Movement-related muscle activity and kinetics affect human vocalization amplitude

Wim Pouw*[1], Lara Burchardt [1, 2], Raphael Werner[1], and Luc Selen[1]

*Corresponding Author: wim.pouw@donders.ru.nl
[1]Donders Institute for Brain, Cognition and Behavior, Radboud University Nijmegen
[2]Leibniz Institute for General Linguistics ZAS Berlin

Voice production can be a whole-body affair: Upper limb movements physically impact the voice in steady-state vocalization, speaking, and singing. This is supposedly due to biomechanical impulses on the chest-wall, affecting subglottal pressure. Unveiling such biomechanics is important, as humans gesture with their hands in a synchronized way with speaking. Here we assess biomechanical interactions between arm movements and the voice, by measuring activity of respiratory-related muscles during different types of upper limb movement. We show that positive peaks in the voice's amplitude increase with movements and more so with a 1 kg weight attached to the wrist. We further report exploratory findings that gesture-related muscle activations scale with positive peaks in the voice's amplitude. These results indicate that the voice aligns with the forces generated by the body and implies that the voice evolved in the context of bodily action.

1. Introduction

In principle, any muscle attached to the human rib cage can act on it and thus affect the subglottal pressure that supports voicing. Consequently, there are many potential respiratory-vocal muscles. Said muscles would include those around the chest (e.g., pectoralis major), abdomen (rectus abdominis), and back (erector spinae, serratus posterior/anterior). Passive breathing and speaking, however, is mainly driven by the diaphragm and the intercostal muscles between the ribs (Aliverti, 2016; Seikel et al., 2019). Only on rare occasions, when coughing, shouting, or breathing deeply, humans recruit other so-called "accessory" respiratory muscles such as the abs and pectoral muscles (Aliverti, 2016; Lasserson et al., 2006; Seikel et al., 2019).

Yet, when humans speak or sing, it is common to move the upper limbs expressively at the same time – called gesture (Pearson & Pouw, 2022; Wagner et al., 2014). Such upper limb movements will recruit a whole range of upper body muscles, including those involved in maintaining posture (Cordo & Nashner, 1982). Several of these muscles attach to the rib cage (e.g., abdominal and pectoral muscles) and are classically listed as accessory to respiratory functioning (Seikel et al., 2019). Given that speaking requires subtle modulations of subglottal pressure (Rubin et al., 1967; Sundberg et al., 1993a; Sundberg et al. 1993b), co-gesture speaking must be in some way coordinated with the respiratory-and-thus-vocal muscles that are activated during gesturing (see Pouw & Fuchs, 2022).

There is considerable evidence that gestural arm movements affect the voice directly, as summarized by Pouw et al. (2019b): More extreme peaks in the acceleration of movements bigger (arm) and smaller (wrist) upper limb movements relate to more chest-circumference changes, which is associated with more extreme acoustic effects on the intensity of vocal sound (via increasing subglottal pressure). Furthermore, acoustic effects of upper limb movements are more pronounced when subjects are in a less stable standing vs. sitting position (Pouw et al., 2019b). This all ties in with the idea that a physical impulse (mass x acceleration), impacts posture (especially when standing), recruiting respiratory-related muscles (that change chest circumference), which impacts respiratory-vocal functioning (such that intensity is affected).[1] These previous gesture-speech physics[2] studies assessed continuous voicing, mono-syllable utterances, and fluent speech production (Pouw et al., 2020). However, direct evidence of mass and muscle activity relating to gesture-speech physics has so far not been reported.

We ask two questions here: 1) Do different upper limb movements lead to dissociable positively peaked deviations of the amplitude envelope of ongoing voicing? 2) Does peak muscle activity predict positively peaked deviations of the amplitude envelope of ongoing voicing? In addition to the results from the two pilot participants, we also report confirmatory pre-registered results with respect to the first question assessing differences in positive amplitude peaks.

Based on the gesture-speech physics account (Pouw & Fuchs, 2022) we predict that movements will increase the magnitude of positive amplitude peaks in the voice and that we should find (for some muscles) that peak muscle activity relates to the amplitude envelope.

2. Method

We report exploratory results with N = 2 participants, and some confirmatory pre-registered results (N = 17).[3] For the current pilot experiment supporting the pre-registration, the first author (Dutch-speaking; male; BMI = 21.7) and a volunteer

[1] Such interactions, between pectoral/upper limb activity and respiratory-vocal states, have been well-studied in non-human animals (Cooper & Goller, 2004; Lancaster et al., 1995; Blumberg, 1992).

[2] Gesture-speech physics (Pouw et al., 2019a) does not require the speaker to move to modulating vocalization, but conversely moving may affect vocalizations, or even voiceless expirations (Werner et al., 2024).

