



HAL
open science

Relations among subglottal pressure, breathing and acoustic parameters of sentence-level 1 prominence in German 2

Caterina Petrone, Susanne Fuchs, Laura L. Koenig

► **To cite this version:**

Caterina Petrone, Susanne Fuchs, Laura L. Koenig. Relations among subglottal pressure, breathing and acoustic parameters of sentence-level 1 prominence in German 2. *Journal of the Acoustical Society of America*, 2017. halshs-01459350

HAL Id: halshs-01459350

<https://shs.hal.science/halshs-01459350>

Submitted on 7 Feb 2017

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.

1 **Abstract**

2 This study investigates whether acoustic correlates of prominence are related to actions of the
3 respiratory system resulting in local changes of subglottal pressure (P_{sub}). Simultaneous recordings
4 were made of acoustics; intraoral pressure (P_{io}), as an estimate of P_{sub} ; and thoracic and abdominal
5 volume changes. Ten German speakers read sentences containing a verb ending with /t/ followed by
6 a noun starting with /t/. These /t#t/ sequences were typically realized as one /t:/ with a long intraoral
7 pressure plateau. Sentence-level prominence was manipulated by shifting the position of contrastive
8 focus within the sentences. The slope and peak value of P_{io} within the /t#t/ sequence was used to
9 estimate differences in P_{sub} across focus positions. Results show that prominence production is
10 related to changes in the slope and maximum value of the pressure plateau. While pressure
11 increases led to higher intensity, they did not relate to f_0 , hence suggesting that local f_0 changes
12 primarily reflect laryngeal activity. Finally, strong individual differences were observed in the
13 respiratory data. These findings confirm past reports of local P_{sub} increases corresponding to
14 sentence-level prominence. Speaker-specific activations of the respiratory system are interpreted in
15 terms of motor equivalence, with laryngeal mechanisms also appearing to contribute to P_{sub}
16 changes.

17

18 200/200 words

19

1 **I. Introduction**

2 In this paper we investigate the physiological mechanisms behind the production of prosodic
3 prominence at the sentence level. Specifically, we test the hypothesis that acoustic correlates of
4 prominence (as induced by focus variations) are primarily related to actions of the respiratory
5 system leading to local changes of subglottal pressure. “Prominence” is a phonological notion,
6 which relates to the metrical alternation of “strong” (more prominent) and “weak” (less prominent)
7 elements within a prosodic unit. Prominence is also hierarchical: e.g., at the lexical level, word
8 stress specifies which syllable in a word is the strongest one; at the sentence level, only stressed
9 syllables are potential docking sites for phrase-level prosodic prominence and pitch accents. A
10 factor affecting prosodic variation at the sentence level is focus. In German, the language under
11 investigation here, words under focus are generally produced with a pitch accent associated with the
12 stressed syllable, which is defined as the most prominent element in the sentence (cf. Grice et al.,
13 2005). Shifts in focus position bring about a different placement of the pitch accents within the
14 sentence.

15 Phonetically, prominence in German is signaled by multiple parameters, such as fundamental
16 frequency (f_0), intensity, segmental duration, voice quality and spatiotemporal extent of articulatory
17 movements (e.g., Féry, 1993; Niebuhr, 2008; Mooshammer, 2010; Mücke and Grice, 2014).
18 Sentence-level prominence is primarily signaled by local and rapid f_0 changes (i.e., pitch accents)
19 on the stressed syllables. Local increases in the duration and intensity further contribute to
20 enhancing prominence (cf. Kügler and Gollrad, 2015 and references therein). In this paper, we are
21 particularly interested in f_0 and intensity, and their relationship with breathing. Cross-linguistically,
22 f_0 and intensity have been found to co-vary in many typologically different languages such as
23 English (e.g., Fry, 1955), French (Alain, 1993), and Swedish (Fant et al., 2000). A likely
24 explanation is that intensity and f_0 have shared physiological sources (e.g., Ladefoged, 1967). As
25 reviewed in the next sections, it is well established that respiration contributes to intensity and in

1 some cases to f0 variations, especially in long temporal windows (at the utterance level). The
2 question remains, however, whether respiratory maneuvers might also play a role in generating
3 intensity and frequency variations in short temporal windows (at level of prominent syllables). The
4 latter idea has a very long history, but the experimental evidence is sparse and controversial.

5

6 **A. Relations among breathing, subglottal pressure, intensity, and f0 in long temporal** 7 **windows**

8 **1. Intensity**

9 Variations in lung volume primarily occur over long temporal intervals (Ohala, 1990). One
10 breathing cycle in speech production is on average approximately 4 seconds (Rochet-Cappelan and
11 Fuchs, 2013). In parallel, during “neutral” (in contrast to emotive) speech, subglottal pressure
12 mainly shows long-term, gradual changes (Leanderson et al., 1986) and it is highly correlated with
13 sound pressure level, i.e. intensity. For example, Björklund and Sundberg (2016) measured
14 subglottal pressure indirectly via intraoral pressure in 31 participants repeating the syllable [pæ]
15 over a long time frame with increasing and decreasing loudness, but keeping f0 constant. Across
16 speakers, the average correlation coefficient between subglottal pressure and intensity was .83.
17 These results are similar to those of Lecuit and Demolin (1998), who investigated subglottal
18 pressure directly via tracheal puncture while two speakers produced sustained vowels with different
19 intensities. The results displayed a strong linear relationship between intensity and the logarithm of
20 subglottal pressure, independent of variations in f0. (Studies of singing provide additional evidence
21 for this relationship such as Herbst et al., 2015; however, since speech and singing may have some
22 differences in their underlying physiological control, our focus here will be on studies of speech.) In
23 general the results are coherent and reveal a tight relationship between intensity and subglottal
24 pressure (Draper et al., 1959; Ladefoged and McKinney, 1963; see also a historical review in
25 Ohala, 1990).

1 Modulation of laryngeal resistance, assessed via the amount of air flow passing the vocal folds for a
2 given subglottal pressure level, may also contribute to intensity variations (e.g., Finnegan et al.,
3 2000; Holmberg et al., 1988). Variation in laryngeal resistance is usually associated in large
4 measure to the thyroarytenoid muscle, which forms the muscular body of the vocal folds (cf.
5 Finnegan et al., 2000). The magnitude of resistance effects may vary with f_0 (e.g., Hirano et al.,
6 1969, Lecuit and Demolin, 1998). In a recent paper, Zhang (in press) used a computational model
7 of the pressure-volume-flow relationship. His results show convincingly that for durations typical
8 of normal speech (breath groups c. 4s long) air flow conservation is required. This can in principle
9 be done by means of increasing glottal resistance when approaching the end of the breath group or
10 using a larger inhalation at the beginning of the breath group.

11 **2. F0**

12 The relation between subglottal pressure and f_0 is less clear. According to myoelastic aerodynamic
13 theory, an increase in subglottal pressure will give rise to some degree of f_0 increase via increased
14 vocal-fold distension and the Bernoulli effect (van den Berg, 1958). Lieberman (1967) drew on this
15 relationship to explain the gradual decrease of f_0 over the course of an utterance in a long temporal
16 window, called f_0 declination. In particular, Lieberman (1967) found a correlation between f_0
17 declination and subglottal pressure in three English speaking subjects, and proposed that f_0
18 declination is a passive result of the subglottal pressure decrease. In contrast, several subsequent
19 authors argue that declination and other f_0 changes in speech involve more complex mechanisms;
20 specifically, it appears that subglottal pressure changes are not sufficient to account for the
21 magnitude of f_0 changes observed under various prosodic conditions. For example, Titze (1989)
22 estimated this relationship (RFP, i.e., the rate of f_0 change relative to the subglottal pressure
23 change) to be on the order of 2–6 Hz/cm H₂O. Other studies have similarly concluded that f_0
24 changes only slightly with differences in subglottal pressure (e.g., Gelfer et al. 1983; Collier, 1987;
25 cf. also Fuchs et al., 2015). This implies that although subglottal pressure and f_0 have a weak

1 physiological relationship, f_0 changes are primarily controlled by laryngeal muscle activity rather
2 than via the respiratory system. For f_0 increases, the cricothyroid muscle, which stretches the vocal
3 folds and increases longitudinal tension, is usually thought to be the main contributor. Even so,
4 Maeda (1976) noted rather wide variation in the estimated RFP across studies, and Strik and Boves
5 (1995) pointed out that measures of f_0 decrease over the course of an utterance depended strongly
6 on the methods used to estimate those changes. Thus, the magnitude of the relationship between f_0
7 and subglottal pressure remains a matter of some debate.

8

9 **B. Relations among breathing, subglottal pressure, intensity, and f_0 in short temporal** 10 **windows**

11 The idea that respiratory maneuvers might play a role in prosodic variations even within short
12 temporal windows is very old, but solid empirical support from numerous subjects is lacking.
13 Jespersen (1913) was one early work attributing stress variation to subglottal pressure. He defined
14 four levels of expiratory pressure (114ff.) for syllables, with four being the heaviest syllable to one
15 being the weakest, and zero with no expiratory pressure. In addition to respiration, Jespersen took
16 glottal resistance into account as a further mechanism to produce accent, and suggested that
17 accented syllables involve increased effort in all muscles. Stetson (1951) also considered
18 respiratory contributions to short-term aspects of speech production, namely syllables. His "chest-
19 pulse theory" proposed that the muscles between the ribs (intercostals) produced rapid movements
20 at the syllable level that "are like ripples on the wave of the expiratory movement of the breath
21 group" (Stetson, 1951: 2).

22 A series of studies carried out by Ladefoged and colleagues (Ladefoged and Loeb, 2002 and
23 references therein) failed to confirm the "chest pulses" proposed by Stetson for syllables in general.
24 These authors did, however, propose that local intensity increases in lexically stressed syllables can
25 be induced by short-term variation in the activity of the respiratory muscles (especially the internal

1 intercostalis muscles.) They found peaks of subglottal pressure and local activity in the internal
2 intercostal muscles immediately before each lexically stressed syllable for three speakers. Finnegan
3 et al. (2000) confirm the general conclusion that the respiratory system is capable of generating
4 rapid pressure changes. On the other hand, Ohala (1990), in a review of his own data, indicated that
5 he had observed local pressure peaks only during "emphatically" stressed syllables, suggesting that
6 the degree or level of prominence may be a relevant consideration.

