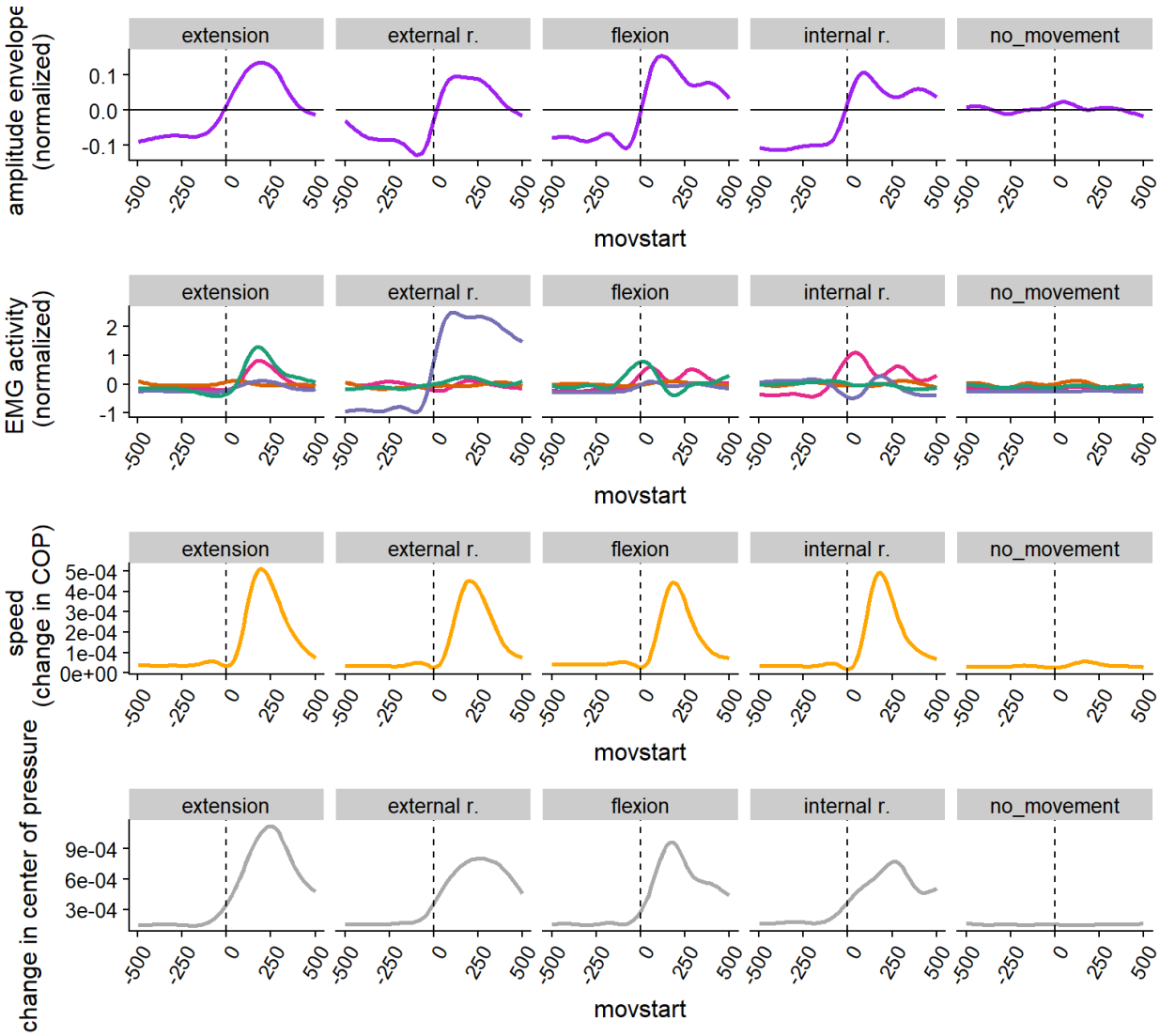

Figure S5: Smoothed conditional means for vocalization and muscle activity (with gam). Smoothed conditional means over time, where time is centered at 0 at the movement onset (as determined by the motion tracking of the wrist). The smoothed conditional means are generated by fitting non-linear smooths using generalized additive models in R-package ggplot2. It can be seen that especially for the extension movement there are clear anticipatory postural muscle activations (4), of the rectus abdominis before the movement onset, which is then followed by postural adjustments of the erector spinae. For the flexion condition, this activation pattern is reversed as one would expect given that the impulse vector should be directed in the opposite direction.

Main Confirmatory Analyses

As pre-registered, for the current confirmatory analyses we focus on the vocalization condition, ignoring the expiration baseline conditions as they are of secondary interest at this point of our inquiry.

Do different movement conditions have different effects on vocalization amplitude?

