



HAL
open science

Exploring the source of short-term variations in respiratory data

Susanne Fuchs, Laura Koenig, Caterina Petrone

► **To cite this version:**

Susanne Fuchs, Laura Koenig, Caterina Petrone. Exploring the source of short-term variations in respiratory data. *Journal of the Acoustical Society of America*, 2019, 145 (1), pp.EL66-EL71. 10.1121/1.5087272 . hal-02091229

HAL Id: hal-02091229

<https://hal.science/hal-02091229>

Submitted on 5 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AUTHOR QUERY FORM

	<p>Journal: J. Acoust. Soc. Am.</p> <p>Article Number: 506901JAS</p>	<p>Please provide your responses and any corrections by annotating this PDF and uploading it according to the instructions provided in the proof notification email.</p>
---	--	---

Dear Author,

Below are the queries associated with your article; please answer all of these queries before sending the proof back to AIP. Please indicate the following:

Figures that are to appear as color online only (i.e., Figs. 1, 2, 3) __ (this is a free service).
 Figures that are to appear as color online and color in print (a fee of \$325 per figure will apply).

Article checklist: In order to ensure greater accuracy, please check the following and make all necessary corrections before returning your proof.

1. Is the title of your article accurate and spelled correctly?
2. Please check affiliations including spelling, completeness, and correct linking to authors.
3. Did you remember to include acknowledgment of funding, if required, and is it accurate?

Location in article	Query / Remark: click on the Q link to navigate to the appropriate spot in the proof. There, insert your comments as a PDF annotation.
AQ1	Please check that the author names are in the proper order and spelled correctly. Also, please ensure that each author's given and surnames have been correctly identified (given names are highlighted in red and surnames appear in blue).
AQ2	Please note journal style only permits one additional affiliation and it must be set as a footnote. Please check changes made.
AQ3	Please check change from "in the next section" to "in Sec. 1.1" to comply with journal style.
AQ4	The name of the manufacturer and their location must be provided for all equipment used in this paper.
AQ5	Please verify the definition of "SE."
AQ6	Ladefoged and McKinney (1963) is not cited in text. Please cite where appropriate or delete from the reference list.
AQ7	Please provide the date this site was last viewed.
AQ8	Please provide a digital object identifier (doi) for Ladefoged and Loeb (2002), Ladefoged <i>et al.</i> (1958). For additional information on doi's please select this link: http://www.doi.org/ . If a doi is not available, no other information is needed from you.

Thank you for your assistance.

Exploring the source of short-term variations in respiratory data

AQ1AQ2

Susanne Fuchs^{a)} and Laura L. Koenig^{b)}

Leibniz-Centre General Linguistics (ZAS), Berlin, 10117, Germany
 fuchs@leibniz-zas.de, koenig@haskins.yale.edu

Caterina Petrone

Aix-Marseille Université, CNRS, Laboratoire Parole et Langage, Aix-en-Provence, France
 caterina.petrone@lpl-aix.fr

Abstract: This study explores short-term respiratory volume changes in German oral and nasal stops and discusses to what extent these changes may be explained by laryngeal-oral coordination. It is expected that respiratory volumes decrease more rapidly when the glottis and the vocal tract are open after the release of voiceless aspirated stops. Two experiments were performed using Inductance Plethysmography and acoustics, varying consonantal properties, loudness, and prosodic focus. Results show consistent differences in respiratory slopes between voiceless vs voiced and nasal stops, which are more extreme in a loud or focused position. Thus, respiratory changes can even occur at a local level.

