RESEARCH   **Open Access**

The aerodynamics of voiced stop closures

Luis M. T. Jesus[1*] and Maria Conceição Costa[2]

Abstract

Experimental data combining complementary measures based on the oral airflow signal is presented in this paper, exploring the view that European Portuguese voiced stops are produced in a similar fashion to Germanic languages. Four Portuguese speakers were recorded producing a corpus of nine isolated words with /b, d, g/ in initial, medial and final word position, and the same nine words embedded in 39 different sentences. Slope of the stop release (SLP), voice onset time (VOT), release and stop durations and steady-state oral airflow amplitude characteristics preceding and following the stop were analysed. Differences between independent groups (three different places of articulation and two vowel contexts) and correlations between variables were studied; generalised linear mixed effects models were developed to study the effects of VOT, SLP and the factors place of articulation and vowel context on the mean oral airflow. A classification of stop's voicing was automatically extracted. Both SLP ($p = .013$) and VOT ($p = .014$) were significantly different for the three places of articulation. Weak voicing was observed for 57% of the stops. It is hypothesised that the high percentages of weakly voiced stops are a consequence of passive voicing and that the feature of contrast in Portuguese is privative [*spread glottis*].

Keywords: Stops, Voicing, Voice onset time, Oral airflow, Portuguese

1 Introduction

The concept of contrast in the phonology of a language is closely linked to the competence of being able to isolate meaningful units such as phonemes or words. More specifically, the phonological laryngeal/voicing contrast is cued by a number of different features [13]: vocal fold vibration, duration of the adjacent phonemes and voice onset time (VOT) are just some of them.

The theoretical framework of this study is grounded on views of the laryngeal feature of contrast for stops that have been considerably enriched over the last decade by new acoustic and articulatory phonetics evidence which strengthened arguments that in some languages, stop voicing is phonologically active and in others, it is passive [6]. A clear relation between phonetic cues and phonological processes that support this has yet to be found, so studies such as ours, based on new aerodynamic data that is more closely related to laryngeal behaviour, could contribute towards clarifying these issues.

Laryngeal contrast has been shown to be highly correlated to VOT in a variety of languages but other parameters such as the duration, the fundamental frequency (f0) and the frequency of the first formant (F1) of adjacent vowels have also been proposed as cues of voicing [3, 6, 9, 13, 14, 20, 32, 33, 40, 46, 51, 54].

Current knowledge concerning the different contributions of acoustic parameters for voicing distinction in European Portuguese (EP) has been the focus of various studies based on adult's and children's acoustic data [8, 33, 40]. It has been shown that stop duration, duration of the preceding and following vowel, duration of voicing during closure, are relevant acoustic properties for the classification of voicing and that the percentage of devoiced exemplars decreases as the place of articulation moves anteriorly for word medial and word final stops [33]. A more recent cross-linguistic (Portuguese, Italian and German) speech production study looked at voicing status during closure based on time-dependent measures computed from voicing profiles [35, 40]. European Portuguese voicing patterns were different from other Romance languages and EP speakers' characteristics resembled those of German speakers. Velar stops from five out of six speakers were least likely to be produced with voicing during closure in low vowel context [40].

The motivation for this study is that although laryngeal articulation strategies used by EP speakers have

* Correspondence: lmtj@ua.pt
[1]School of Health Sciences (ESSUA) and Institute of Electronics and Informatics Engineering of Aveiro (IEETA), University of Aveiro, 3810-193 Aveiro, Portugal
Full list of author information is available at the end of the article

recently been recognised to differ from other Romance languages, inherent aerodynamic processes remain to be clarified [49]. This paper contributes towards clarifying what Solé ([49], p. 237) recently pointed out: "voicing patterns (and targets) may differ in language families and, therefore, a word of caution is in order when making generalizations about genetically related languages". Therefore, Portuguese language-specific features and attributes are explored and, how these mediate the speech outputs in relation to the place of articulation, the preceding and following phone is determined, providing a new insight into voicing contrast in EP. The corpus design and analysis methodology of complementary experimental measures based on the oral airflow signal of voiced stops and adjacent phones are presented in great detail. Novel results are discussed in the context of the most recent literature and conclusions are presented supporting the view that voicing, in Portuguese, results from speech mechanisms that have also been observed for German and English [36].