From inspecting the summarized trajectories, and the descriptive exploratory analyses above, we obtain that the internal rotation of the arm seems to start with a negative peak around the onset of the movement, which is followed by a positive peak. Furthermore, we observe that all other movements primarily have a positive peak. A straightforward test of whether the amplitude envelope has positive or negative peaks is to assess differences in peaks in the vocalization conditions per movement (and weight) condition.

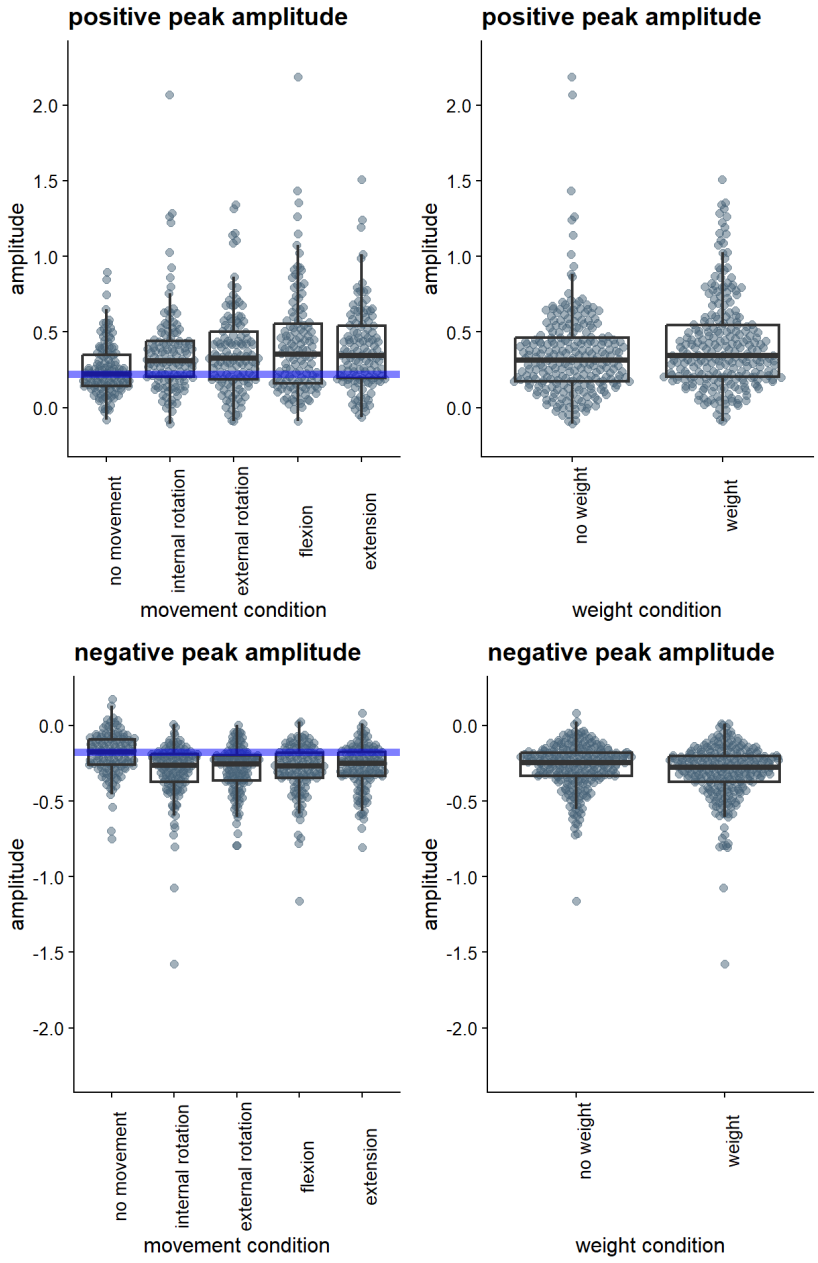

Figure S6: Effects of movement condition of positive and negative peaks in the voice amplitude. Note: The upper part of Fig. S6 shows the positive peaks in the amplitude envelope during the different movement and weight conditions. The lower part shows the negative peaks (hence the negative values; note, in the modeling we will absolutize these values). It is clear that relative to vocalization of no movement, there are especially positive, but also more negative peaks in the amplitude envelope for the different movement conditions.

We first modeled with a mixed linear regression the variation in *positive* peaks in the amplitude envelope (using R-package lme4), with participant as random intercept (for more complex random slope models did not converge). A model with weight and movement

conditions explained more variance than a base model predicting the overall mean, Change in Chi-Squared (5.00) = 36.06, $p < .001$). The model coefficients are given in Table S3. It can be seen that there is a positive but not statistically reliable effect of wrist weight in this sample. Further, all movements (extension, flexion, internal rotation, external rotation) lead to statistically reliable increases in positive peaks in the amplitude envelope relative to the no movement condition (with flexion and external rotation leading to more extreme effects).