© 2019 Acoustical Society of America

[DDO'S]

Date Received: August 4, 2018 Date Accepted: December 29, 2018

1. Introduction

During speech production, chest wall displacements and subglottal pressures mainly show *slow, long-term variations* (e.g., Leanderson *et al.*, 1987), but *short-term changes* in respiratory signals, i.e., brief excursions from the long-term baselines, have also been observed. As elaborated in Sec. 1.1, these short-term changes have most typically been associated with word or sentence stress (Ladefoged, 1968; Ohala, 1990), but a few authors have noted the possibility of segmental effects on respiratory measures, particularly in cases where both the glottis and upper vocal tract are open, leading to rapid venting of air. Data reporting on such segmental effects have been sparse and anecdotal, however, and it is not clear whether such results can consistently be observed across multiple speakers. This paper systematically assesses respiratory displacement variation as a function of consonantal characteristics, loudness, and prosodic focus in several speakers.

1.1 Syllables, stress, and segments

Early reports of short-term excursions (or “pulses”) in respiratory system data came from Stetson (1951, originally published in 1928), who carried out a range of studies exploring respiratory control for speech. He collected various signal types, including electromyography (EMG), torso wall movements, and subglottal and esophageal pressures. Stetson proposed that individual syllables were associated with a “chest pulse” generated by the internal intercostal muscles.

Subsequent studies (e.g., Ladefoged *et al.*, 1958) challenged this chest pulse theory. To a greater extent than Stetson, Ladefoged and colleagues employed EMG methods along with measures of chest wall movements and esophageal pressures. These authors argued that the internal intercostals did not show ballistic activity for individual syllables. They did, however, report increases in subglottal pressure for stressed syllables (Ladefoged, 1968), and presented one figure quantifying the frequency of single motor unit firing in the internal intercostals to argue that such activity increased before stressed syllables (Ladefoged *et al.*, 1958).¹ Ohala (1990) subsequently summarized counterevidence to the claim that stressed syllables are associated with greater respiratory system activity, and suggested that lung volume changes correlate with syllables that have emphatic rather than lexical stress.

Ladefoged (1968) made an additional observation that has received considerably less attention, namely, subglottal pressure could show short-term decreases during

^{a)}Author to whom correspondence should be addressed.

^{b)}Also at: Haskins Laboratories, New Haven, Connecticut 06511, USA.

voiceless consonants. [Ohala \(1990\)](#) also acknowledged an interaction between the respiratory and oral systems and reported having observed rapid changes in lung volume during regions of high airflow for consonants.

Despite these suggestions in the literature, there has been little subsequent consideration of the interaction between specific segmental and prosodic properties and their effect on respiratory kinematics. To the extent that previous authors have explored this possibility, the data presentation has been anecdotal and qualitative, in that the authors did not consistently quantify respiratory changes as a function of consonant type. Moreover, the number of speakers investigated in this work, although not consistently documented, was likely quite limited.

In a recent study ([Petrone *et al.*, 2017](#)) we observed a regular rapid drop in respiratory volume after oral release (burst) in voiceless alveolar stops produced by women. However, that study did not compare voiceless stops to segments with a closed glottis in similar contextual and prosodic environments. From a purely logical perspective, it would make sense for segmentally-induced loss of air to affect respiratory system volumes. On the other hand, respiratory system volumes are quite large in comparison to whatever quantity of air might be released during an individual consonant such as an aspirated stop, particularly when one considers speakers with smaller glottal apertures (*viz.*, women and children, who have smaller laryngeal structures than the men who were mainly represented in the early work). Thus, this study represents a systematic exploration of whether multiple speakers, mostly adult females, consistently show segmental effects on respiratory system displacements. In particular, we evaluate respiratory volume during consonants varying in glottal and velar opening (nasal stops, voiced oral stops, voiceless aspirated oral stops). Along with consonant type, we also consider possible interactions with prosodic changes at sentence and word levels via loudness (first experiment) and focus (second experiment) manipulations. Prosodic variation is of interest given that prosody can be manifested in articulatory changes at a segmental level. For German speakers, a larger glottal opening for voiceless stops has been observed in the production of loud in comparison to normal speech ([Fuchs *et al.*, 2004](#)) and for focused in comparison to an unfocused position ([Hoole and Bombien, 2017](#)). It is not known, however, whether such prosodic effects on consonantal characteristics are reflected in respiratory patterns.

