Speech during physical activity: Effect on f0 and vocal intensity

Heather Weston,[1,2] Susanne Fuchs[1] & Amélie Rochet-Capellan[3]

[1]Leibniz-Centre General Linguistics (ZAS), Berlin, Germany
[2]Humboldt-Universität zu Berlin, Berlin, Germany
[3]Université Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France

Stress and the voice

A

- Physical activity is a type of physical stress
  STRESS: "the non-specific response of the body to any demand" **[1]**
- Many studies report an increase in fundamental frequency (f0) during stress, but most research concerns psychological stress **[2]**
- Less is known about physical stress, and methods vary widely **[3]**:
  - Speech tasks: counting 1 to 10, sustained vowels, reading a sentence
  - Physical tasks: moderate to extreme (e.g., running till exhaustion)
- Consequently, it is unclear to what extent findings generalize
- To facilitate comparison and extend ecological validity, we have created a corpus of multiple speech tasks and levels of activity.
- The first results presented here investigate mean f0 and vocal intensity.

→ How does physical activity affect f0 and vocal intensity?

Method

B

Participants

N = 12

- **female**
- **German native speakers**
- **age: 19–31y ($\bar{x}$ = 23)**

Conditions

control	light cycling	moderate cycling
no cycling (sitting still)	35% max. heartrate (HR)	65% max. heartrate

Calculating exercise intensity: Karvonen method

208 - (age*0.7) = age-predicted max HR

- resting HR = HR reserve

*exercise intensity + resting HR

Speech task

- 126-word passage: transcribed monologue
- "read as you speak"
- 3 trials per condition
- approx. 45s/trial

Vocal intensity

- calibrated mic (MiniSPL) (30cm from mouth)
- sound pressure level meter (Acoustilyzer; NTi Audio)
- equivalent continuous sound level (Leq) in dB per trial

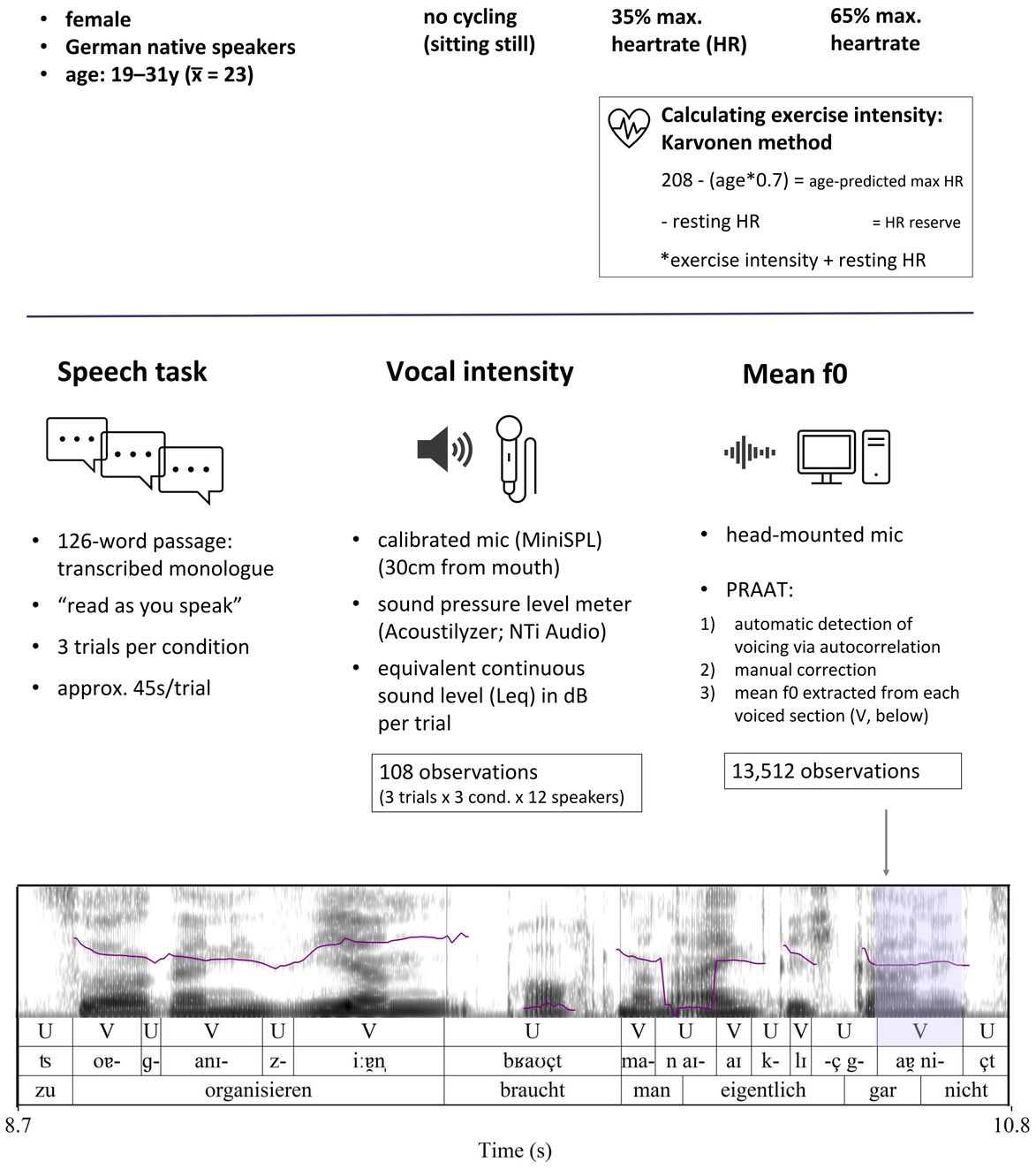
108 observations
(3 trials x 3 cond. x 12 speakers)