1.1 The aerodynamics of stops

The aerodynamics of transient speech sounds such as stops, [and] more particularly, their intraoral pressure and the nasal airflow, have been extensively described in the literature [47, 48, 55]. We focus here on studies that have used parameters based on oral airflow because valid glottal airflow mean amplitude values, inferred from oral airflow measures, have been shown to be a reliable indicator of laryngeal characteristics [4, 11, 16, 27].

Peak oral airflow values have been reported in consonant vowel (CV), vowel consonant (VC) and vowel consonant vowel (VCV) syllables where C was one of the stops /p, b, t, d, k, g/ and /i, ɑ/ were selected as vowels—V [16]. Results showed voiced stops' peak oral airflow values significantly lower than their voiceless cognates, nonsignificant vowel context effects in CV and VCV sequences and a tendency for female values to be lower than male's [16]. The lowest peak oral airflow average values were measured for /ibi/ syllables (66 cm^3/s for females and 112 cm^3/s for males) and the highest for /tɑ/ syllables (1162 cm^3/s when produced by female speakers and 1324 cm^3/s for male speakers). This was also one of the first papers to discuss relative flow values (linguistically more relevant than absolute values and in line with one of the central goals of speech production: to achieve broad aerodynamic targets), concluding that "air flow differences between voiced and voiceless productions may be largely attributable to the flow resistance imposed by vocal action in voicing" ([16], p. 253).

Additional mean peak values during closure reported in the literature include those of Stathopoulos and Weismer's [50] study ([b]—284 ± 123 cm^3/s; [d]—634 ± 164 cm^3/s; [g]—293 ± 111 cm^3/s).

Moreover, Cho et al. [11] studying fortis, lenis and aspirated bilabial stops in three real words (the bilabial stops were in word-initial position and followed by the vowel /e/) reported maximum oral airflow after stop release of more than 500 cm^3/s (up to 3500 cm^3/s for Seoul Korean speakers), and significant effect of stop category (fortis, lenis and aspirated) was found.

1.2 Contextual effects on stops' production

Various effects of vowel context on VOT, closure and release duration have been reported in the literature, but most of them have no systematic influence across languages and some results are even contradictory [1, 18, 33, 40, 41]. In French, /p, t, k/ closures have been found to be significantly longer than those of /b, d, g/, only between /a/ vowel contexts; and short-lag (positive) VOT values significantly longer in between voiceless fricative /s/ context than in between vowel /a/ context [1]. Italian short-lag VOT results showed a distinct behaviour for voiceless and voiced stops suggesting different laryngeal articulations to sustain vocal fold vibration [18]. Whereas in German closure durations reported were not systematically affected by vowel context and the percentage of devoiced stops was higher in low to mid-vowel context [41]. In EP, conflicting results have also been reported, with average VOT values not exhibiting any clear pattern regarding the influence of vowel height, and a more recent study showing that EP and German (not Italian) stops in low vowel context were more likely to be devoiced than in high vowel context [33, 40].

The effect of the place of articulation on acoustic correlates of voicing contrast in stops has also been the subject of various studies [2, 12, 15, 18, 25, 33, 40, 41], and the observation of language-specific variations in VOT has guided modifications [12, 18] to classical models of stop voicing [28] and motivated new studies on the "interaction of universal and language specific process" ([2], p. 68). In French, voiced stops' short-lag VOT has been found to significantly increase as the place of articulation moves more posteriorly but place of articulation does not seem to have a significant effect on closure durations [2].