	Estimate	SE	df	t-value	p-value
Intercept (no movement/no weight)	0.233	0.037	42.576	6.361	< .001
vs. weight	0.058	0.021	603.790	2.825	0.005
vs. internal rotation	0.104	0.033	602.729	3.204	< .001
vs. external rotation	0.116	0.032	602.740	3.586	< .001
vs. flexion	0.164	0.033	602.778	5.057	< .001
vs. extension	0.130	0.032	602.718	4.005	< .001

Table S3. Effects of weight and movement condition on magnitude positive peaks in amplitude envelope.

Note that in the previous model we assess the role of weight vs. no weight. But during the no movement condition the weight is not affecting the participant. Thus, a better estimate of the effect of wrist weight is to assess weight vs. no weight conditions for only the movement conditions. Below it shows indeed that this comparison further increases the effect strength as compared to a model with the no movement condition included.

	Estimate	SE	df	t-value	p-value
Intercept (no weight)	0.361	0.035	20.170	10.296	< .001
vs. weight	0.062	0.024	484.657	2.570	0.01

Table S4: Effects of weight on magnitude of positive peaks in amplitude envelope (no movement condition filtered out).

Secondly, we modeled in a similar way the negative peaks in the amplitude envelope, and found that a model with weight and movement conditions explained more variance than a base model predicting the overall mean, Change in Chi-Squared (5.00) = 48.54, $p < .001$). Model

coefficients are shown in Table S5. We find that some movement conditions (flexion, internal rotation, and especially external rotation) had higher magnitude negative peaks relative to no movement. No reliable effect of weight condition was found, nor did the extension movement lead to negative peaks relative to the no movement condition. Note, in our models we have absolutized the values of negative peaks, such that positive effects of some condition means higher magnitude negative peaks (i.e., more negative peaks).

	Estimate	SE	df	t-value	p-value
Intercept (no movement/no weight)	0.185	0.024	30.928	7.747	< .001
vs. weight	0.020	0.011	603.514	1.784	0.075
vs. internal rotation	0.110	0.018	602.850	6.102	< .001
vs. external rotation	0.094	0.018	602.857	5.238	< .001
vs. flexion	0.094	0.018	602.879	5.243	< .001
vs. extension	0.080	0.018	602.843	4.478	< .001

Table S5: Effects of weight and movement condition on magnitude negative peaks in amplitude envelope.

	Estimate	SE	df	t-value	p-value
Intercept (no weight)	-0.274	0.024	18.500	-11.611	< .001
vs. weight	-0.029	0.013	484.464	-2.264	0.024

Table S6: Effects of weight on magnitude of negative peaks in amplitude envelope (no movement condition filtered out).

Is different muscle activity differently affecting amplitude of vocalization?

Since each movement and weight condition is designed to recruit different muscle activations, we can also directly relate muscle activity peaks with the positive and negative peaks in the amplitude envelope. We use a similar linear mixed regression approach to model variance in vocal amplitude peaks with peaks in muscle activity for the different muscles measured. Since there are correlations between these muscle activities (see Tables S10, S11, S12, S13), we first assessed the VIF's between the muscle activity peaks, which yielded a maximum VIF value of 1.29. Since this is considered a low value (VIF > 5 is generally considered problematic), we can

combine the different muscle activity measurements in one model to predict amplitude envelope peaks. Figure S7 shows the graphical results of these relationships.