7 In an investigation of sentence-level prominence, Gelfer et al. (1983) used recordings of respiratory
8 volumes, subglottal pressure via tracheal puncture, and electromyography of the cricothyroid
9 muscle in one speaker. Their results suggested a primary role for local increases in cricothyroid
10 activity in prominence production. Subglottal pressure changes spanned longer temporal windows
11 than the f_0 variations (see also Collier, 1987). In contrast, in another single-speaker study, Fant et
12 al. (1996) reported an increase of subglottal pressure in short temporal windows (accented
13 syllables), measured by tracheal puncture. However, they noted that the pressure changes seemed to
14 be more closely related in time to intensity than to f_0 changes. In a similar vein, Finnegan et al.
15 (2000) employed tracheal puncture and electromyographic recordings of the thyroarytenoid muscle
16 to study mechanisms of prominence variation in six English speakers. They found that sentence-
17 level prominence was associated both with changes in respiratory driving pressure and laryngeal
18 airway resistance, but respiratory driving pressure played the largest role. The results of this study
19 are stronger than the previous ones in that more speakers were involved; on the other hand the
20 authors did not evaluate f_0 or cricothyroid activity, limiting the degree to which the data speak to
21 physiological control of prominence.

22 A methodological challenge in all studies lies in the separation of laryngeal and respiratory effects
23 on sentence level prominence. Short-term variations in subglottal pressure (or "ripples", following
24 Stetson, 1951) could arise from laryngeal sources as well as respiratory ones. For example, glottal
25 opening behind a closed vocal tract may lead to a small drop in subglottal pressure, because the oral

1 cavity is added to the lower airways as an additional volume. Furthermore, when the glottis is open
2 and the vocal tract released, rapid venting of air through the mouth can yield a transient drop in
3 subglottal pressure. Thus, careful consideration of the speech material is needed to tease apart
4 different effects and derive valid conclusions.

5 One way of separating laryngeal and respiratory contributions would be simultaneous recordings of
6 subglottal pressure along with laryngeal muscle activity using electromyography (EMG). EMG
7 studies are increasingly rare given their invasive nature. Historical methods of assessing the
8 contributions of subglottal pressure to intensity and f_0 variation were likewise invasive for both
9 direct (tracheal puncture, e.g. Finnegan et al. 2000; catheter with pressure transducer, cf. Lecuit and
10 Demolin, 1998) and indirect (oesophageal balloon, Draper et al., 1959, Slifka, 2000) measurements
11 of subglottal pressure. In recent decades the most typical method has been to derive subglottal
12 pressure from intraoral pressure during stop consonants (Löfqvist et al., 1982; Hertegård et al.,
13 1995; Demolin et al., 1997; Björklund and Sundberg, 2016). Our study follows this method, using
14 intraoral pressure as an indirect estimate of subglottal pressure (see details below), combined with
15 non-invasively-obtained data on thoracic and abdominal movements.

16

17 **C. Speaker-specific behavior**

18 Several authors have reported speaker differences in prosodic control. For example, Adams and
19 Munro (1973) reported that one of their four speakers—who happened to be a trained singer and
20 public speaker—used longer periods of internal intercostal activity over the course of his utterances
21 than the others. Leanderson et al. (1987) reported that a few speakers maintained active
22 diaphragmatic contraction during phonation, and among those speakers the muscle activation
23 patterns varied as a function of pitch level. Maeda's (1976) two speakers differed in their use of
24 laryngeal contributions to focus (specifically, activation of the adductory lateral cricoarytenoid
25 muscle).

1 Ladefoged and Loeb (2002) point out in their conclusion that a variety of neuromotor combinations
2 can be used to control subglottal pressure. The same holds true for intensity and fundamental
3 frequency. Differing results across past studies lend support to this notion. Given the possibility of
4 motor equivalence, i.e. varying underlying actions leading to equivalent output, we may also expect
5 speaker-specific behavior in the underlying mechanisms that lead to differences in subglottal
6 pressure or in the relation between respiratory maneuvers and acoustic output. Together, the
7 possibility of speaker-specific behavior, combined with the cross-study variation evident in the
8 results reviewed above, underscore the need for continued study of breathing control for speech,
9 using methods that facilitate the use of multiple speakers and extensive datasets.

10

11 **D. Goals and hypotheses**

12 The goal of this paper is to use minimally invasive methods to quantify the relationships among
13 respiratory system activity, inferred subglottal pressure, and acoustic parameters of sentence-level
14 prominence in several speakers and in a language not yet studied via such methods (German).

15 Based on previous literature on German, we expect stressed syllables carrying sentence-level
16 prominence (as a result of focus) to be acoustically marked by higher f_0 , higher intensity, and
17 longer duration. If subglottal pressure (P_{sub}) is locally controlled, sentence-level prominence will
18 be accompanied by changes in subglottal pressure. P_{sub} will be inferred from intraoral pressure
19 measured during the oral closure plateau in stop consonants. We expect stops to be realized with an
20 intraoral pressure plateau whose attributes will change depending on the position of focus relative
21 to the surrounding vowel. If focus affects the vowel at the right side of the stop consonant, we
22 expect subglottal pressure to rise from the consonant to the second vowel and yield a slightly rising
23 pressure slope. If focus affects the vowel at the left side of the stop consonant, we expect a higher
24 pressure on the first vowel and a decrease in subglottal pressure from the vowel towards the

1 consonant, thus a negative pressure slope. Moreover, both left focus and right focus conditions
2 should show higher maximum intraoral pressure levels as compared to no focus.
3 If the respiratory system is involved in generating local subglottal variations, the slope of the rib
4 cage movement should be steeper under focus compared to the no focus condition as a mechanism
5 to increase subglottal pressure during focus production. We have no particular predictions for the
6 abdominal volume changes. On the one hand they could work independently of the rib cage, but on
7 the other hand, coordinated activity between the two has been reported by Ladefoged and Loeb
8 (2002) (but see Hixon and Weismer, 1995). Concerning the relationship among parameters, local
9 increases in subglottal pressure should lead acoustically to a higher intensity. F₀ and intensity might
10 be correlated weakly via automatic laryngeal mechanisms (van den Berg, 1958) or more strongly in
11 the sense that both are cues for focus. If local subglottal pressure changes are related to respiratory
12 kinematics, a correlation is expected between subglottal pressure values and respiratory slopes.
13 Finally, if the activity of the thorax and the abdomen are interdependent, the respective slopes will
14 correlate either positively (for coordinated activity) or negatively (for a trading-off relationship).

15

16 **II. Methodology**

17

18 **A. Corpus and participants**

19 The speech material consisted of five sets of German sentences containing a sequence of two target
20 words, viz., a monosyllabic verb ending with /t/ followed by a bisyllabic noun starting with /t/
21 (phonetically [t^h]). The noun was lexically stressed on the initial syllable, such as “Tanja” in the
22 sentence “Er malt Tanja, aber nicht Sonja” (*He paints Tanja, but not Sonja*). The target word sets
23 were “malt Tanja” (*paints Tanja*); “hat Tassen” (*has cups*); “kennt Tine” (*knows Tine*); “nimmt
24 Tiegel” (*takes pans*); “sieht Timmy” (*sees Timmy*). The /t#t/ sequence was chosen since German
25 speakers frequently realize it as one /t/, produced with a long closure duration (Fuchs et al., 2007)

1 which is comparable to the duration of a geminate. (In the current data, over 90% of productions
2 showed this pattern.) Whereas a single /t/ closure may not be long enough for intraoral pressure to
3 reach the level of subglottal pressure, especially in rapid or connected speech, the /t##t/ production
4 typically yields a long intraoral pressure plateau that allows us to infer Psub indirectly, as well as its
5 change over time (see section II.C).

6 An elicitation procedure was designed to obtain variation in focus position naturalistically using a
7 question-answer paradigm. A native speaker of German (the second author of the paper) read a
8 question intended to trigger a contrastive interpretation in the participant's answer. For example, if
9 the utterance "*Er nimmt Tiegel, aber wäscht sie nicht*" (He takes cups, but he does not wash them)
10 was prompted with the question "Wäscht er Tiegel?" (Does he wash cups?), the participant would
11 be led to emphasize the verb "nimmt" in the response (He *takes* cups, but does not wash them).
12 Three focus conditions were created: (1) focus on the leftmost word in the target sequence (e.g., the
13 verb "nimmt", "left focus" condition, **lf**); (2) focus on the rightmost word in the target sequence
14 (e.g., the noun "Tiegel", "right focus" condition, **rf**); and (3) and no focus on the first or second
15 target word ("no focus" condition, **nf**), but on the preceding personal pronoun ("Er"). We expected
16 a rising pitch accent to be realized on the stressed syllable of the focussed word (e.g., the first
17 syllable of "Tiegel"), which is typical of contrastive focus in German (Mücke and Grice, 2014). The
18 no focus condition served as a baseline, since early focus triggers post-focus deaccentuation in both
19 target words (e.g., Baumann, 2006).

20 Ten speakers were recorded: seven women and three men. They were between 22 and 36 years of
21 age and had normal Body Mass Index (between 19 and 25), and vital capacities within normal
22 limits. No participants reported any history of breathing or speech difficulties, and all spoke a
23 Northern variety of German. The speakers repeated the sentences four times in a randomized order
24 leading to 600 sentences (5 targets * 3 focus conditions * 4 repetitions * 10 speakers). After the
25 recording session participants were asked to do vital capacity manoeuvres (maximum inhalation

1 followed by maximum exhalation) 3–4 times with pauses in between. The maximum of the set was
2 taken as the participant’s vital capacity.

3

4 **B. Data recording and pre-processing**

5 Speakers sat in front of a music stand and read from a printed page where the words in focus were
6 marked in bold. They were asked to read without moving their arms and legs to avoid distortions of
7 the breathing signal. A super cardio condenser microphone (Sennheiser HKH50 P48) was
8 positioned about 30 cm from the mouth. The recording took approximately 15–20 minutes.
9 Acoustic, aerodynamic and breathing data were simultaneously recorded to a multi-channel data
10 recording system with a sampling frequency of 11025 Hz and imported into Matlab.