1.2 Hypotheses

We hypothesize that respiratory volume changes reflect laryngeal-oral coordination. Specifically, we expect that the slope of the respiratory volume declines more steeply at the release of voiceless aspirated stops than at the release of voiced and nasal stops. The steeper decline in voiceless stops should result from an open vocal tract after oral release which is coordinated with maximal glottal aperture. In voiced stops, phonetically realized as voiceless unaspirated or voiced word-initially in German, the glottis should be (almost) closed and less air can escape than with an open glottis, so that the slope of respiratory volume should not change to the same extent. Similarly, phonation in nasal stops should limit air loss and have only a marginal effect on respiratory volume changes.

We also expect effects to be larger in loud speech (experiment 1) and focused position (experiment 2) than in normal speech and unfocused position, because the degree of glottal opening for voiceless aspirated stops is larger in loud speech and under focus. Finally, we suppose that thoracic volume changes may be affected by loudness and focus, because the thorax is close to the larynx and the upper vocal tract and it has been discussed with respect to local respiratory changes ([Ladefoged and Loeb, 2002](#)). We do not expect consistent abdominal volume changes, since the abdomen is anatomically more distant from the larynx and upper vocal tract, and we generally associate it with slower motions ([Thomasson and Sundberg, 2001](#)).

2. Experiment 1

2.1 Methodology

Participants were 11 native speakers of German (all female) with an age between 20 and 37 yr and a body mass index between 18 and 23. All were recorded in a seated position. Thoracic and abdominal displacements (obtained using Inductive Plethysmography) were recorded simultaneously with speech acoustics (Sennheiser microphone HKH50 P48) using a multi-channel system that prevented the need for post-synchronization. The data were recorded with Edwin, software provided by the manufacturer. The data were then converted to MATLAB (version 2017b) for analysis.

The speech material consisted of bisyllabic target words containing initial /m b p/, a medial alveolar obstruent, and various vowels. Word with initial /m/ were: “Mieten”

AQ4

(rents), “Mitte” (section of Berlin), “Mate” (a tea), “Mützen” (caps), “München” (Munich). Words with /b/ were “Butter” (butter), “Büsten” (busts), “Büsum” (an island) and, with /p/, “Paddeln” (to canoe), “Pudel” (poodle), “Pita” (pita), “Pizza” (pizza), “Paten” (god-parents), “Pasta” (pasta), “Pute” (turkey), “Pudding” (pudding). Thus, phonetically, the vowels following /m/, /b/, and /p/ were, respectively, /i i y a:/, /u y/, and /i i u o a: a/, where /a:/ represents the tense vowel and /a/ is the lax counterpart, but both differ frequently only in vowel quantity.

The respective target words occurred sentence-initially. A question–answer paradigm was employed. For example, the experimenter asked the question: “Magst du X?” *Do you like X?* and the participant answered: “X mag ich, aber nicht Y.” *X I like, but not Y.* The participants supplied Y. The inclusion of Y made the experiment more engaging for participants, because they could partially create their own responses. Speakers produced utterances in normal and loud conditions. Louder speech was elicited by increasing speaker-experimenter distance.

Segments were labeled acoustically. For /m/ the onset was defined as the beginning of vocal fold oscillation. The offset was determined as the beginning of prominent formant structure corresponding to the beginning of the following vowel. For the oral stops the onset was defined as the first visible burst and the offset as the beginning of vocal fold oscillations for the following vowel. The slopes of thoracic and abdominal volume changes were calculated from the acoustically annotated onset (x1) to the offset (x2) of the segment (see Fig. 1 lower plots). Figure 1 shows the experimental setup and the annotation of the acoustic signals.