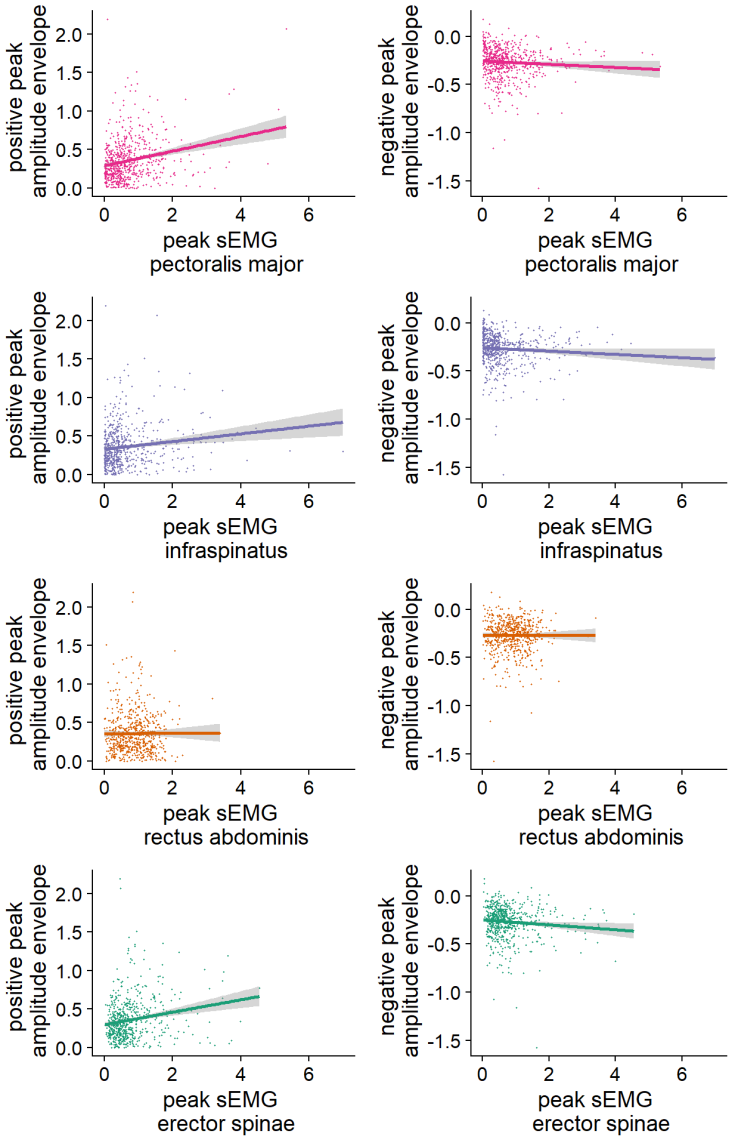

Figure S7: Relations between peak muscle activity and positive peaks in the amplitude envelope.

A model with participant as random intercept (a more complex random slope model did not converge) the different peak muscle activities explained more variance than a base model predicting the overall mean, Change in Chi-Squared = 59.21 (4.00), $p < .001$). The model coefficients are given in Table S7. Further, Table S7 shows that peak EMG activity in all the muscles (except for the rectus abdominis) leads to statistically reliable increases in positive peaks in the amplitude envelope (with stronger effects).

	Estimate	SE	df	t-value	p-value
Intercept	0.240	0.037	54.499	6.429	< .001
erector spinae	0.016	0.004	617.693	4.207	< .001
infraspinatus	0.006	0.002	612.279	2.832	0.005
pectoralis major	0.013	0.003	617.423	4.819	< .001
rectus abdominis	-0.003	0.007	612.589	-0.371	0.711

Table S7: Linear mixed regression model assessing the relation of peak muscle activity with the positive peak in the amplitude envelope.

We similarly modeled the variance of the magnitude in negative peaks in the amplitude envelope with muscle activity peaks (see right side of Figure S7 for graphical results), and we similarly observed that such a model performed better than a base model predicting the overall mean, Change in Chi-Squared = 16.48 (4.00), $p = 0.002$). The model coefficients are given in Table S8. Further, as shown in Table S8 for the infraspinatus and the pectoralis major, increases in peak EMG activity lead to statistically reliable increases in the magnitude of negative peaks in the amplitude envelope (with stronger effects), while the effects of erector spinae and the rectus abdominis do not reach statistical reliability.

	Estimate	SE	df	t-value	p-value
Intercept	0.227	0.026	36.153	8.830	< .001
erector spinae	0.004	0.002	613.122	1.943	0.053
infraspinatus	0.003	0.001	608.863	2.328	0.020
pectoralis major	0.003	0.002	612.841	2.152	0.032
rectus abdominis	0.001	0.004	619.957	0.359	0.719

Table S8: Linear mixed regression model assessing the relation of peak muscle activity with the negative peak in the amplitude envelope. The negative peaks in the amplitude envelope are absolutized.

Is a particular muscle related to postural stability?

Finally, we will assess which muscle activity can be related to changes in the center of pressure, which would directly confirm that gesture-speech biomechanics relate to postural stability (12).

Figure S8 shows the graphical results. We similarly performed a linear mixed regression (with participant as random intercept) with a model containing peak EMG activity for each muscle which was regressed on the peak in change in the center of pressure (COPc). We obtained that a base model predicting the overall mean of COPc was outperformed relative to said model, Change in Chi-Squared (4.00) = 180.01 , $p < .001$. Table S9 provides the coefficient information. We find that only the postural muscles (rectus abdominis, erector spinae) indeed reliably predict the magnitude of changes in the center of pressure, while the pectoralis major and the infraspinatus do not reliably relate to the changes in ground reaction forces.