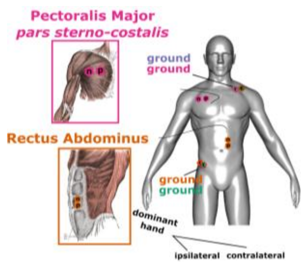
[3] The current study has been approved by the Ethics Committee Social Sciences (ECSS) of the Radboud University (reference nr.: 22N.002642).

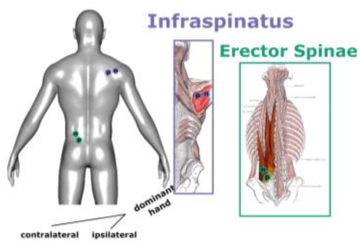
female (Dutch-speaking; BMI = 21.5) performed the experiment.[4] We also report a selection of the confirmatory results we pre-registered. This full dataset consisted of N = (17) participants (7 f, 10 m), M (SD) age = 28.5 (6.5), BMI = 23.40 (2.20).

The 1-hour study involved a two-level within-subject factor wrist-weight manipulation (no weight, 1 kg weight), a two-level within subject vocalization condition (expire, vocalize), and a five-level within-subject movement condition (‘no movement’, ‘elbow extension’, ‘elbow flexion’, ‘shoulder external rotation’, ‘shoulder internal rotation’). With 4 trial repetitions over the experiment, we yield 80 trials per participant. Trials were blocked by weight condition and vocalization condition. Within blocks, all movement conditions were randomized.

To manipulate the mass set in motion, we apply a wrist weight. We use a TurnTuri sports wrist weight of 1 kg. The experiment was coded in Python using functions from PsychoPy. The participants were recorded via a video camera. We used Mediapipe (Lugaresi et al., 2019) to track the skeleton and facial movements, which is implemented in Masked-piper which we also use for masking the videos (Owoyele et al., 2022).

We measured surface ElectroMyoGraphy (sEMG) using a wired BrainAmp ExG system (sampling rate: 2,500 Hz). Disposable surface electrodes were used, and for each of the four target muscles we had 3 (positive, negative, ground) electrodes. Positive and negative electrodes were attached with a 15 mm distance center to center. We applied electrodes for focal muscles which directly participate in the internal (pectoralis major) and external rotation (infraspinatus) of the humerus. We attached the electrodes for focal muscles ipsilaterally (relative to the dominant hand) to the muscle belly of the clavicular head of the pectoralis major, with a ground electrode on the clavicle on the opposite side. We recorded postural muscles: applying electrodes to muscles that anticipate and react to postural perturbations due to upper limb movements. Electrodes for these muscles were attached contralaterally to the moving dominant hand to the rectus abdominis and the erector spinae muscle group (specifically, the iliocostalis romborum), with ground electrodes on the iliac crest on the opposite side.





For audio recordings, we used a headset microphone sampling at 16 kHz. The gain levels of the condenser power source were set by the hardware. We also record ground reaction forces, but these will not be discussed here.

[4] More detail on participants, equipment, and data processing can be found in the pre-registration.

We use LabStreamLayer as an interface for data synchronization across signals. After body measurements, we applied the surface EMG. After practice trials, participants performed 80 blocked trials. For each trial, participants were closely guided by the information on the monitor.

Participants were instructed to adopt the start position of the movement, which is a 90° elbow flexion, with either an externally rotated humerus (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 were asked to rest their arms alongside their bodies. Upon trial start, participants inhaled deeply with a timer counting down from 4 seconds. Subsequently, participants were asked to continuously 'vocalize' with a *schwa* sound, or 'expire', with a screen appearing after 3 seconds to perform the movement with visual guidance to where the movement's end position is so that participants are reminded of the movement. After an additional 4 seconds, the trial ends, which gives enough time to perform the movement and stabilize vocalization after the perturbation. Participants were explicitly instructed to keep their vocalization as stable as possible during the different movement conditions. To cater for vocal amplitude decrease as the lungs deflate over time, we detrended the amplitude envelope time series and expressed positive peaks relative to this trend line.

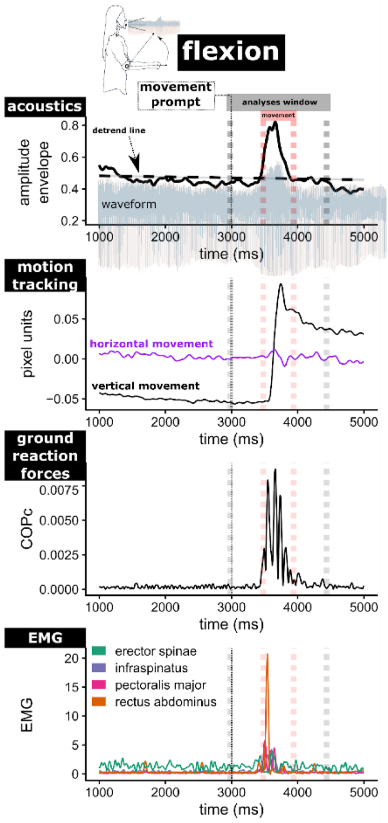


Figure 2: Time series example for a flexion movement + vocalization trial.