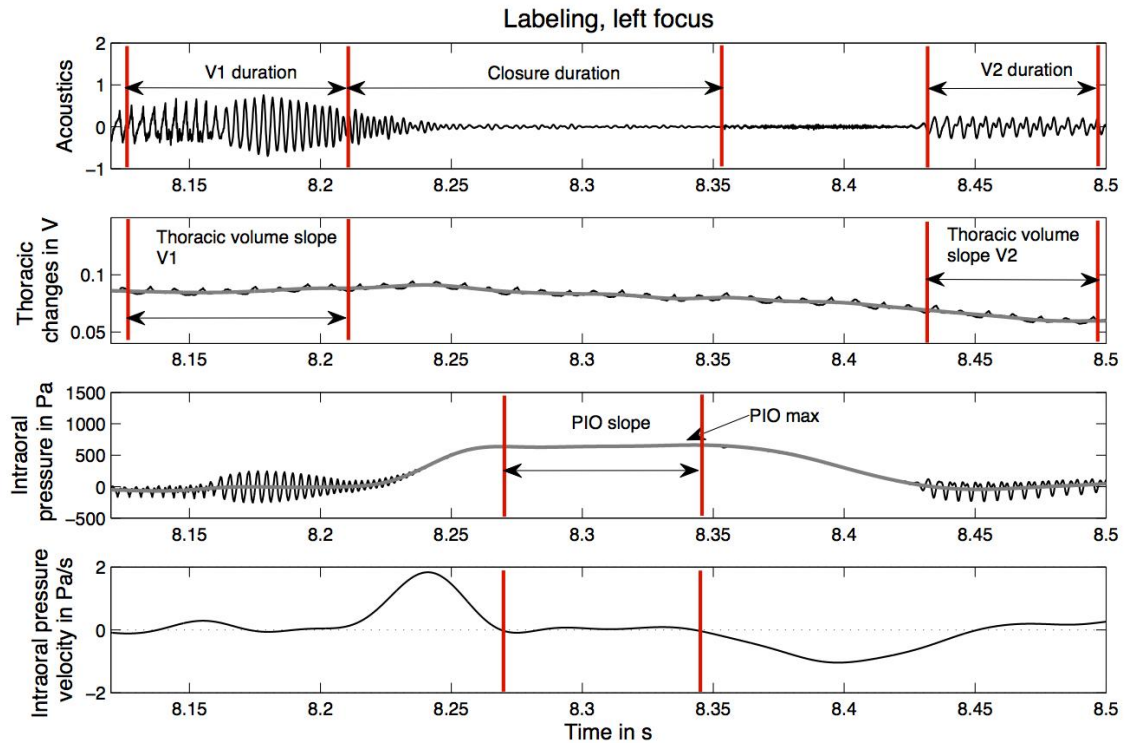
11 Intraoral pressure (Pio) data were obtained using a pressure transducer (Endevco 8507C-2) which
12 was glued onto the posterior end of the hard palate. This position placed the transducer posterior to
13 the alveolar closures in the target words. The pressure signal was calibrated after each recording
14 session using a water manometer. Pio data were smoothed using a zero-delay low-pass filter (*filtfilt*
15 function in Matlab) with 40 Hz passband and 100 Hz stopband edges, and a 50 dB damping factor.
16 This processing eliminated glottal pulses as well as effects of low-frequency electrical interference.
17 From the smoothed Pio signal, the first derivative (velocity) was calculated.

18 Respiratory movements (measured in Volts) were obtained via inductance plethysmography, using
19 two elastic bands, one around the rib-cage and one around the abdomen, to register volume changes
20 via changes in the electrical resistance of small wires located in the bands. Respiratory data were
21 downsampled by a factor of 10 and filtered using a 6th-order butterworth filter. The vital capacity
22 (VC) manoeuvres provided a measure of each speaker’s maximal compression and expansion of the
23 rib cage as well as of the abdomen, and allowed us to measure the rib cage and abdomen slope in
24 %VCs.

25 **C. Labelling and measurements**

1 Figure 1 illustrates the labelling in the acoustic, intraoral pressure and thoracic data. In the acoustic
2 signals, we labelled the onsets and offsets of the vowels preceding (V1) and following (V2) the
3 /t#/t/. In three cases (*malt*, *kennt*, *nimmt*), the vowel V1 was not immediately adjacent to the /t/
4 because of the presence of an intervening sonorant consonant, which was included in the
5 measurement of the vowel portion. V2 always followed the /t/. Intensity effects were measured as
6 the difference in mean intensity between the vowel following and preceding the /t#/t/ (**Int_diff**).
7 This relative intensity measurement minimized effects of slow changes in mouth-to-microphone
8 distance over the course of the recording session. (As indicated above, speakers were sitting and
9 explicitly asked to avoid major bodily movements during the recording to limit noise in the
10 respiration signal.) Analogously, relative measurements were taken for f0 and duration. For f0, we
11 calculated the mean difference between V2 and V1 (**f0_diff**); the presence of a pitch accent in the
12 focused word should be reflected in a higher f0 in the vowel compared to the unfocused one. For
13 duration, we extracted the mean difference between V2 and V1 (**Dur_diff**), as we expected the
14 vowel in the focused word to undergo lengthening compared to that of the unfocused word. Finally,
15 the closure duration (**Clos_dur**) was calculated as the difference between the onset of the
16 consonantal burst and the end of the preceding vowel.

17



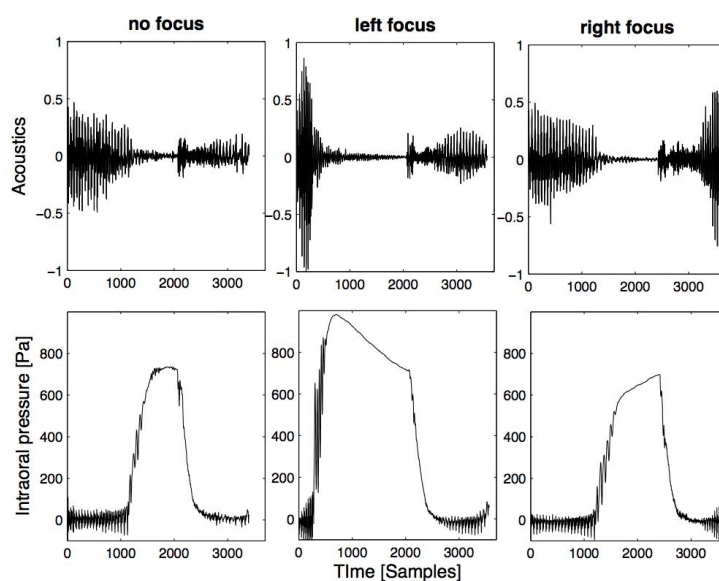
1

2 **Figure 1.** Example for labeling acoustic, respiratory and intraoral parameters for the target phrase
 3 “*Er nimmt Tiegel*” from the sentence “*Er nimmt Tiegel, aber wäscht sie nicht*” (“He takes cups, but
 4 does not wash them”). The phrase is produced with a left focus on the verb “*nimmt*” (“takes”). First
 5 track acoustic signal, second track thoracic volume changes, third and fourth tracks intraoral
 6 pressure and intraoral pressure velocity. Black lines are raw data, while grey lines superimposed on
 7 thoracic and intraoral pressure data represent filtered data. Vertical lines (colour online: red vertical
 8 lines) and horizontal arrows are interval boundaries. Pressure maximum (Pio_max) is indicated by
 9 an arrow.

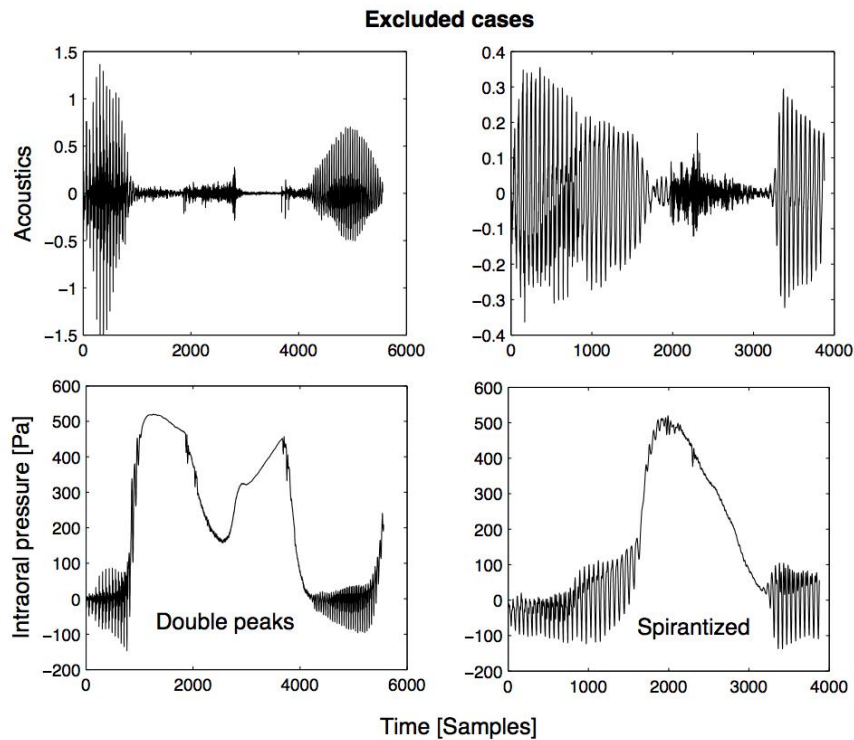
10

11 In the Pio data, we searched automatically for the pressure maximum (**Pio_max**) in the acoustically
 12 defined consonant closure. We also labelled the onset and offset of the pressure plateau in the /t#t/
 13 using zero crossings in the velocity signal, and measured the intraoral pressure slope (**Pio_slope**)
 14 between these time points (calculated as the difference in Pascals between the onset and offset
 15 divided by the duration of the plateau). The Pio slope measure used here follows past work

1 comparing intraoral and subglottal pressure and is based on the assumption that P_{io} quickly rises to
 2 the level of P_{sub} when the vocal tract is completely closed and the glottis is open as in voiceless
 3 aspirated stops. Data comparing P_{io} with directly-obtained subglottal or tracheal pressure shows a
 4 strong correlation between the two values, at least for driving pressures typical of speech (see
 5 Löfqvist et al., 1982; Kitajima and Fujita, 1990; Hertegård et al., 1995). In sequences of repeated
 6 equally-stressed syllables, P_{io} contours show a rather flat plateau, and authors use this value to infer
 7 the (stable) P_{sub} . However, P_{io} may also change during a voiceless stop closure interval if P_{sub} is
 8 changing (Löfqvist et al., 1982; Hertegård et al., 1995). In this study we drew on that observation to
 9 use the slope of P_{io} as an indication of changing P_{sub} depending on the focus position. Some
 10 examples from the data are shown in Figure 2. Data were excluded (7.7% of the database) when no
 11 clear plateau phase could be defined since speakers either realized two /t/s or they spirantized the
 12 stop (see Figure 3).