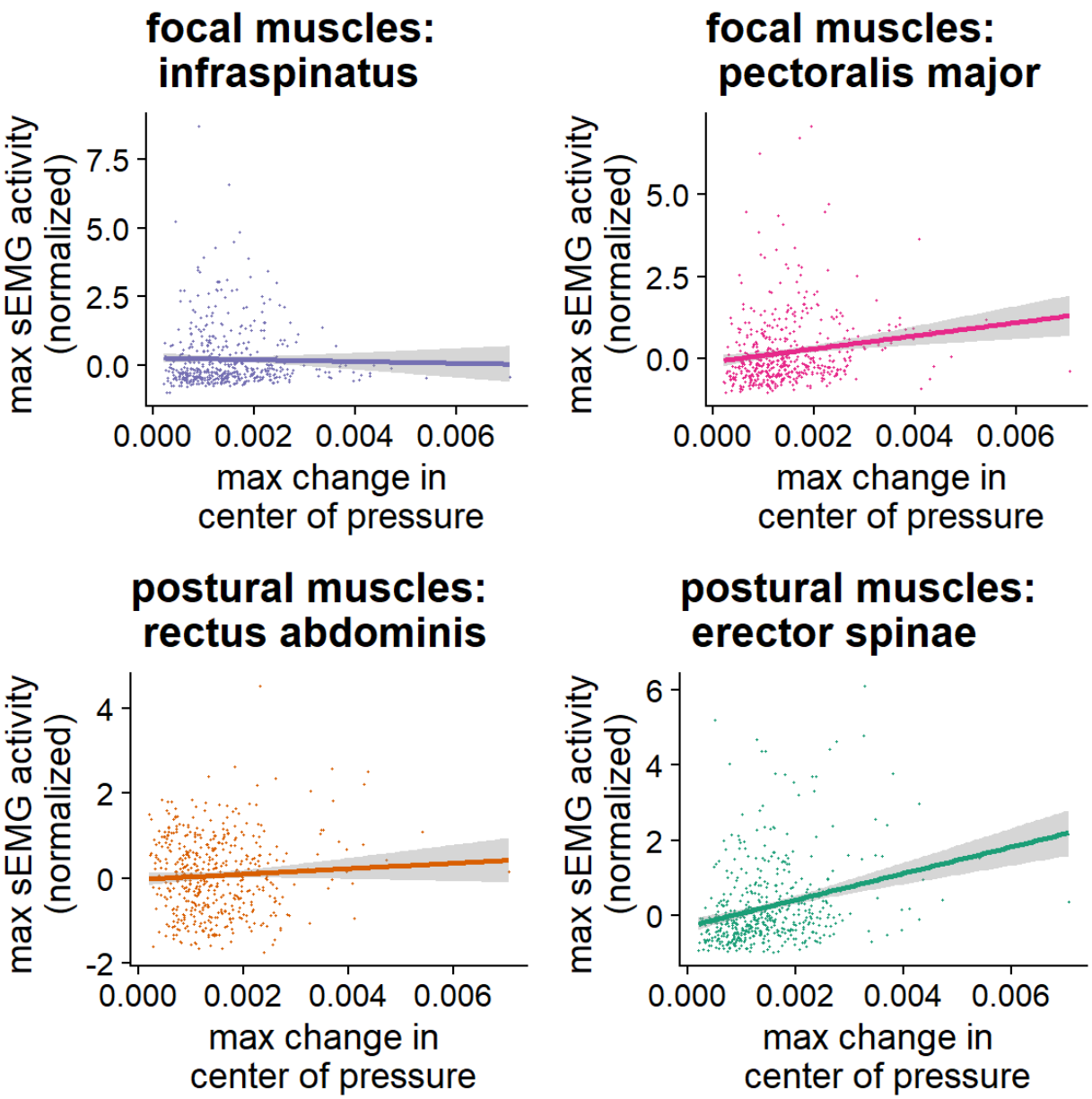

Figure S8: Confirmation of postural muscle activation during changes in center of mass.

	Estimate	SE	df	t-value	p-value
Intercept	-0.875	0.122	55.418	-7.169	< .001
erector spinae	0.118	0.012	617.831	9.752	< .001
infraspinatus	0.028	0.007	612.430	4.168	< .001
pectoralis major	0.063	0.009	617.566	7.065	< .001
rectus abdominis	0.025	0.023	611.990	1.109	0.268

Table S9: Linear mixed regression model for predicting peak change in center of pressure based on muscle activity.

Other Exploratory Analyses

Along with the confirmatory analyses reported above that relate to questions posed in the pre-registration, we report on additional, exploratory analyses. Firstly we assess for possible context-dependent interrelationships of posture, muscle activity, and voice, by producing correlation matrices per movement condition (see Tables S10, S11, S12, S13) which are represented in our synergies plot in Fig. 2 (panel F) in the main article. We see that producing different movements can recruit variable interrelationships between muscle, posture and voice. Striking to pick out is for example that we find a direct relationship between the degree of center of pressure change and the magnitude of positive amplitude peaks in the voice.

Table S10 shows the correlation matrix for the peaks in EMG (row and columns 1-4), center of pressure (5) and amplitude (6). For this table, the data only include the **internal rotation** condition.