Thoracic and abdominal slopes were calculated using formula (1) where x denotes the on- and offset of the segment (see Fig. 1) and y1 and y2 are the respiratory signals obtained at times x1 and x2 from either the rib cage or the abdomen.

$$\text{Slope} = (y2 - y1)/(x2 - x1). \quad (1)$$

The units for the respiratory data are arbitrary, but are consistent within each speaker.²

2.2 Results

Linear mixed effect models [version R 3.4.3, R Core Team, 2018], *lme4* (Bates *et al.*, 2015), *lmerTest* (Kuznetsova *et al.*, 2017) were run with thoracic slope or abdominal slope as the dependent variable, loudness, phoneme, and their interaction as independent factors and speaker-specific slopes for loudness and phoneme. The reference level was set to /p/ and normal speech. For pairwise comparisons involving other levels than /p/ (/b/ vs /m/), the reference level was changed and only *p*-values smaller than 0.025 [*p* = 0.05 divided by the number of models (2) run] were treated as significant. Thoracic slope for /p/ was significantly steeper than in /m/ [$\beta = -0.316$, standard error

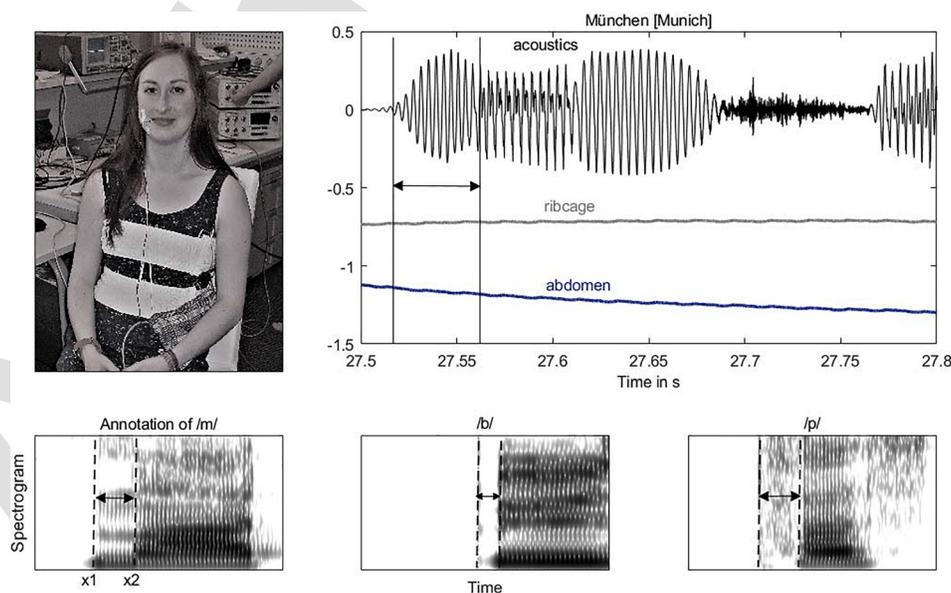


Fig. 1. (Color online) The upper plots show the experimental setup (participant sitting on a chair wearing the two respiratory belts) and the corresponding signals [acoustics (black line), thoracic volume changes (gray line), and abdominal volume changes (dark gray line below the ribcage signal) in arbitrary units]. The graphs below show the acoustic annotation which served as the input for obtaining the respiratory data during these regions.