Mean f0

- head-mounted mic
- PRAAT:

1) automatic detection of voicing via autocorrelation
2) manual correction
3) mean f0 extracted from each voiced section (V, below)

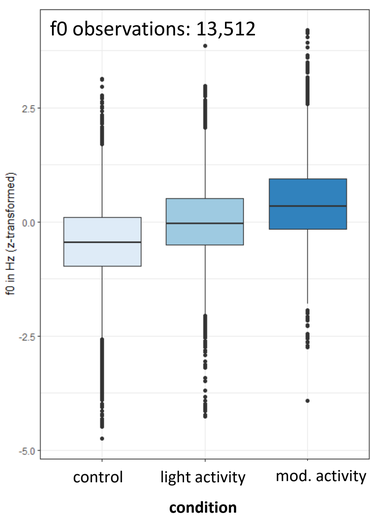
13,512 observations



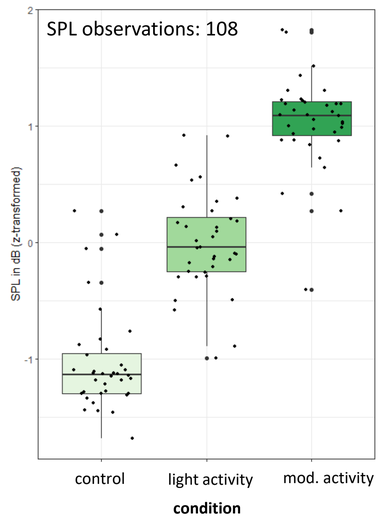
Group results (z-transfomed by speaker)

C

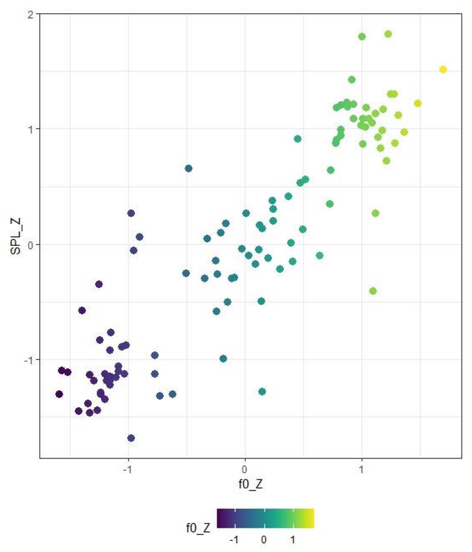
change in f0 (normalized)



change in SPL (Leq, normalized)



SPL and f0 trial mean (normalized)



f0: mean difference btw. conditions

		95% CI
light-con	**18.4 Hz**	(16.5–20.3)
mod-con	**33.4 Hz**	(31.5–35.3)
mod-light	**15.0 Hz**	(13.2–16.9)

Statistics: Linear mixed model on raw data (rand. var.: speaker); ANOVA; post hoc Tukey test showed significant differences between all conditions ($p < 0.001$)

SPL: mean difference btw. conditions

light-con	**2.8 dB**
mod-con	**4.2 dB**
mod-light	**2.3 dB**

Descriptive results: Differences between conditions calculated per speaker and averaged

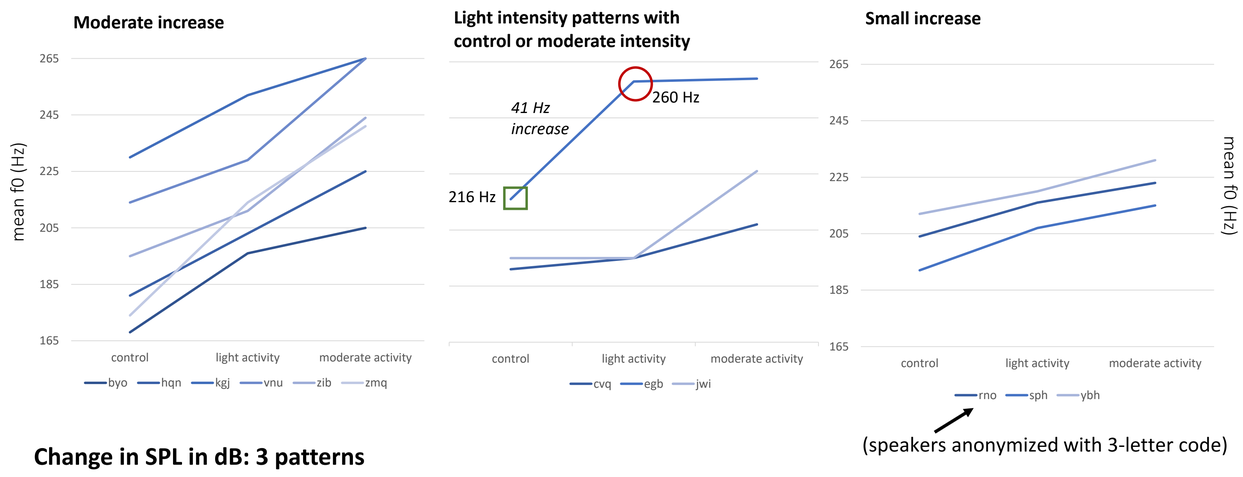
f0 plotted against SPL: descriptive results