	1	2	3	4	5
1. max. EMG pectoralis major	-				
2. max. EMG infraspinatus	.34***	-			
3. max. EMG rectus abdominis	.00	-.16	-		
4. max. EMG erector spinae	.17	.09	.16	-	
5. max. change in center of pressure	.30***	.23*	-.07	.21*	-
6.max. positive amplitude vocalization	.42***	.22*	-.02	.10	.51***

Table S10. Correlation matrix for internal rotation.

Table S11 shows the correlation matrix for the maximum peaks in EMG (row and columns 1-4), change in center of pressure (5) and maximum vocal amplitude (6). For this table, the data only include the **external rotation** condition.

	1	2	3	4	5
1. max. EMG pectoralis major	-				

2. max. EMG infraspinatus	.04	-			
3. max. EMG rectus abdominis	-.14	.06	-		
4. max. EMG erector spinae	.20*	.07	.15	-	
5. max. change in center of pressure	.12	.12	-.09	.27**	-
6.max. positive amplitude vocalization	.27**	.18*	-.05	.25**	.04

Table S11. Correlation matrix for external rotation.

Table S12 shows the correlation matrix for the maximum peaks in EMG (row and columns 1-4), change in center of pressure (5) and maximum vocal amplitude (6). For this table, the data only include the **extension** condition.

	1	2	3	4	5
1. max. EMG pectoralis major	-				
2. max. EMG infraspinatus	.18*	-			
3. max. EMG rectus abdominis	.06	-.02	-		
4. max. EMG erector spinae	.20*	.05	.15	-	
5. max. change in center of pressure	.21*	.32***	.18*	.17	-
6.max. positive amplitude vocalization	-.02	.29***	-.01	.26**	.17

Table S12: Correlation matrix for extension.

Table S13 shows the correlation matrix for the maximum peaks in EMG (row and columns 1-4), change in center of pressure (5) and maximum vocal amplitude (6).. For this table, the data only include the **flexion** condition.

	1	2	3	4	5
1. max. EMG pectoralis major	-				
2. max. EMG infraspinatus	.18*	-			
3. max. EMG rectus abdominis	-.19*	-.02	-		
4. max. EMG erector spinae	.03	.26**	.05	-	
5. max. change in center of pressure	.13	.34***	.02	.04	-
6.max. positive amplitude vocalization	.09	.23**	.00	.07	.16

Table S13: Correlation matrix for flexion.

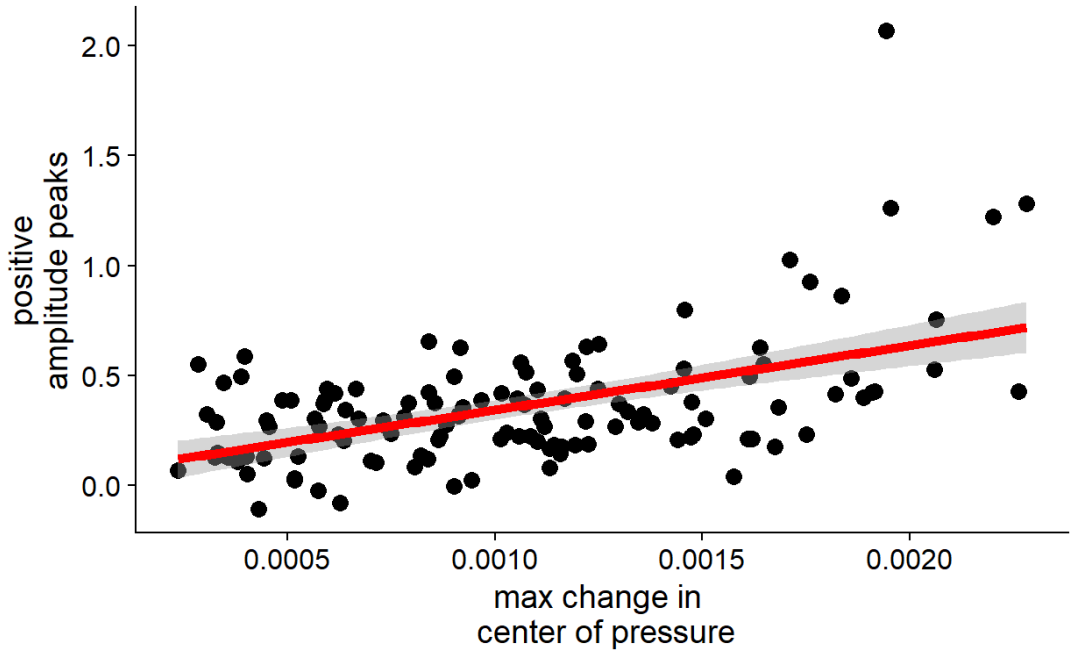

Figure S9. Correlation of posture and amplitude for internal rotation.

Figure S9 visualizes the positive correlation between posture (max. change in center of pressure) and vocalization amplitude (positive peaks in amplitude) found for the internal rotation condition in Table S10.