Although initial evidence has suggested that stops in the context of high vowels would be less likely to devoice than stops in the context of nonhigh vowels, not much support for this was found in phonology [38, 39]. However, the degree of articulatory constraint (DAC) model of speech production predicts that different stop places of articulation result in various degrees of resistance to contextual effects [24, 44].

Results on the aerodynamic effect of vowel context on stops (e.g., the need to control for backness and the other contextual questions) and what effect it has on airflow are yet unclear, especially when real words are considered. Previous studies, presenting speaker-

specific vowel effects, were based on nonsense word productions of stops which have been recently shown to differ in terms of the observed patterns from real words [1, 18, 31, 40, 41, 52].

Table 1 shows key stop production-related results in the literature based on oral airflow amplitude measures. Significant vowel effects have been found in French, but in American English (AE) contextual effects are still unclear, although oral airflow signals conveying idiosyncratic elements of voicing onset and offset have been reported [10, 21, 22, 30, 31, 37].

1.3 Purpose of this study and research hypothesis

This paper contributes to a novel view on laryngeal contrast, focussing on intervocalic stops, in word initial, word medial or word final but not on utterance initial stops after a pause, because the objective was to investigate a number of vocalic and consonantal contexts, previously reported in the literature, that influence the maintenance of voicing in obstruents (including stops). Additionally, relative oral airflow parameters were developed for intervocalic stops.

The study's objectives (a first objective O.1 concerned with the definition of features of contrast and the second objective O.2 to define language-specific phonotactics) and the hypothesis (H1.1 to H4.10) that support them are:

- Objective 1 (O.1): Identify language-specific aerodynamic and voicing behaviours supporting the position that, contrary to other Romance languages, the EP feature of contrast is *privative* [*spread glottis*]

and that *passive voicing* (resulting from contextual effects, rather than laryngeal gestures by the speaker) can be used to describe voicing mechanisms in EP.

○ Hypothesis 1.1 (H1.1): More than 40% of the stops will be weakly voiced (devoiced). This threshold is based on previous results for German and English [6]. We will be analysing patterns of decreasing amplitude during stop closures that have been previously associated with *passive voicing* [6].

○ Hypothesis 1.2 (H1.2): Less than 5% of the stops will be produced with an average airflow above the Phonation Threshold Flow (PTF).

- This hypothesis supports the idea that the laryngeal feature is *privative*, i.e. that it is defined by the occurrence or absence of a laryngeal gesture [6]. Mean oral airflow is approximately equal to the glottal airflow ([27], p. 2880), so a rough estimate of the PTF can be noninvasively evaluated using an oro-nasal circumferentially vented mask setup (used in this study). Estimated values for this parameter vary between 180 cm^3/s and 1200 cm^3/s [23, 45].

○ Hypothesis 1.3 (H1.3): More than 95% of the EP voiced stops are produced with regular variations in magnitude (as seen in the oral airflow) that are generated by mucosal oscillations for a *spread glottis*.

- Objective 2 (O.2): Explore associations between the slope of the stop release (SLP), VOT, release duration (RLS), stop duration (STP), steady absolute