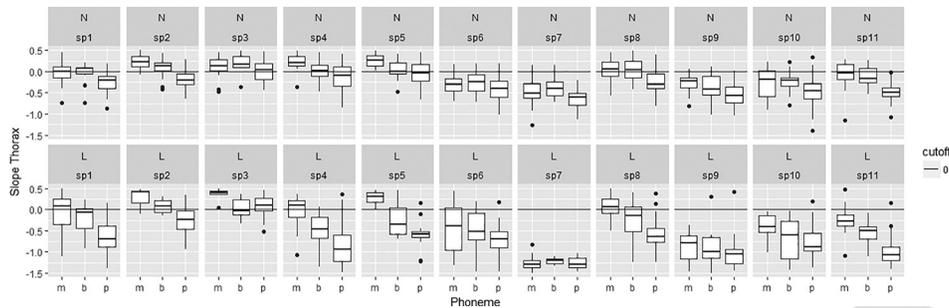


Fig. 2. Slope of the thoracic volume between the acoustically annotated on- and offsets (y -axis) split by phoneme (x axis) and loudness (upper panels: normal, lower panels: loud). Individual speakers' results (sp1–sp11) are displayed in the subplots. The horizontal line depicts a threshold. All values below this line refer to a decrease in thoracic volume (negative slope) while values above correspond to an increase in thoracic volume (positive slope).³ The number of all samples (n) is 1705.

AQ5 (SE)=0.051, $t=-6.16$, $p<0.001$] and also steeper in /p/ than /b/ ($\beta=-0.24$, SE=0.055, $t=-4.4$, $p<0.001$) (see Fig. 2). No significant differences were found between /b/ and /m/.

Loudness also revealed an effect with steeper slopes for loud speech than normal in /p/ ($\beta=-0.49$, SE=0.195, $t=-2.49$, $p=0.021$), but no significant differences in loudness were found for /b/ and /m/.

The abdominal slope revealed a difference between /p/ and /b/ with a shallower (less negative) slope for /b/ ($\beta=0.359$, SE=0.122, $t=2.95$, $p=0.005$), and a shallower slope in /b/ than /m/ ($\beta=0.391$, SE=0.128, $t=3.05$, $p=0.003$). No differences between /p/ and /m/ were found and there was no effect of loudness.

3. Experiment 2

3.1 Methodology

Seven women and three men, all native speakers of German, were recorded. Respiratory equipment and data annotation were the same as in experiment 1. Speakers had an age between 22 and 36 yr and body mass index between 19 and 25 (see Petrone *et al.*, 2017). The speech material consisted of sentences with contrastive focus, i.e., a word in a target utterance was contrasted with another word in the preceding context. The target sentences were elicited using a question–answer paradigm. The experimenter asked a question and the participant read the answer from a sheet of paper. Focused words were written in capital letters. For example, the question “Wäscht er Tiegel?” (Does he wash cups?) was used to prompt contrastive focus in the target utterance “*Er NIMMT Tiegel, aber wäscht sie nicht*” (He TAKES cups, but does not wash them). Prompts that put the final noun (e.g., “Tiegel”) in focus yielded the no-focus condition for the verbs. For analysis, we selected three verbs starting with a nasal or oral stop, i.e. /n/ in “nimmt” (takes), /m/ in “malt” (paints), and /k/ in “kennt” (knows); the vowels in the three words were /ɪ a: ɛ/, respectively. Note that the current analysis differs from the original study because here we assess the initial consonant of the verb whereas Petrone *et al.* (2017) investigated the /t/ at the end of the verb.

3.2 Results

Two linear mixed-effects models (using the packages *lme4*, *lmerTest*) were run with thoracic slope or abdominal slope as the dependent variable and focus, phoneme, and their interaction as independent factors. The random structure included speaker specific slopes for focus and phoneme. The voiceless stop /k/ in the focus condition served as the reference level and it was changed when comparing other phoneme pairs.

Statistical results showed that the thoracic slope was significantly shallower in /n/ than /k/ ($\beta=0.36$, SE=0.043, $t=8.35$, $p<0.001$), shallower for /m/ than /k/ ($\beta=0.34$, SE=0.043, $t=8.02$, $p<0.001$), and did not differ between /n/ and /m/ (see Fig. 3). Thoracic slope was affected by focus in /k/, with a shallower slope for the no focus than for the focus condition ($\beta=0.14$, SE=0.041, $t=3.36$, $p=0.0016$). Focus did not affect thoracic slope in /m/ and /n/, and individual variation is evident (Fig. 3).

Results revealed no effect of focus on abdominal slope, but an effect of phoneme with a shallower slope for /n/ than /k/ ($\beta=0.176$, SE=0.062, $t=2.85$, $p=0.0078$), but not for /k/ versus /m/. There were no differences between /n/ and /m/.