We finally explored the timing relationship of negative and positive peaks in the amplitude of the voice relative to the peak in speed of the hand. The density plot in Fig. S10 visualizes the temporal relationship between peak speed in movement and negative (left) and positive (right) peaks in the amplitude envelope, relative to the moment the moving hand reaches a peak in speed during the trial (i.e., t=0, at peak speed of the wrist). Our findings corroborate our general observation that positive (rather than negative) peaks seem to be more robustly driven by gesture kinetics. In the timing relations obtained, the negative peaks that we observe are more variably distributed around peak speed, while for the positive peaks we see clear occurrences just around the movement onset (just before the peak speed at t = 0). This result, next to our more robust relationship between positive peaks and the muscle and posture signals, makes us conclude that in general the movements performed in our experiment are increasing subglottal pressures increasing the voice's amplitude.

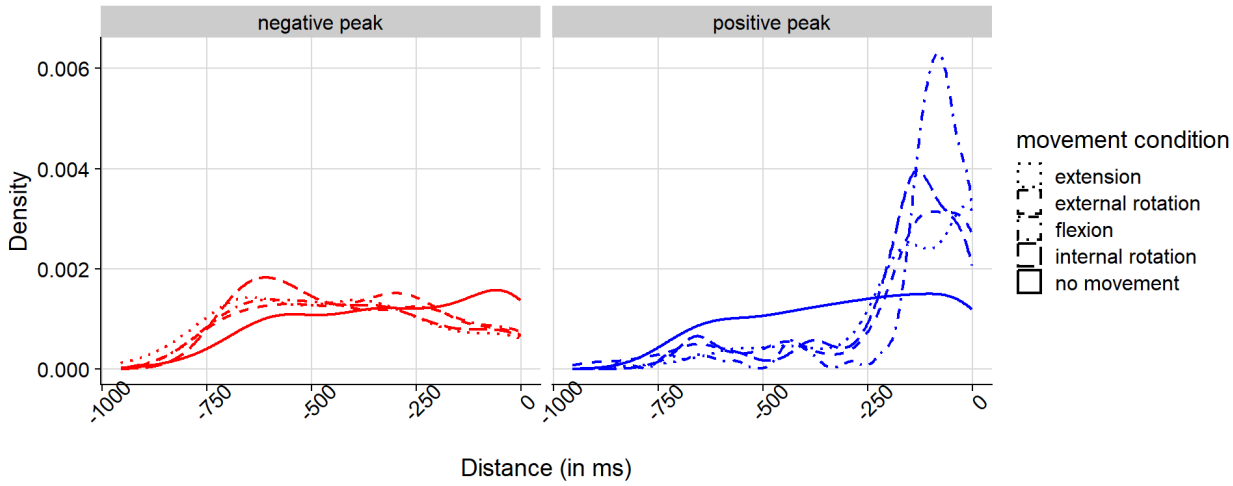

Figure S10: Density plot showing temporal distance (in milliseconds) of positive and negative peaks in amplitude from peak speed in wrist movement.