We only analyze vocalization trials here.

An example time series is shown in Fig. 2. At time = 0, 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. At 3,000 ms there is a movement prompt. We assess signal peaks in a time window running from 500 ms before movement onset to 500 ms after movement offset, as indicated by gray dashed bars.

3. Exploratory Results

3.1 Effects of different movements on positive peaks in vocalization

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.[5] The model coefficients are given in Table 1. There is a positive, but not statistically reliable effect of wrist weight in this *exploratory* sample. Further, all movements (extension, flexion, internal rotation, external rotation) lead to statistically reliable increases in positive peaks in the vocalization amplitude envelope relative to the no movement condition (with flexion and external rotation leading to more extreme effects).

Table 1: Effects of weight and movement condition on positive peaks in the voicing amplitude envelope.

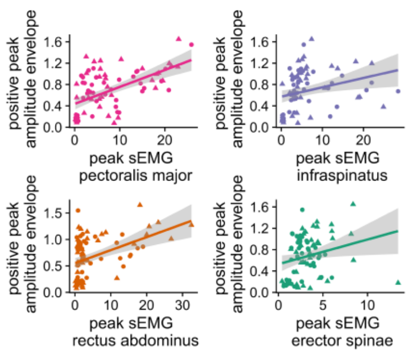
	Est.	SE	t-value	p-value
Intercept	0.164	0.101	1.63	0.181
vs. weight	0.062	0.063	0.98	0.328
vs. extension	0.436	0.100	4.35	<.001
vs. flexion	0.618	0.100	6.17	<.001
vs. internal r.	0.568	0.100	5.67	<.001
vs. external r.	0.794	0.100	7.92	<.001

3.2 Effects of muscle activity on positive peaks in vocalization

We also directly relate muscle activity peaks with the positive peaks in the amplitude envelope (after checking for collinearity). The model coefficients (for a model with the different peak muscle activities as predictor and participant as random intercept), are given in Table 2 and show that peak EMG activity in all the muscles (but especially the rectus abdominis, a well-known expiratory muscle) leads to statistically reliable increases in positive peaks in the amplitude envelope.

Table 2 and Figure 3: Muscle activity effects on magnitude positive peaks in vocalization. In the figure triangles indicate movements with wrist-weight.

	Est.	SE	t-value	p-value
Intercept	0.311	0.094	3.315	0.072
erector spinae	-0.019	0.017	-1.112	0.270
infraspinatus	0.009	0.005	1.825	0.072
pectoralis major	0.029	0.005	5.973	<.001
rectus abdominis	0.030	0.006	5.266	<.001



4. Confirmatory Results Research Question 1

We here report the *confirmatory* results for the effects of different movements on positive peaks in vocalization, in a model including weight and movement

[5] Every model reported was compared to a base model predicting the overall mean and explained more variance than the base model.

conditions. We confirm that different types of movement, as well as adding a weight to the moving wrist, increases the positive peaks in the amplitude envelope relative to a no movement/no weight condition.

	Est.	SE	t-value	p-value
Intercept	0.233	0.037	6.361	<.001
vs. weight	0.058	0.021	2.825	0.005
vs. extension	0.130	0.032	4.005	<.001
vs. flexion	0.164	0.033	5.057	<.001
vs. internal r.	0.104	0.033	3.204	0.001
vs. external r.	0.116	0.032	3.586	<0.001

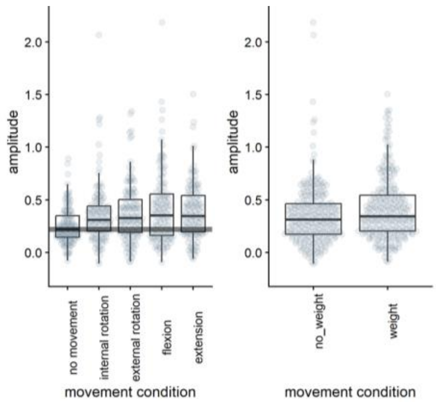


Table 4 & Figure 5: Confirmatory effects of weight and movement condition on positive peaks in the amplitude envelope.