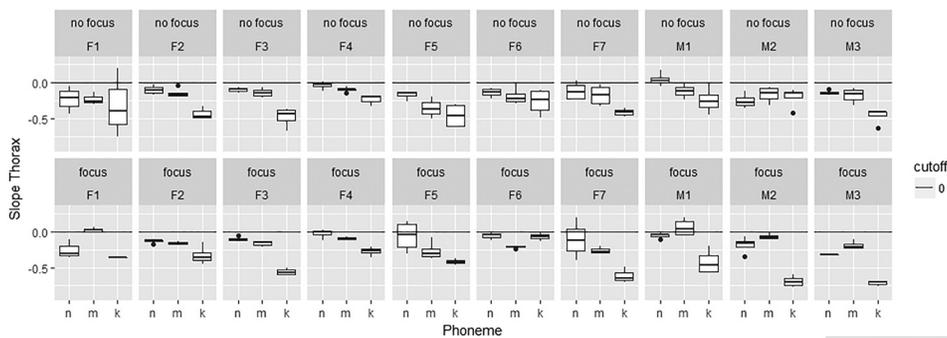


Fig. 3. Slope of the thoracic volume between the acoustically annotated on- and offsets (y -axis) split by consonant (x axis) and focus (upper track: no focus, lower track: focus). Individual speakers' results are displayed in the subplots (seven females, three males). All values below the black horizontal line at zero indicate a decrease in thoracic volume (negative slope) while values above correspond to an increase in thoracic volume (positive slope). The number of all samples (n) is 221.

4. Discussion and conclusion

Results for two experiments with different participants revealed consistent local effects on thoracic volume in voiceless aspirated consonants, which differ from voiced stops and nasals in their laryngeal-oral coordination. Specifically, a larger amount of air can escape when the vocal tract and the glottis are open in comparison to configurations where the glottis is closed or the vocal folds vibrate. Patterns for voiceless stops were rather consistent across speakers, with only a few speakers differing from the overall pattern (Fig. 2: sp7, loud condition; Fig. 3, M2 and F6). A review of the acoustic data for these few unusual speakers did not immediately show a reason for their atypical behaviour.

We also obtained evidence for prosodic effects (loudness, focus) in the slope of the thorax, but not the abdomen, in the vicinity of voiceless aspirated stops. The thorax is closer to the larynx and thoracic muscles may also be more flexible and adaptable to short temporal changes in comparison to the global abdominal motions (e.g., Ladefoged *et al.*, 1958). It may also be that abdominal movements reflect greater inertia than those of the thorax (cf. Thomasson and Sundberg, 2001). However, we note that kinematic data do not necessarily allow firm conclusions about underlying physiological processes. In our study the results could reflect an active involvement of the thorax muscles, mechanical properties of the thorax and abdomen, or simply a larger glottal aperture. A study adding measures of vocal-fold abduction or subglottal pressure could help disentangle the possibilities here.

Our results also demonstrate that prosody affects the degree to which differences in consonantal aerodynamics are reflected in respiratory data. The effects are not restricted to specific conditions, but are more extreme in loud speech and in words under focus. This is in line with literature on prosodic strengthening and in particular on laryngeal kinematics, where a larger glottal opening has been reported for strong prosodic conditions than in weak ones (Fuchs *et al.*, 2004; Hoole and Bombien, 2017).

Finally, we suggest that the short term negative excursions seen here may provide an alternative explanation for Stetson's chest pulses. That is, when one observes an undulating signal, one might focus on the local increases (pulses) rather than the local decreases ("valleys") as a phenomenon calling for explanation. Stetson's speech material consisted to a large extent of CV syllables with C being voiceless stops; hence it may not be surprising to find excursions on every syllable. Further, Stetson's experimental design may have led to rather careful or staccato speech, i.e., the syllables could also have received accentuation.