13
 14 **Figure 2.** Examples of tokens for the no focus, left focus and right focus conditions. Top panels:
 15 Acoustics. Bottom panels: Intraoral pressure. In these examples the slope of the pressure plateau
 16 during the oral stop appears different across the focus conditions (flat in nf, falling in lf and rising
 17 in rf).



2

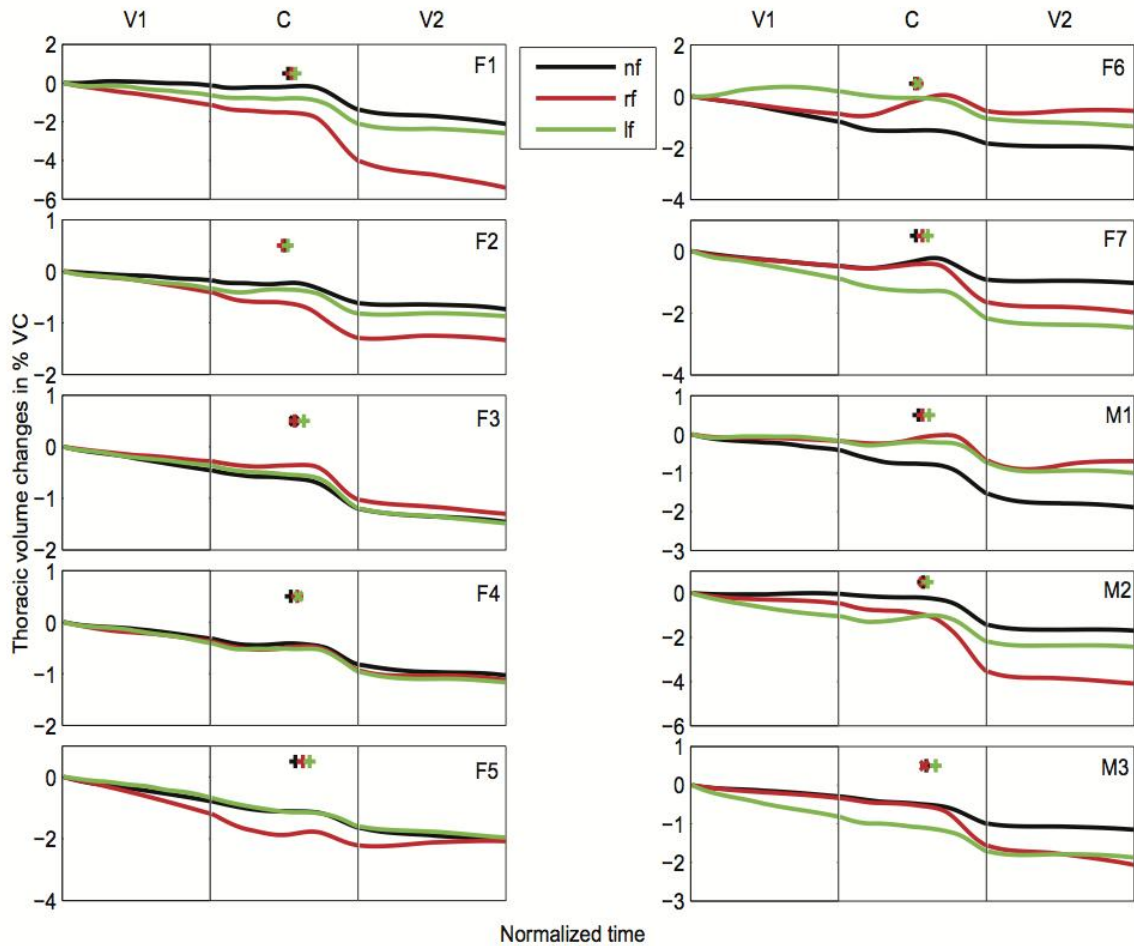
3 **Figure 3.** Examples of excluded tokens. Top panels: Acoustics. Bottom panels: Intraoral pressure.
 4 The left panels show a token where the speaker released the final /t/ in the verb. The right panels
 5 show a spirantized stop (notice the lack of a clear closure or burst, the noise in the waveform, and
 6 the slow discharge of P₀).

7

8 Concerning the respiratory contributions to focus, after visual inspections of the data (Figure 4), we
 9 decided to focus only on the first and second vowel and exclude the consonant portion. This is
 10 because we found a small, but very consistent drop in the thoracic movements occurring at the
 11 release burst. We interpreted the drop at the release burst as an automatic consequence of the
 12 increased airflow which coincides with maximal glottal aperture at the oral release (cf. Ohala, 1990
 13 for similar observations) and wanted to eliminate the impact of glottal opening on the respiratory
 14 signals. That is, in the vowels it is reasonable to assume a closed glottis, whereas during the

1 consonantal region we expect an abduction whose precise timing may vary across tokens and which
 2 may additionally perturb Pio data for reasons unrelated to respiratory actions.

3
 4



5

6 **Figure 4.** Time normalized averaged thoracic volume changes for each speaker (F = female, M =
 7 male), vowel1-consonant-vowel2 sequence (V1-Cons-V2) and focus condition (in gray scale) in
 8 percent vital capacity. The consonant is always a /t/. “+” signs indicate to the average burst location
 9 in each focus condition based on the acoustic measures. Time normalization was accomplished by
 10 obtaining 100 temporally equidistant data points for each target segment (V1, Cons, V2). All data
 11 were vertically aligned by subtracting out the first thoracic value in V1 so that all contours start
 12 with 0.

13

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Thus, to factor out such potential laryngeal effects, we measured thoracic and abdominal slope values between the on- and offset of V1 and V2 (in %VC) divided by V1 or V2 duration (in s) respectively. Analogously to the other measures we calculated the differences in slope of the respiratory signals between the first and the second vowel (**Thor_diff**, **Abd_diff**). For example, for focus on the left, we would expect a steeper slope for V1 (greater degree of thoracic compression) than V2 and consequently a negative value (V2 thoracic slope – V1 thoracic slope) of Thor_diff. For focus on the right vowel, we would expect a steeper slope for V2 and henceforth a positive Thor_diff.

D. Statistics

Linear mixed models were run with the statistical software R (R Development Core Team, 2016, version 3.3.0). For each model, the dependent variables were a) acoustics: Dur_diff, Clos_dur, Int_diff and f0_diff; b) intraoral pressure: Pio_slope, Pio_max; c) breathing: Thor_diff, Abd_diff. The pressure maximum was logarithmically transformed to achieve a normal distribution. The factor Focus (nf/lf/rf) was included as a fixed effect, with the condition no focus as the reference level. For f0_diff, sex was additionally included as a fixed factor to account for f0 differences across male and female speakers.

To study the relationships among acoustics (f0_diff, Int_diff) and aerodynamics (Pio_slope, Pio_max), four covariance models were run with f0 differences (f0_diff) depending on intensity differences (Int_diff), focus and sex; f0_diff depending on Pio_slope, focus and sex; intensity differences depending on Pio_slope and focus; and intensity differences depending on Pio_max and focus. Pio_slope was separately correlated to thoracic and abdomen slopes. Finally, a covariance model was run to evaluate the relationship between thoracic and abdominal volume changes. We started the statistical analysis by fitting each model with all possible random effect components

1 included (Barr et al., 2013). Since the full models showed some over-parameterization (e.g., when
2 the variance explained by a specific factor is close to zero), backward elimination based on
3 likelihood-ratio tests was used to decide which components should be retained in the models
4 (Pinheiro and Bates, 2000). Likelihood-ratio tests were run comparing full models (e.g., which
5 contained a random component) with simpler ones (e.g., without that component). As for p -values,
6 the standard functions in R to calculate linear mixed models do not provide them when the
7 dependent variables are continuous (e.g., duration differences between the target vowels). Hence,
8 we employed bootstrapping methods with replacement from the original sample (number of
9 samples = 500) to estimate p -values and 95% confidence intervals (CI) for each mixed model
10 (package R *lme4*, function `confint.merMod`, Wald method). The cut-off point for significance was
11 set at $p < 0.05$. Given that multiple parameters were collected from the same dataset and they might
12 not be independent, p -values were adjusted through the “false discovery rate” correction (Benjamini
13 and Yekutieli, 2001).

14

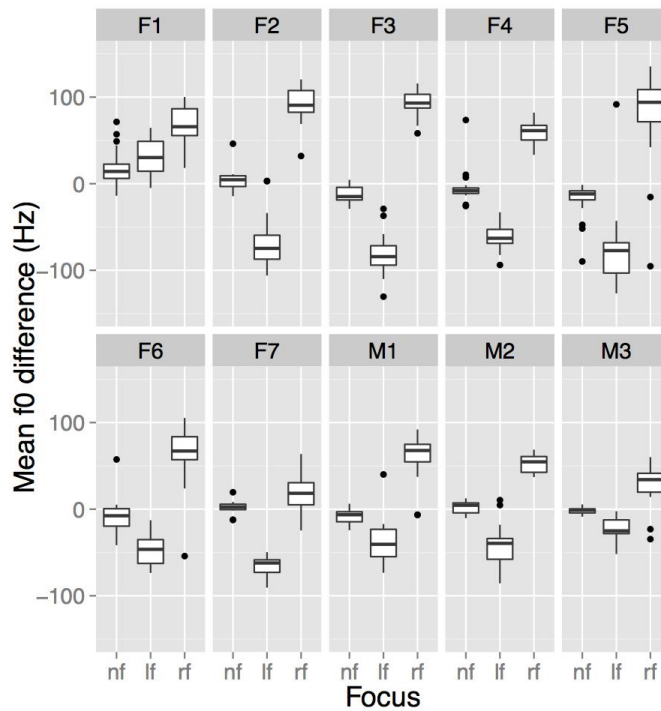
15 **III. Results**

16 **A. Acoustics**

17 Speakers realized contrastive focus successfully. When focus was present, f_0 differences became
18 larger than in the no focus condition (see Figure 5). When focus was placed on the left, f_0_diff was
19 larger with respect to the reference level [$\beta = -40.58$, $SE = 11.99$, $t = -3.38$, $p = 0.003$, $CI = -62.62$; -
20 15.36], with f_0 values being higher in V1 (217.5 Hz) than in V2 (162.2 Hz). f_0_diff was
21 significantly larger in the right focus than in no focus condition [$\beta = 51.21$, $SE = 12.08$, $t = 4.23$, p
22 $= 0.003$, $CI = 26.17$; 76.20], with f_0 values being higher in V2 (225.9 Hz) than in V1 (164.9 Hz).
23 Sex did not affect the f_0 differences significantly and showed no interaction with focus. Figure 5
24 shows that the f_0_diff patterns were consistent in 9 out of 10 speakers. At first glance, speaker F1
25 does not seem to show the expected f_0 patterns for focus. That is, for this speaker, Figure 5 shows

1 no difference between the target vowels in lf condition. However, a closer inspection of the acoustic
 2 data revealed that F1 realized focus by means of f0, but with a different f0 contour than the other
 3 speakers. This is because the left focus was characterized by high f0 plateau spanning the first to the
 4 second target word (corresponding to a continuation contour), which resulted in the lack of f0
 5 differences between V2 and V1 in this focus condition.