5. Discussion & Conclusion

In summary: Movement of the upper limb yields unintentional positive peaks on the amplitude envelope of vocalization. This can be labeled 'unintentional' as the task is to produce a stable vocalization output. Further, we show promising exploratory results that activity of specific muscles is reliably related to these positive peaks in the voice's amplitude. Though in the exploratory sample we did not find an effect of weight, in the confirmatory study we were able to confirm this small but reliable effect of wrist weight on voice peaks, suggesting that the vocalization peaks are related to the required forces (kinetics) to move the segment and not necessarily to the movement itself (i.e., to the muscle activity involved; Pouw et al., 2019b).

The results reported here will be further confirmed by analyzing the data regarding muscular and postural effects on the voice (Pouw et al., 2023). Such analyses will illustrate whether it is possible to make clear predictions about what type of upper limb movements have a certain effect on the voice that could then be functionally integrated with speech. To conclude, we hereby show that voice production is a dynamically open system that will be affected by other communicative actions such as hand gestures. This has deep implications for why gesture and speech are often produced in synchrony in humans in specific, and how evolution of the mammalian voice in general might be related to the whole-body movement system (Pouw & Fuchs, 2022). Voices do not operate in a vacuum, they are produced by bodily elements that can be synergistically recruited for moving one's body too.

References

Aliverti, A. (2016). The respiratory muscles during exercise. *Breathe*, *12*(2), 165-168.

Blumberg, M. S. (1992). Rodent ultrasonic short calls: locomotion, biomechanics, and communication. *Journal of Comparative Psychology*, *106*(4), 360-365.

Cooper, B. G., & Goller, F. (2004). Multimodal signals: enhancement and constraint of song motor patterns by visual display. *Science*, *303*(5657), 544-546.

Cordo, P. J., & Nashner, L. M. (1982). Properties of postural adjustments associated with rapid arm movements. *Journal of neurophysiology*, *47*(2), 287-302.

Lancaster, W. C., Henson Jr, O. W., & Keating, A. W. (1995). Respiratory muscle activity in relation to vocalization in flying bats. *Journal of Experimental Biology*, *198*(1), 175-191.

Lasserson, D., Mills, K., Arunachalam, R., Polkey, M., Moxham, J., & Kalra, L. (2006). Differences in motor activation of voluntary and reflex cough in humans. *Thorax*, *61*(8), 699-705.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W. Georg, M. & Grundmann, M. (2019). Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*.

Owoyele, B., Trujillo, J., De Melo, G., & Pouw, W. (2022). Masked-Piper: Masking personal identities in visual recordings while preserving multimodal information. *SoftwareX*, *20*, 101236.

Pearson, L., & Pouw, W. (2022). Gesture–vocal coupling in Karnatak music performance: A neuro–bodily distributed aesthetic entanglement. *Annals of the New York Academy of Sciences*, *1515*(1), 219-236.

Pouw, W., & Fuchs, S. (2022). Origins of vocal-entangled gesture. *Neuroscience & Biobehavioral Reviews*, *141*, 104836.

Pouw, W., Harrison, S. J., & Dixon, J. A. (2019a). Gesture–speech physics: The biomechanical basis for the emergence of gesture–speech synchrony. *Journal of Experimental Psychology: General*, *149*(2), 391.

Pouw, W., Harrison, S. J., Esteve-Gibert, N., & Dixon, J. A. (2019b). Energy flows in gesture-speech physics: exploratory findings and pre-registration of confirmatory analysis.

Pouw, W., de Jonge-Hoekstra, L., Harrison, S. J., Paxton, A., & Dixon, J. A. (2020). Gesture–speech physics in fluent speech and rhythmic upper limb movements. *Annals of the New York Academy of Sciences*, *1491*(1), 89-105.

Pouw, W., Werner, R., Burchardt, L. S., & Selen, L. (2023). The human voice aligns with whole-body kinetics. *bioRxiv*, 2023-11.

Rubin, H. J., LeCover, M., & Vennard, W. (1967). Vocal intensity, subglottic pressure and air flow relationships in singers. *Folia Phoniatrica et Logopaedica*, *19*(6), 393-413.

Seikel, J. A., Drumright, D. G., & Hudock, D. J. (2019). *Anatomy & physiology for speech, language, and hearing*. Plural Publishing, Incorporated.

Sundberg, J., Elliot, N., Gramming, P., & Nord, L. (1993a). Short-term variation of subglottal pressure for expressive purposes in singing and stage speech: a preliminary investigation. *Journal of Voice*, *7*(3), 227-234.

Sundberg, J., Titze, I., & Scherer, R. (1993b). Phonatory control in male singing: A study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source. *Journal of Voice*, *7*(1), 15-29.

Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech communication*, *57*, 209-232.

Werner, R., Selen, L., & Pouw, W. (2024). Arm movements increase acoustic markers of expiratory flow. *bioRxiv*, 2024-01.