Acknowledgments

This work was supported by Grant No. 01UG1411 from the Ministry for Education and Research (BMBF) and the Leibniz Society to S.F. at ZAS and by financial aid (Bonus Qualité Recherche) from the Laboratoire Parole et Langage. We thank Jörg Dreyer for technical support and our speakers for their participation.

References and links

- ¹As pointed out by Ladefoged and Loeb (2002), the technology available at the time did not allow calculating averaged EMG data; the only cases where EMG data could be quantified was when a single motor unit was recorded.

- ²Since we did not merge thorax and abdomen kinematics to obtain overall lung volume, we cannot provide the slope values in percent vital capacity. 236
237
- ³It is possible that some cases of positive slopes reflect preparatory chest wall positioning (cf. Hixon *et al.*, 1988). However, we also obtained a few cases of positive slopes in experiment 2, where the analyzed words were not in utterance-initial position. 238
239
240
- Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). "Fitting linear mixed-effects models using lme4," *J. Stat. Software* 67(1), 1–48. 241
243
- Fuchs, S., Hoole, P., Pelorson, X., van Hirtum, A., Perrier, P., Dahlmeier, K., and Creutzburg, J. (2004). "Laryngeal adjustment in voiceless consonant production: I. An experimental study of glottal abduction in loud versus normal speech," in *Proceedings of the Conference on Vocal Fold Physiology and Biomechanics*, Marseille, France, pp. 1–5. 244
245
246
247
- Hixon, T. J., Watson, P. J., Harris, F. P., and Pearl, N. B. (1988). "Relative volume changes of the rib cage and abdomen during prephonatory chest wall posturing," *J. Voice* 2(1), 13–19. 248
249
- Hoole, P., and Bombien, L. (2017). "A cross-language study of laryngeal-oral coordination across varying prosodic and syllable-structure conditions," *J. Speech Lang. Hear. Res.* 60(3), 525–539. 250
251
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). "lmerTest Package: Tests in linear mixed effects models," *J. Stat. Software* 82(13), 1–26. 252
253
- Ladefoged, P. (1968). "Linguistic aspects of respiratory phenomena," in *Sound Production in Man*, edited by A. Bouhuys (New York Academy of Sciences, New York), pp. 141–151. 254
255
- Ladefoged, P., and Loeb, G. (2002). "Preliminary studies on respiratory activity in speech," *UCLA Work. Pap. Phon.* 101, 50–60. 256
257
- Ladefoged, P., and McKinney, N. (1963). "Loudness, sound pressure, and subglottal pressure in speech," *J. Acoust. Soc. Am.* 35, 454–460. 258
259
- Ladefoged, P., Draper, M. H., and Whitteridge, D. (1958). "Syllables and stress," *Le Maître Phonétique* 36, 1–14. 260
261
- Leanderson, R., von Euler, C., and Sundberg, J. (1987). "Breathing muscle activity and subglottal pressure dynamics in singing and speech," *J. Voice* 1, 258–262. 262
263
- Ohala, J. J. (1990). "Respiratory activity in speech," in *Speech Production and Speech Modelling*, edited by W. J. Hardcastle and A. Marchal (Kluwer, Dordrecht, The Netherlands), pp. 23–53. 264
265
- Petrone, C., Fuchs, S., and Koenig, L. L. (2017). "Relations among subglottal pressure, breathing, and acoustic parameters of sentence level prominence in German," *J. Acoust. Soc. Am.* 141, 1715–1725. 266
267
- R Core Team (2018). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org/>. 268
269
- Stetson, R. H. (1951). *Motor Phonetics: A Study of Speech Movements in Action* (North-Holland, Amsterdam, the Netherlands), 212 pp. 270
271
- Thomasson, M., and Sundberg, J. (2001). "Consistency of inhalatory breathing patterns in professional operatic singers," *J. Voice* 15, 373–383. 272
273
- AQ8
- AQ6
- AQ7