6



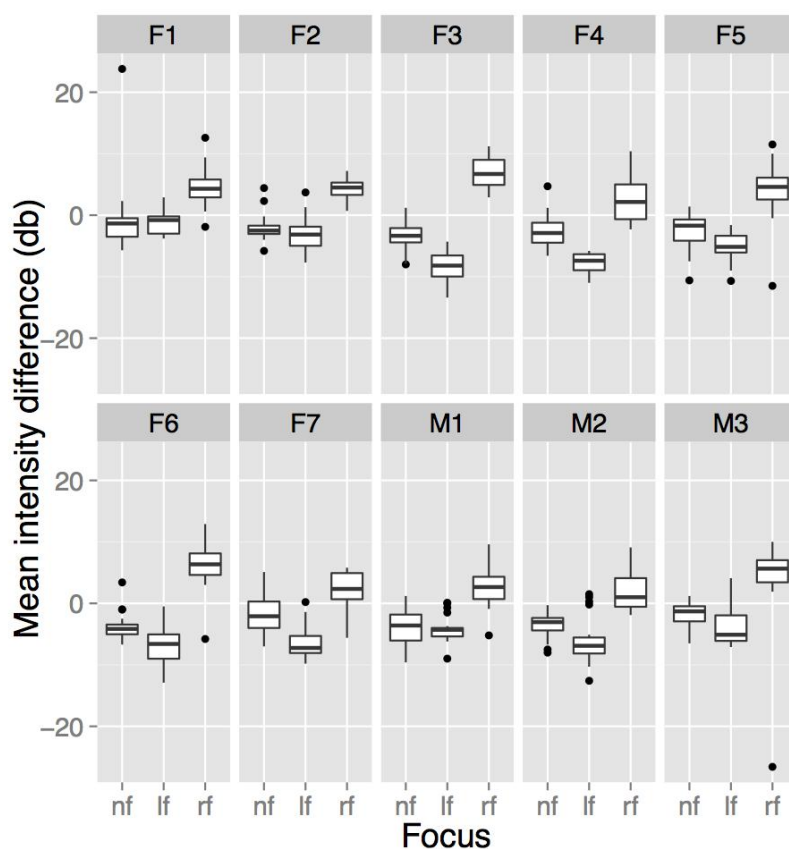
7

8 **Figure 5.** Boxplots for mean f0 differences between V2 and V1 (y-axis) against the three focus
 9 conditions (x-axis) and split by speakers (F = female, M = male). Values at zero indicate no
 10 difference between the two vowels. Positive values indicate that f0 is higher in V2 than in V1,
 11 negative values that it is higher in V1. Here and in all subsequent boxplots the median value is
 12 indicated as a horizontal bar, the boxes correspond to the 25th–75th percentile range, and the
 13 whiskers correspond to the ± 1.5 interquartile range.

14

15 Intensity differences between V2 and V1 behaved similarly to f0. When focus was present, intensity
 16 differences became larger relative to the no focus condition (see Figure 6). The direction of the

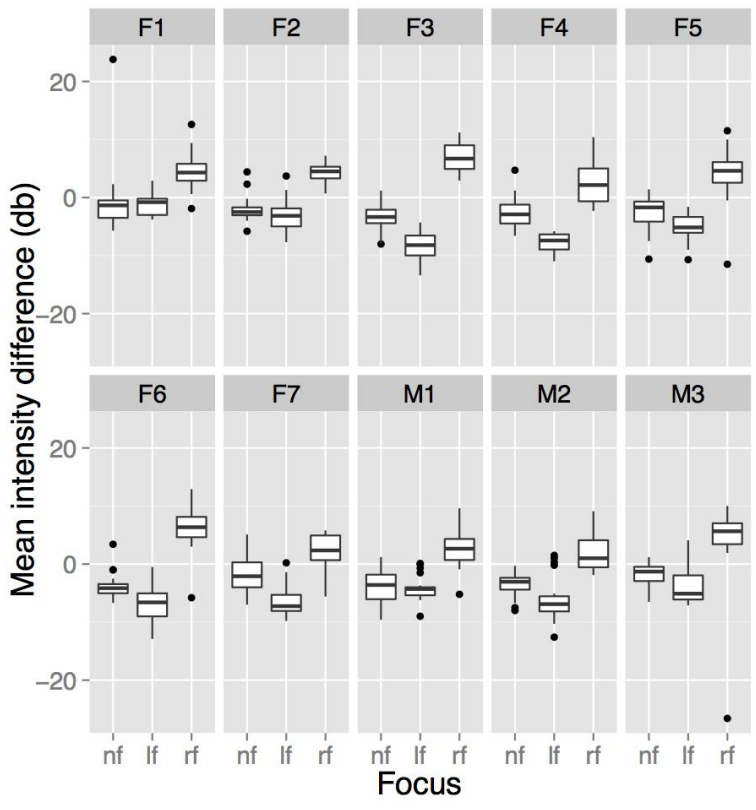
1 effect depended on the focus position. In left focus, the first vowel had higher intensity (82.3 dB)
 2 than the second one (76.7 dB). Hence, the intensity difference between V2 and V1 was negative
 3 and the effect was significant compared to the no focus condition [$\beta = -2.22$, SE = 0.82, $t = -2.69$, p
 4 = 0.011, CI = -3.91; -0.70]. In right focus, Int_diff was positive, with higher values in the second
 5 vowel (81.6 dB) than in the first one (77.7 dB). The comparison between right focus and no focus
 6 was also significant [$\beta = 5.80$, SE = 0.81, $t = 7.09$, $p = 0.003$, CI = 4.01; 7.44]. Figure 6 shows that
 7 the effect of intensity is rather consistent across speakers. F1 and M1 do not use intensity to
 8 distinguish no focus and left focus. Given a possible co-variance between f0 and intensity, the
 9 different intensity pattern for F1 could be related to the different f0 pattern employed by this
 10 speaker.



11
 12 **Figure 6.** Boxplots for mean intensity differences between V2 and V1 (y-axis) against the three
 13 focus conditions (x-axis) and split by speakers (F = female, M = male). Values at zero indicate no

1 difference between the two vowels. Positive values indicate that intensity is higher in V2 than in
 2 V1, negative values that it is higher in V1.

3
 4 As for vowel duration, we found a significant difference between V2 and V1 in left focus (V1= 139
 5 ms; V2 = 81.6 ms) compared to the no focus condition (V1= 102.5 ms; V2 = 80.09 ms) [$\beta = -32.19$,
 6 SE =14.49, $t = -2.2$, $p = 0.05$, CI = -59.27; -2.94]. This difference was consistent across speakers.
 7 The contrast between right focus and no focus condition was not significant, but Figure 7 reveals
 8 some speaker-specific variation for this contrast. Furthermore, closure duration was consistently
 9 longer under focus (nf = 86 ms; lf = 111 ms; rf = 99 ms) [for left focus: $\beta = 0.28$, SE = 0.02, $t =$
 10 10.1, $p = 0.003$, CI = 0.22; 0.34; for right focus: $\beta = 0.15$, SE = 0.02, $t = 5.57$, $p = 0.003$, CI = 0.10;
 11 0.20].



13
 14 **Figure 7.** Boxplots for mean duration differences between V2 and V1 (y-axis) against the three
 15 focus conditions (x-axis) and split by speakers (F = female, M = male). Values at zero indicate no

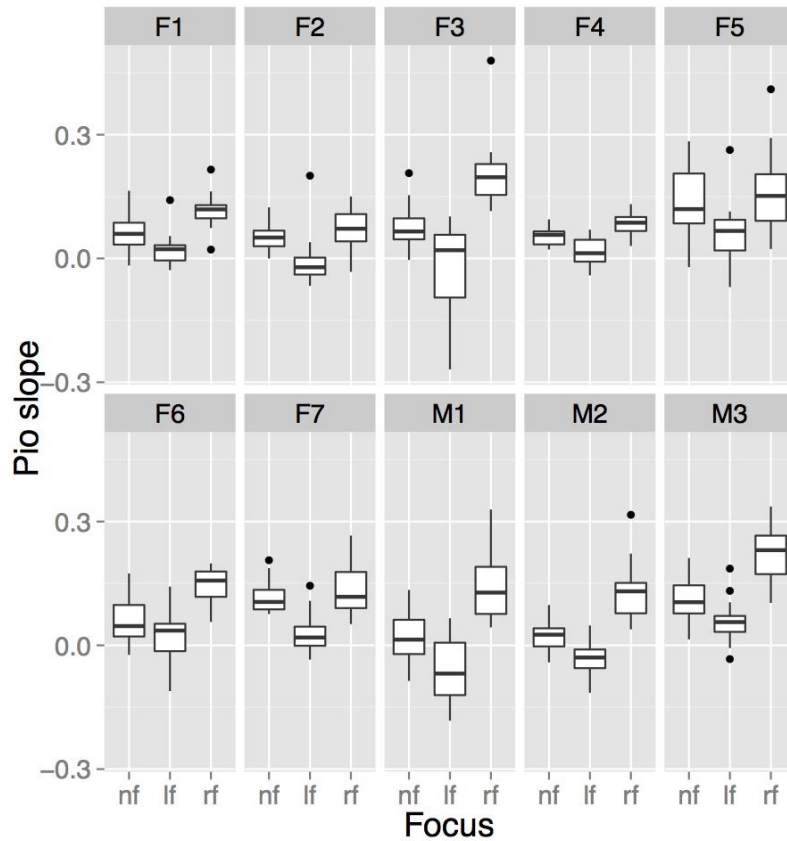
1 difference between the two vowels. Positive values indicate that duration is higher in V2 than in
2 V1, negative values that it is higher in V1.