Mean f0 was calculated for each speaker per trial per condition (9 trials) and plotted against SPL to give an indication of their correlation

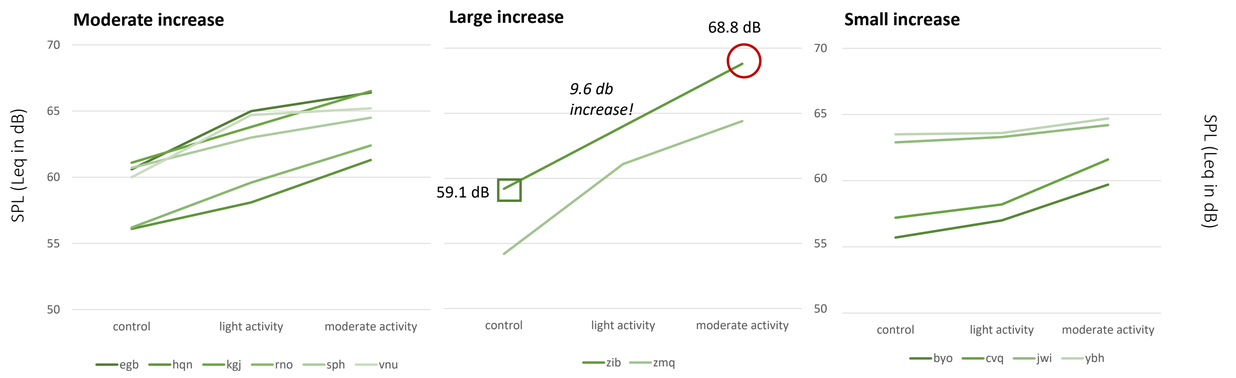
Individual results

D

Change in f0 in Hz: 3 patterns



Change in SPL in dB: 3 patterns



Discussion

E

Finding	Previous findings	What's next?
f0 and SPL increase with light and moderate activity	We found higher increases in f0 than previous studies: e.g., for moderate activity **[4]** report 19 Hz (vs. 33 Hz). But speech task was sustained phonation of /a/ (we used connected speech) → Is magnitude of effect task-related?	Analyze our data on sustained /a/: *Does our effect size decrease or stay the same?*
	[5] report an increase similar to ours (31 Hz), but speakers ran till exhaustion then read a passage – this exercise level was more demanding than ours. But mean f0 was calculated differently. → How much does mean-f0 method affect results?	Analyze our reading data with their method (mean f0 per utterance): *Significant difference between results obtained with each method?*
considerable speaker differences	Speaker differences are rarely addressed, but are considerable when reported; e.g., **[6]** found that 40% of speakers had no change or decreased f0 during activity. → Due to sample size? N=51 (vs. 14 in [4], 23 in [5])	Analyze other 36 speakers: *Does distribution of individual results (above) change?*

References **[1]** Selye H. 1974. *Stress without Distress*. Philadelphia, PA: Lippincott, 14. **[2]** Kirchhübel C, Howard DM & Stedmon AW. 2011. Acoustic correlates of speech when under stress: research, methods and future directions. Int J Speech Lang Law 18(1). **[3]** Van Puyvelde M, Neyt X, McGlone F & Pattyn N. 2018. Voice stress analysis: a new framework for voice and effort in human performance. Front Psych 9:1994. **[4]** Primov-Fever A, Lidor R, Meckel Y & Amir O. 2013. The effect of physical effort on voice characteristics. Folia Phoniatr Logo 65(6). 288-93. **[5]** Trouvain J & Truong KP. 2015. Prosodic characteristics of read speech before and after treadmill running. Interspeech-2015, Dresden, 3700-4. **[6]** Godin KW. 2009. *Classification based analysis of speech under physical task stress*. Dallas: U Texas. (MSc thesis.)

Acknowledgements

This research is jointly supported by the French National Research Agency (ANR) and the German Research Foundation (DFG) as part of the SALAMMBO project (http://salammbo-anr-dfg.ovh/).

The authors thank Jörg Dreyer and Theres Weißgerber for their support.