References

1. J. A. Seikel, D. G. Drumright, D. J. Hudock, *Anatomy & Physiology for Speech, Language, and Hearing* (Plural Publishing, Incorporated, 2019).
2. P. W. Hodges, S. C. Gandevia, Changes in intra-abdominal pressure during postural and respiratory activation of the human diaphragm. *J. Appl. Physiol.* **89**, 967–976 (2000).
3. P. W. Hodges, C. A. Richardson, Feedforward contraction of transversus abdominis is not influenced by the direction of arm movement. *Exp Brain Res* **114**, 362–370 (1997).
4. A. S. Aruin, M. L. Latash, Directional specificity of postural muscles in feed-forward postural reactions during fast voluntary arm movements. *Exp Brain Res* **103**, 323–332 (1995).
5. S. L. Morris, B. Lay, G. T. Allison, Transversus abdominis is part of a global not local muscle synergy during arm movement. *Hum Mov Sci* **32**, 1176–1185 (2013).
6. D. Lasserson, *et al.*, Differences in motor activation of voluntary and reflex cough in humans. *Thorax* **61**, 699–705 (2006).
7. T. E. Dolmage, *et al.*, Arm elevation and coordinated breathing strategies in patients with COPD. *Chest* **144**, 128–135 (2013).
8. X. Liu, *et al.*, Evaluation of isokinetic muscle strength of upper limb and the relationship with pulmonary function and respiratory muscle strength in stable COPD patients. *COPD* **14**, 2027–2036 (2019).
9. M. Zedka, A. Prochazka, Phasic activity in the human erector spinae during repetitive hand movements. *J Physiol* **504**, 727–734 (1997).
10. S. Fuchs, A. Rochet-Capellan, The Respiratory Foundations of Spoken Language. *Annual Review of Linguistics* **7**, null (2021).
11. P. Wagner, A. Ćwiek, B. Samlowski, Exploiting the speech-gesture link to capture fine-grained prosodic prominence impressions and listening strategies. *Journal of Phonetics* **76**, 100911 (2019).
12. D. Bolinger, Intonation and gesture. *American Speech* **58**, 156–174 (1983).
13. W. Pouw, S. Fuchs, Origins of vocal-entangled gesture. *Neuroscience & Biobehavioral Reviews* **141**, 104836 (2022).
14. W. Pouw, S. J. Harrison, N. Esteve-Gibert, J. A. Dixon, Energy flows in gesture-speech physics: The respiratory-vocal system and its coupling with hand gestures. *The Journal of the Acoustical Society of America* **148**, 1231–1247 (2020).
15. W. Pouw, S. J. Harrison, J. A. Dixon, Gesture-speech physics: The biomechanical basis of the emergence of gesture-speech synchrony. *Journal of Experimental Psychology: General* **149**, 391–404 (2019).
16. W. Pouw, A. Paxton, S. J. Harrison, J. A. Dixon, Acoustic information about upper limb movement in voicing. *PNAS* **117**, 11364–11367 (2020).
17. L. Pearson, W. Pouw, Gesture–vocal coupling in Karnatak music performance: A neuro–bodily distributed aesthetic entanglement. *Annals of the New York Academy of Sciences* **n/a** (2022).
18. W. Pouw, L. de Jonge-Hoekstra, S. J. Harrison, A. Paxton, J. A. Dixon, Gesture-speech physics in fluent speech and rhythmic upper limb movements. *Annals of the New York Academy of Sciences* **1491**, 89–105 (2020).
19. P. W. Hodges, R. Sapsford, L. H. M. Pengel, Postural and respiratory functions of the pelvic floor muscles. *Neurourol. Urodyn.* **26**, 362–371 (2007).

20. V. Pettersen, Preliminary findings on the classical singer's use of the pectoralis major muscle. *FPL* **58**, 427–439 (2006).
21. P. Cavallari, F. Bolzoni, C. Bruttini, R. Esposti, The Organization and Control of Intra-Limb Anticipatory Postural Adjustments and Their Role in Movement Performance. *Front. Hum. Neurosci.* **10** (2016).
22. S. Bouisset, M.-C. Do, Posture, dynamic stability, and voluntary movement. *Neurophysiologie Clinique/Clinical Neurophysiology* **38**, 345–362 (2008).
23. F. G. Baldissera, L. Tesio, APAs Constraints to Voluntary Movements: The Case for Limb Movements Coupling. *Frontiers in Human Neuroscience* **11** (2017).
24. W. C. Lancaster, O. W. Henson, A. W. Keating, Respiratory muscle activity in relation to vocalization in flying bats. *Journal of Experimental Biology* **198**, 175–191 (1995).
25. M. T. Turvey, S. T. Fonseca, The medium of haptic perception: A tensegrity hypothesis. *Journal of Motor Behavior* **46**, 143–187 (2014).
26. M. Brysbaert, M. Stevens, Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition* **1**, 9 (2018).
27. S. P. Coundouris, A. G. Adams, S. A. Grainger, J. D. Henry, Social perceptual function in Parkinson's disease: A meta-analysis. *Neuroscience and Biobehavioral Reviews* **104**, 255–267 (2019).
28. A. K. Ho, R. Iansek, J. L. Bradshaw, Motor Instability in Parkinsonian Speech Intensity. *Cognitive and Behavioral Neurology* **14**, 109–116 (2001).
29. S. Sharma, K. Fleck, S. Winslow, K. Rothermich, The Impact of Parkinson's Disease on Social Communication: An Exploratory Questionnaire Study (2021) https:/doi.org/10.31235/osf.io/3x42a (August 24, 2022).
30. M. J. Richardson, K. L. Marsh, R. C. Schmidt, Effects of visual and verbal interaction on unintentional interpersonal coordination. *J Exp Psychol Hum Percept Perform* **31**, 62–79 (2005).
31. J. J. Gibson, *The senses considered as perceptual systems* (Houghton Mifflin, 1966).
32. C. Lugaresi, *et al.*, MediaPipe: A Framework for Building Perception Pipelines (2019) https:/doi.org/10.48550/arXiv.1906.08172 (August 30, 2022).
33. B. Owoyele, J. Trujillo, G. de Melo, W. Pouw, Masked-piper: Masking personal identities in visual recordings while preserving multimodal information. *SoftwareX* (2022) https:/doi.org/10.31234/osf.io/bpt26 (June 3, 2022).
34. J. D. M. Drake, J. P. Callaghan, Elimination of electrocardiogram contamination from electromyogram signals: An evaluation of currently used removal techniques. *Journal of Electromyography and Kinesiology* **16**, 175–187 (2006).