3

4 **B. Subglottal pressure**

5 Results for intraoral pressure slope (Pio_slope) within the closure portion of the /t#/ provide
6 consistent evidence of subglottal pressure changes in the different focus conditions. As shown in
7 Figure 8, both the no focus (mean slope = 0.07) and the right focus (mean slope = 0.14) conditions
8 were characterized by a positive slope value. Hence, in those conditions, the intraoral pressure
9 plateau was rising from the beginning to the end of the closure, with a larger change in the focused
10 condition. The slope value in left focus was around zero or negative, indicating a flat or slightly
11 falling plateau (mean slope = 0.004). Patterns of Pio_slope were consistent across speakers. The
12 statistical analysis confirmed that the slope value in the no focus condition (the reference level in
13 our model) was significantly different from zero [$\beta = 0.073$, SE = 0.018, $t = 4.03$, $p = 0.003$, CI =
14 0.037; 0.108]. The slope was significantly steeper in right than in no focus [$\beta = 0.068$, SE = 0.015, t
15 = 4.49, $p = 0.003$, CI = 0.038; 0.094]. The slope in the left focus condition was significantly flatter
16 compared to the no focus condition [$\beta = -0.067$, SE = 0.015, $t = -4.4$, $p = 0.003$, CI = -0.09; -0.03].

17



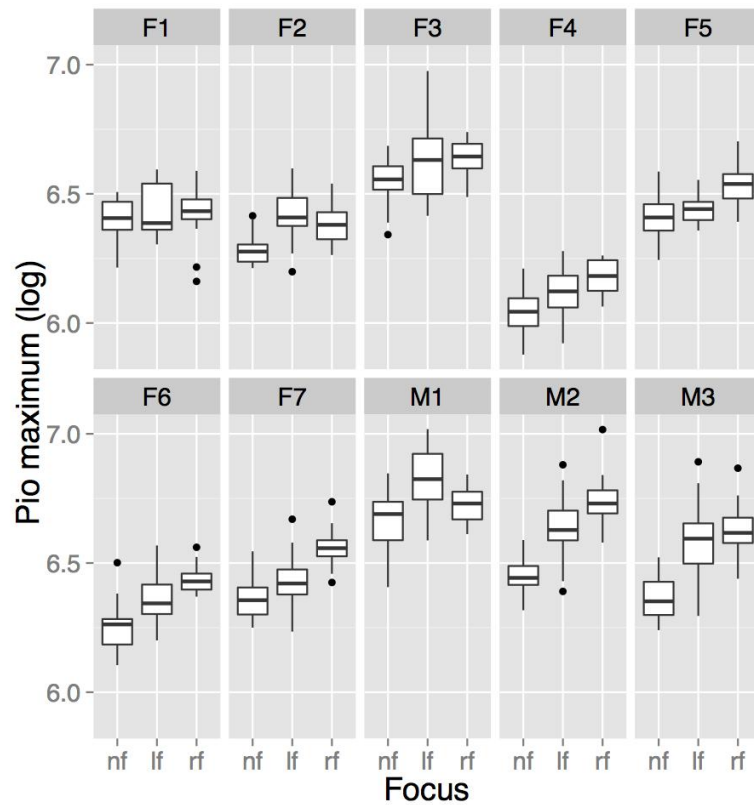
1

2 **Figure 8.** Boxplots for the slope of the intraoral pressure plateau (y-axis) against the three focus
 3 conditions (x-axis), split by speakers (F = female, M = male). Values at zero indicate that the
 4 plateau is flat. Positive values indicate that the Pio signal is rising, negative values that it is falling.

5

6 The involvement of Psub in focus production is also substantiated by the results from the pressure
 7 maximum (Pio_max) within the closure portion of the /t#t/ (Figure 9). The Pio_max value was
 8 higher under focus than with no focus (nf = 595 Pa; lf = 681 Pa; rf = 694 Pa). Compared to the no
 9 focus condition, the (logarithmically transformed) pressure maximum was significantly higher, both
 10 in the left focus [$\beta = 0.12$, SE = 0.02, $t = 4.27$, $p = 0.003$, CI = 0.07; 0.18] and in the right focus
 11 conditions [$\beta = 0.15$, SE = 0.02, $t = 5.19$, $p = 0.003$, CI = 0.09; 0.21]. Whereas the Pio_slope results
 12 were quite consistent across speakers, Figure 9 shows some differences in Pio_max across speakers.
 13 In 9 out of 10 speakers, Pio_max is higher under focus than in no focus. Five speakers (F4, F5, F6,
 14 F7, M2) have higher pressure maximum values in the right than in the left focus; F2 and M1 have

1 the reverse pattern, with higher values in the left focus; and F3 and M3 have similar values for left
 2 and right focus. F1 is the only speaker who shows no differences at all across the three focus
 3 conditions.

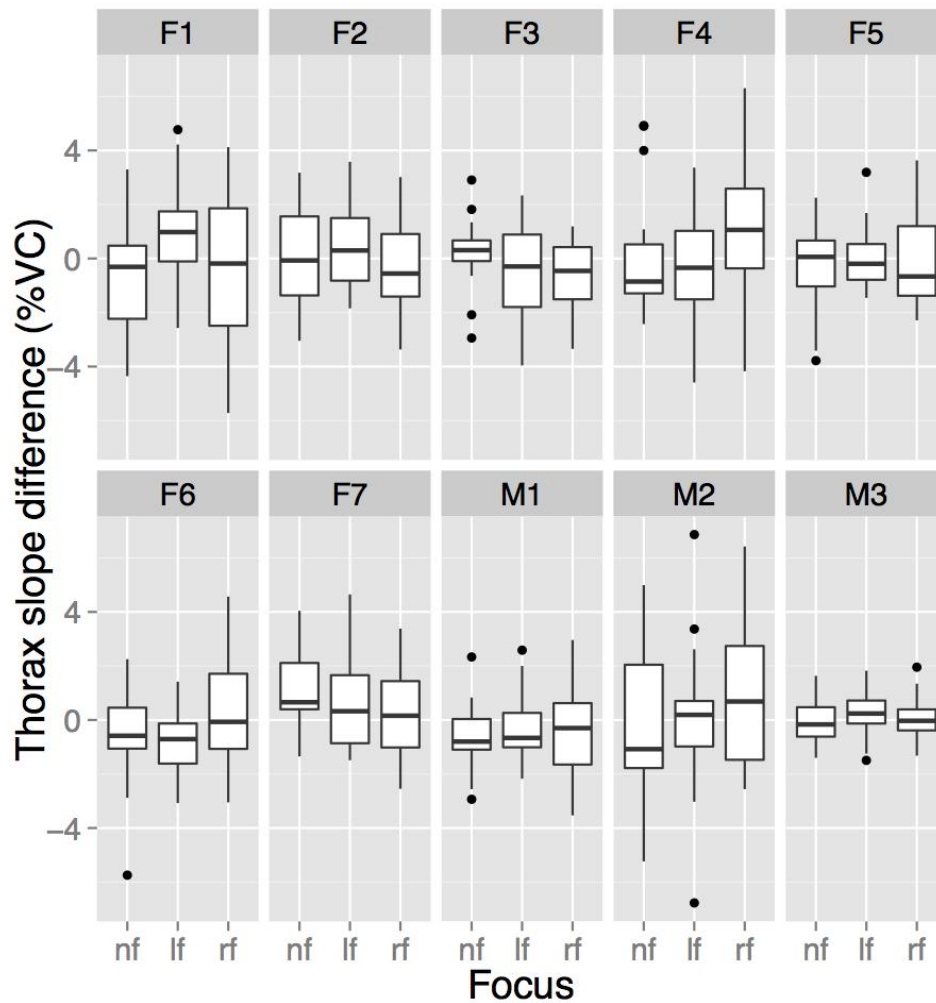


4
 5 **Figure 9.** Boxplots for the pressure maximum (log) of the intraoral pressure plateau (y-axis) against
 6 the three focus conditions (x-axis), split by speakers (F = female, M = male).

7
 8 **C. Respiratory kinematics**

9 The results for the respiratory kinematics are less clear than the results reported thus far. We found
 10 no effect of focus on the slope differences between V2 and V1, both for thoracic and abdominal
 11 volume changes. However, the kinematic data showed considerable speaker specific differences in
 12 volume changes. For instance, for the thorax (Figure 10), three speakers showed an increase of
 13 thorax slope when focus was on the right (F4, F6, M2); four speakers displayed a decrease in thorax
 14 slope (F2, F3, F5, F7), and three speakers had no effect of focus at all (F1, M1, M3). Results for
 15 abdominal changes are comparable.

1
2



3

4 **Figure 10.** Boxplots for the slope difference of rib cage movement in V2 and V1 (y-axis) against
5 the three focus conditions (x-axis), split by speakers (F = female, M = male). Values at zero
6 indicate that there is no difference in slope between V2 and V1. Positive values indicate that the
7 slope in V2 is steeper than in V1 and negative values that the slope in V1 is steeper than in V2.

8

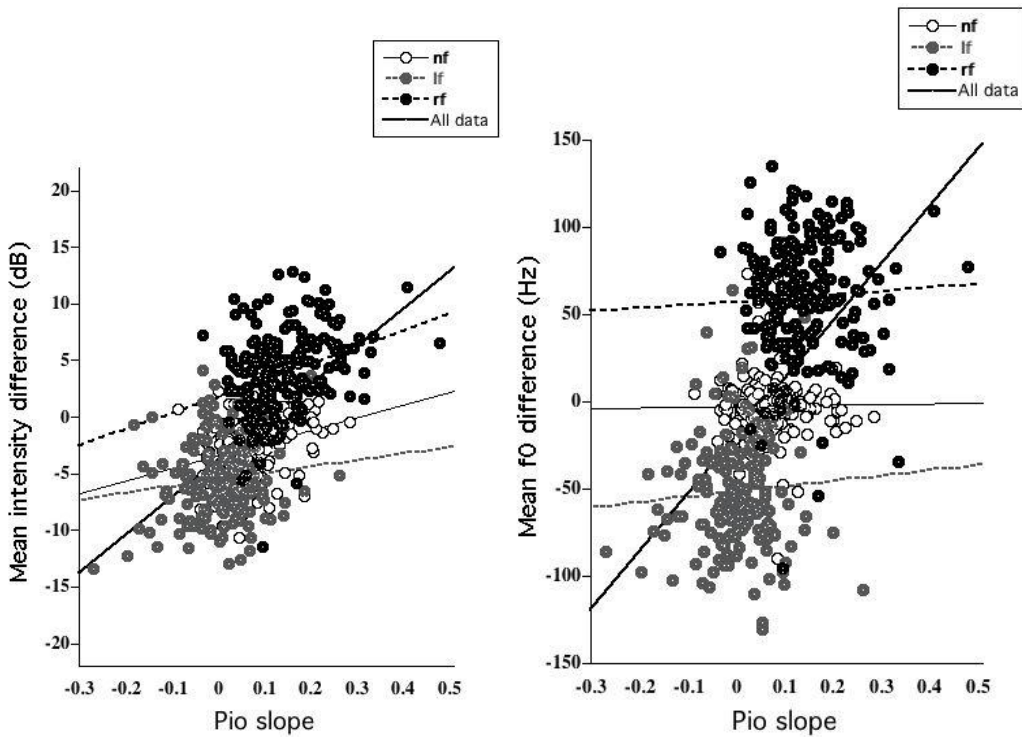
9 **D. Relations among parameters**

10 The V2–V1 differences in f0 and intensity were positively correlated with one another [$\beta = 1.3$, SE
11 = 0.45, $t = 2.93$, $p = 0.008$, CI = 0.42; 2.19]. There was no interaction between intensity and focus.

12 Intensity differences were also significantly correlated with Pio slope (Figure 11, left) [$\beta = 9.52$, SE

1 = 2.28, $t = 4.17$, $p = 0.003$, $CI = 5.24; 13.98$], whereas we found no correlation with Pio maximum.
2 We did not find any significant relation between f_0 differences and the Pio slope (Figure 11, right).
3 Pio slope was also independent of thoracic and abdominal volume changes. Finally, thoracic and
4 abdominal slope differences were positively correlated across all focus conditions [$\beta = 1.31$, $SE =$
5 0.06 , $t = 20.00$, $p = 0.003$, $CI = 1.19; 1.44$]. This is illustrated in Figure 12.

6
7
8
9

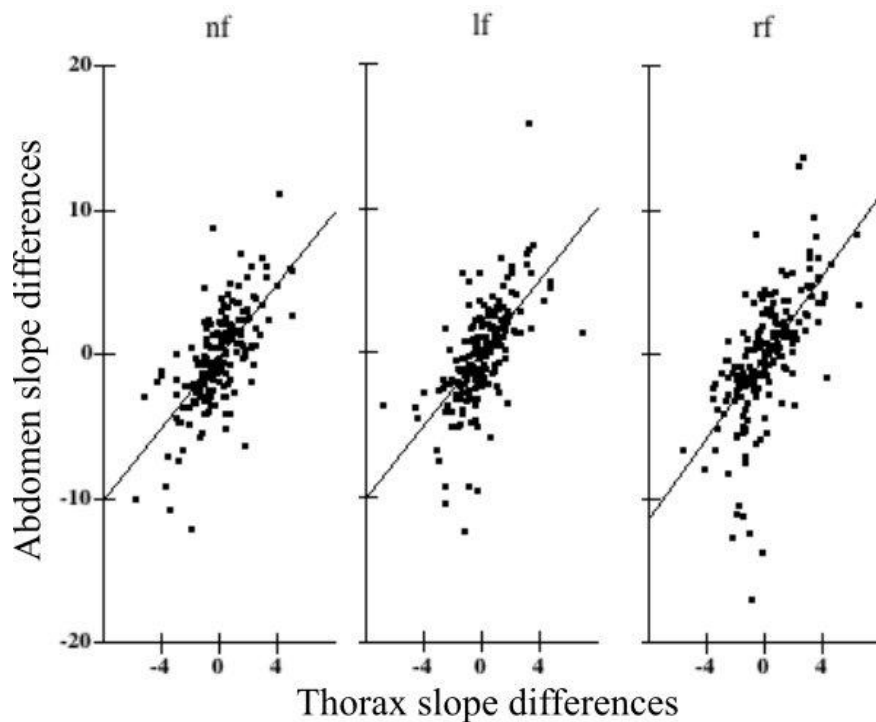


10

11

12 **Figure 11.** Scatterplots with the mean differences (dots) in intensity (y-axis) and intraoral pressure
13 slope (x-axis) between the V2 and V1 (left); and with mean differences (dots) in f_0 (y-axis) and
14 intraoral pressure slope (x-axis) between the V2 and V1 (right). Trendlines are superimposed.
15 Different shades of gray and line types indicate different focus conditions. A trendline for all data
16 (collapsing across focus conditions) was added.

17



1
2
3 **Figure 12.** Scatterplots with the mean differences (dots) in abdomen (y-axis) and thorax (x-axis)
4 slopes between V2 and V1 for the three focus conditions. Trendlines are superimposed.

5

6 **IV. Discussion and conclusion**

7 The acoustic measurements confirmed the correct elicitation of contrastive focus, by which stressed
8 syllables in the target words bear sentence-level prominence reflected in f_0 , intensity, and
9 durational differences. In line with the literature (e.g., Kügler and Gollrad, 2015), f_0 is the most
10 robust acoustic cue of sentence-level prominence in German, in that changes due to focus were
11 essentially consistent across speakers. At the aerodynamic level, focus production was related to
12 different intraoral pressure slopes in the voiceless alveolar stop. Furthermore, the maximum
13 pressure value was higher in the focus conditions than in the no focus one. These indirect estimates
14 of P_{sub} confirm the relationship between sentence-level prominence and subglottal pressure
15 suggested in the previous literature (e.g., Fant et al., 1996; Finnegan et al., 2000) and allow
16 generalization across a higher number of speakers. While the peak of the intraoral pressure is
17 correlated with the presence/absence of prominence, the pressure slope measurements suggest that

1 the Psub reaches a maximum within the focused word and decays afterwards. Moreover, we found
2 that different intraoral pressure slopes accompany differences in V2 and V1 intensity. This provides
3 evidence that prominent syllables are produced with local increases of Psub and that the slope of the
4 intraoral pressure contour is related to intensity differences between the two vowels surrounding the
5 consonant.

6 Also, we found a positive correlation between intensity and f0 differences between V2 and V1.
7 Given that prominent syllables carried a rising pitch accent, it is not surprising that higher f0 values
8 are related to higher intensity. However, this finding cannot be generalized to other intonation
9 contours, in which prominent syllables carry different pitch accents (e.g., low or falling accents).
10 On the other hand, the lack of relationship between f0 and intra-oral pressure speaks to their relative
11 independence in the production of prominence and supports the view that local f0 changes are
12 related to laryngeal mechanisms, such as changes in the activity of intrinsic laryngeal muscles (e.g.,
13 Hirano et al., 1969; Ohala, 1990).

14 Concerning the respiratory data, we evaluated separately the respiratory signal from the thorax and
15 the abdomen. The literature has suggested that muscles of the rib cage (internal intercostals) play a
16 major role in prominence production (Ladefoged and Loeb, 2002), and we hypothesized that the
17 thoracic signal might show reliable relationships with subglottal pressure and acoustic measures. By
18 recording both thoracic and abdominal volume changes, but analyzing them separately, we were
19 able to test both whether the thorax movements relate to local changes brought about by focus and
20 whether speakers use the thorax and abdomen in a trading relation. While we found a positive
21 relation between thorax and abdomen slopes (which speaks for a coordinated activity), there were
22 no consistent effects of prominence for either of the two respiratory signals. Graphical inspections
23 suggested, rather, a more complex scenario (Figure 4), where speakers adopted different strategies
24 across the three focus conditions. Thus, involvement of the respiratory muscles, and the division of
25 labor among thoracic and abdominal mechanisms, appears to be speaker-specific and not

1 mandatory in generating the local acoustic changes associated with focus. That is, some speakers
2 may realize sentence-level prominence via respiratory means and other speakers via other
3 mechanisms. For example, laryngeal lowering may lead to a slight increase in subglottal pressure,
4 since lung volume is decreased to a small extent. Differences in glottal resistance may also explain
5 speaker-specific patterns. This is in line with the idea that speakers can achieve the same acoustic
6 result with different articulatory strategies (“motor equivalence”).

7 One might ask whether the lack of consistent correspondences between the respiratory data and the
8 intraoral pressure and acoustic data speak to the adequacy of Inductance Plethysmography.
9 However, we could observe (Figure 4) a dip in the respiratory signals after the stop burst that was
10 consistent across speakers and focus conditions but small (at 1-2% thoracic VC level). This
11 observation implies that Inductance Plethysmography is sensitive to small changes in lung volume.

12 In sum, our findings provide evidence that sentence-level prominence is accompanied quite
13 consistently by local increases of subglottal pressure, especially as measured via P_{io_slope} , and
14 these correlate well with intensity but not with f_0 . They do not indicate a prominent and general
15 role for thoracic volume changes in generating local subglottal pressure changes or changes in f_0 ,
16 one of the main acoustic correlates of focus. It may be that emphatic stress would show clearer
17 involvement of respiratory mechanisms (cf. Ohala, 1990), but for sentential focus we conclude that
18 thoracic mechanisms do not play a decisive role for most speakers.

19

20 **Acknowledgements**

21 This work was supported by a grant (01UG1411) from the Ministry for Education and Research
22 (BMBF) to PB1 at ZAS and by a financial aid from the Laboratoire Parole et Langage. We thank
23 Jörg Dreyer for technical support and our speakers for their participation. This work is dedicated to
24 Giovanni Petrone and Dieter Fuchs.

25

1 **References**

- 2 Adams, C., and Munro, R. R. (1973). "The relationship between internal intercostal muscle activity
3 and pause placement in the connected utterance of native and non-native speakers of
4 English," *Phonetica* **28**, 227–250. <http://dx.doi.org/10.1159/000259457>
- 5 Alain, C. (1993). "The relation among fundamental frequency, intensity, and duration varies with
6 accentuation," *J. Acoust. Soc. Am* **94**(4), 2434–2436. <http://dx.doi.org/10.1121/1.407464>
- 7 Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). "Random effects structure for
8 confirmatory hypothesis testing: Keep it maximal," *J. Mem. Lang.* **68**, 255–278.
9 <http://dx.doi.org/10.1016/j.jml.2012.11.001>
- 10 Baumann, S. (2006). *The Intonation of Givenness. Evidence from German* (Niemeyer, Tübingen,
11 GE), 183 pp.
- 12 Benjamini, B., and Yekutieli, D. (2001). "The control of the false discovery rate in multiple testing
13 under dependency," *Ann. Statistics* **29**, 1165–1188.
14 <http://dx.doi.org/10.1214/aos/1013699998>
- 15 Björklund, S., and Sundberg, J. (2016). "Relationship between subglottal pressure and sound
16 pressure level in untrained voices," *J. Voice* **30**, 15–20.
17 <http://dx.doi.org/10.1016/j.jvoice.2015.03.006>
- 18 Collier, R. (1987). "F0 declination: The control of its setting, resetting and slope," in *Laryngeal*
19 *function in phonation and respiration*, edited by T. Baer, C. Sasaki and K. S. Harris
20 (College-Hill Press, Boston, MA), pp. 403–421.
- 21 Demolin, D., Giovanni, A., Hassid, S., Heim, C., Lecuit, V., and Soquet, A. (1997). "Direct and
22 indirect measurements of subglottic pressure," in *Proceedings of LARYNX'97*, Marseille,
23 France, pp. 69–72. http://www.isca-speech.org/archive_open/larynx_97.
- 24 Draper, M. H., Ladefoged, P., and Whitteridge, D. (1959). "Respiratory muscles in speech," *J.*
25 *Speech Hear. Res.* **2**(1), 16–27. <http://dx.doi.org/10.1044/jshr.0201.16>

- 1 Fant, G., Hertegård, S., and Kruckenberg, A. (1996). "Focal accent and subglottal pressure," *Quart.*
2 *Prog. Stat. Rep.* **37**(2), 29–32.
- 3 Fant, G., Kruckenberg, A., Liljencrants, J., and Hertegård, S. (2000). "Acoustic phonetic studies of
4 prominence in Swedish," *TMH-QPSR* **41**(2–3), 1–52.
- 5 Féry, C. (1993). *German Intonational Patterns* (Niemeyer, Tübingen, GE), 194 pp.
- 6 Finnegan, E. M., Luschei, E. S., and Hoffman, H. T. (2000). "Modulations in respiratory and
7 laryngeal activity associated with changes in vocal intensity during speech," *J. Speech Hear.*
8 *Res.* **43**, 934–950. <http://dx.doi.org/10.1044/jslhr.4304.934>
- 9 Fry, D. B. (1955). "Duration and intensity as physical correlates of linguistic stress," *J. Acoust. Soc.*
10 *Am* **27**, 765–768. <http://dx.doi.org/10.1121/1.1908022>
- 11 Fuchs, S., Koenig, L. L., and Winkler, R. (2007). "Weak clicks in German?," in *Proceedings of the*
12 *16th International Congress of Phonetic Sciences*; Saarbrücken, Germany, pp. 449–452.
- 13 Fuchs, S., Petrone, C., Rochet-Capellan, A., Reichel, U. D., and Koenig, L. L. (2015). "Assessing
14 respiratory contributions to f₀ declination in German across varying speech tasks and
15 respiratory demands," *J. Phon.* **52**, 35–45. <http://dx.doi.org/10.1016/j.wocn.2015.04.002>
- 16 Gelfer, C. E., Harris, K. S., Collier, R., and Baer, T. (1983). "Is declination actively controlled?," in
17 *Vocal fold physiology: Biomechanics, Acoustics and Phonatory Control*, edited by I. Titze
18 and R. Scherer (Denver Center for the Performing Arts, Denver, CO), pp. 113–126.
- 19 Grice, M., Baumann, S., and Benzmüller, R. (2005). "German intonation in autosegmental-metrical
20 phonology," in *Prosodic Typology: The Phonology of Intonation and Phrasing*, edited by
21 S.-A. Jun (Oxford University Press, Oxford, UK), pp. 55–83.
- 22 Herbst, C. T., Hess, M., Müller, F., Svec, J. G., and Sundberg, J. (2015). "Glottal adduction and
23 subglottal pressure in singing," *J. Voice* **29**, 391–402.
24 <http://dx.doi.org/10.1016/j.jvoice.2014.08.009>

- 1 Hertegård, S., Gauffin, J., and Lindestad, P.-Å. (1995). "A comparison of subglottal and intraoral
2 pressure measurements during phonation," *J. Voice* **9**, 149–155.
3 [http://dx.doi.org/10.1016/S0892-1997\(05\)80248-6](http://dx.doi.org/10.1016/S0892-1997(05)80248-6)
- 4 Hirano, M., Ohala, J., and Vennard, W. (1969). "The function of laryngeal muscles in regulating
5 fundamental frequency and intensity of phonation," *J. Speech Hear. Res.* **12**, 616–628.
6 <http://dx.doi.org/10.1044/jshr.1203.616>
- 7 Hixon, T. J., and Weismer, G. (1995). "Perspectives on the Edinburgh study of speech breathing," *J.*
8 *Speech Hear. Res.* **38**, 42–60. <http://dx.doi.org/10.1044/jshr.3801.42>
- 9 Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (1988). "Glottal airflow and transglottal air
10 pressure measurements for male and female speakers in soft, normal, and loud voice," *J.*
11 *Acoust. Soc. Am* **84**(2), 511–529. <http://dx.doi.org/10.1121/1.396829>
- 12 Jespersen, O. (1913). *Lehrbuch der Phonetik* [The Handbook of Phonetics] (2nd ed.) (Teubner,
13 Leipzig, GE), 276 pp.
- 14 Kitajima, K., and Fujita, F. (1990). "Estimation of subglottal pressure with intraoral pressure," *Acta*
15 *oto-laryng.* **109**(5–6), 473–478. <http://dx.doi.org/10.3109/00016489009125172>
- 16 Kügler, F. and Gollrad, A. (2015). "Production and perception of contrast: The case of the fall-rise
17 contour in German," *Frontiers in Psychology* **6**, 1254.
18 <http://dx.doi.org/10.3389/fpsyg.2015.01254>
- 19 Ladefoged, P. (1967). *Three Areas of Experimental Phonetics* (Oxford University Press, London,
20 UK), 180 pp.
- 21 Ladefoged, P., and Loeb, G. (2002). "Preliminary studies on respiratory activity in speech," *UCLA*
22 *Work. Pap. Phon.* **101**, 50–60.
- 23 Ladefoged, P., and McKinney, N. (1963). "Loudness, sound pressure, and subglottal pressure in
24 speech," *J. Acoust. Soc. Am* **35**, 454-460. <http://dx.doi.org/10.1121/1.1918503>

- 1 Leanderson, R., Sundberg, J., and von Euler, C. (1986). "Breathing muscle activity and subglottal
2 pressure dynamics in singing and speech," *STL-QPSR* **27**(4), 57–64.
- 3 Leanderson, R., Sundberg, J., and von Euler, C. (1987). "Role of diaphragmatic activity during
4 singing: a study of transdiaphragmatic pressures," *J. App. Phys.* **62**, 259–270.
- 5 Lecuit, V., and Demolin, D. (1998). "The relationship between intensity and subglottal pressure
6 with controlled pitch," *Proceedings of the 5th International Conference on Spoken Language
7 Processing*, Sydney, Australia, pp. 3079–3082.
- 8 Lieberman, P. (1967). *Intonation, Perception, and Language* (Massachusetts Institute of
9 Technology, Cambridge, MA), 240 pp.
- 10 Löfqvist, A., Carlborg, B., and Kitzing, P. (1982). "Initial validation of an indirect measure of
11 subglottal pressure during vowels," *J. Acoust. Soc. Am* **72**(2), 633–635.
12 <http://dx.doi.org/10.1121/1.388046>
- 13 Maeda, S. (1976). *A characterization of American English intonation*. (Massachusetts Institute of
14 Technology, Cambridge, MA), 334 pp.
- 15 Mooshammer, C. (2010). "Acoustic and laryngographic measures of the laryngeal reflexes of
16 linguistic prominence and vocal effort in German," *J. Acoust. Soc. Am* **127**, 1047–1058.
17 <http://dx.doi.org/10.1121/1.3277160>
- 18 Mücke, D., and Grice, M. (2014). "The effect of focus marking on supralaryngeal articulation – Is it
19 mediated by accentuation?," *J. Phon.* **44**, 47–61.
20 <http://dx.doi.org/10.1016/j.wocn.2014.02.003>
- 21 Niebuhr, O. (2008). "The effect of global rhythms on local accent perceptions in
22 German," *Proceedings of the 4th International Conference of Speech Prosody*, Campinas,
23 Brazil, pp. 331–334.
- 24 Ohala, J. (1990). "Respiratory activity in speech," In *Speech Production and Speech Modelling*,
25 edited by W. J. Hardcastle and A. Marchal (Kluwer, Dordrecht, the Netherlands), pp. 25–53.

- 1 Pinheiro, J. C., and D. M. Bates (2000). *Mixed-Effects Models in S and S-PLUS* (Springer-Verlag,
2 New York, NY) 528 pp.
- 3 R Development Core Team (2016). R: A language and environment for statistical computing. R
4 Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL:
5 <http://www.R-project.org>.
- 6
- 7 Rochet-Capellan, A., and Fuchs, S. (2013). "The interplay of linguistic structure and breathing in
8 German spontaneous speech," *Proceedings of the 14th Annual Conference of the*
9 *International Speech Communication Association [Interspeech]*, Lyon, France, pp. 1228–
10 1231.
- 11 Slifka, J. (2000). Respiratory constraints at prosodic boundaries in speech (Massachusetts Institute
12 of Technology, Cambridge, MA), 140 pp.
- 13 Stetson, R. H. (1951). *Motor Phonetics: A Study of Speech Movements in Action* (North-Holland
14 Publishing Company, Amsterdam, the Netherlands), 212 pp.
- 15 Strik, H., and Boves, L. (1995). "Downtrend in f₀ and P_{sub}," *J. Phon.* **23**, 203–220.
16 [http://dx.doi.org/10.1016/S0095-4470\(95\)80043-3](http://dx.doi.org/10.1016/S0095-4470(95)80043-3)
- 17 Titze, I. R. (1989). "On the relation between subglottal pressure and fundamental frequency in
18 phonation," *J. Acoust. Soc. Am* **85**(2), 901–906. <http://dx.doi.org/10.1121/1.397562>
- 19 van den Berg, J. (1958). "Myoelastic-aerodynamic theory of voice production," *J. Speech Hear.*
20 *Res.* **1**, 227–244. <http://dx.doi.org/10.1044/jshr.0103.227>
- 21 Zhang, Z. (in press). "Respiratory laryngeal coordination in airflow conservation and reduction of
22 respiratory effort of phonation," *J. Voice.*
23 <http://dx.doi.org.gate3.inist.fr/10.1016/j.jvoice.2015.09.015>
- 24