Table 1 Key literature results

	Tokens	Measures	Oral airflow effect of vowel context
Bucella et al. [10]	Sustained vowels /i, a, u/	Mean mid-vowel oral airflow amplitude	Significant vowel effect
Cho et al. [11]	Bilabial stops (fortis, lenis and aspirated) in three real words (same vowel context)	Maximum oral airflow after stop release	Significant effect of stop category (fortis, lenis and aspirated)
Higgins et al. [21]	Isolated repetitions of /pi/ and /pɑ/, and repetitions of <buy pip again> and <buy pap again>	Mean mid-vowel oral airflow amplitude	No significant intrasubject differences across vowel type
Higgins et al. [22]	Isolated repetitions of /pi/ and /pɑ/, and repetitions of <buy pip again> and <buy pap again>	Mean mid-vowel oral airflow amplitude	No significant interactions involving the factor of vowel for mean airflow
Koenig et al. [30]	Repetitions of <poppa popper> and <poppa bopper>	VCV oral airflow signals used to identify voicing onset and offset during stop closure	Some of the stops' oral airflow signals "did not show abrupt changes associated with vocal-tract closure and release, suggesting lenition or spirantization" ([30], p. 1080)
Koenig et al. [31]	Voiceless glottal fricative /h/ produced in repetitions of the sentences <a papa hopper>, <a papa hippie> and <a papa hooper>	Oral airflow amplitude at voicing onset and offset during stop closure	"…individual speakers have unique methods of achieving phonatory goals during running speech." ([31], p. 2535) "… the direction of vowel effects differed across speakers …" ([31], p. 2548)
Netsell et al. [37]	CV syllables where C was /p/, and V was /i/ or /a/	Mean mid-vowel oral airflow amplitude	No significant difference between oral airflow for [i] versus [a]

oral airflow amplitudes (OA1, OA2 and OA3, defined in Table 2) and relative oral airflow amplitude (A12, A23 and MOA, also defined in Table 2), place and vowel related variations to EP language-specific phonotactics.

　○ Hypothesis 2 (H2): There will be significant differences between the mean (or median) values of some of the variables, considering the three different places of articulation and considering the two vowel contexts.

Section 2 of this paper describes the method used to design the corpus, how data were collected and the criteria to annotate the oral airflow signals. This section also defines the relative measures and voicing classification procedures. Section 3 characterises VCV oral airflow waveforms, reports slopes of the stops' releases, VOT and stop durations, absolute and relative amplitudes of the oral flow waveforms and the results of a voicing classification. We then discuss results and contextualise them in light of the most recent studies on glottal behaviour during stop closures. We finally conclude the paper with a reflection on the main contributions of the current study, its limitations and perspectives for future work.

2 Method

Four Portuguese speakers produced voiced stops /b, d, g/, while oral airflow and electroglottographic (EGG) data were collected. The corpus included a rich variety of phonetic contexts that are known to condition voicing in obstruents [14].

2.1 Speakers, corpus and data acquisition

Data were collected from two adult male (SP1 and SP2) and two adult female (SP3 and SP4) speakers of EP with an age range of 20 to 39 years. None of the speakers had reported speech, language or hearing impairments.

Speakers SP2 and SP4 were certified speech and language therapists, SP1 a phonetician and SP3 a speech and language therapy student.

Speakers were asked to seat on a chair and read 48 prompts, displayed randomly on a sheet of paper held on a musical stand, which was placed in front of them, with normal effort and as close as possible to their natural speech: nine isolated words contained the EP voiced stops /b, d, g/ in word-initial, word-medial and word-final positions; the same nine words were embedded in 39 different carrier sentences of the form: <Diga X Y por favor>. The Appendix details the words and carrier sentences used, which were designed to include word and crossword contexts that might elicit devoicing (or help maintain voicing) in stops.

A number of different factors have been reported in the literature as having an influence in the maintenance of voicing. These have determined our choice of word and sentence contexts, namely: place of articulation—voicing may cease earlier for more posterior places of articulation; word-position—from an aerodynamic point of view a voiced obstruent is more likely to be produced in medial position, whereas in utterance initial and final position it is more probable to produce devoiced items; consonant duration—the longer the consonant duration, the more probable is that voicing ceases; context—stops coarticulated with high vowels maintain voicing longer than when coarticulated with low vowels [42].

The following segmental environments were used, with the number of tokens in parentheses:

- Environment 1—nine (9) words without a frame sentence.
  ○ We used these words to establish a baseline, because here stops are better controlled and easier to analyse than those occurring in frame sentences.
- Environment 2—thirty (30) words with stops in final position are produced in frame sentences where the

Table 2 Measured/calculated variables

Name	Abbreviation	Units	Variable description
Slope	SLP	%	Slope of the stop release
Voice onset time	VOT	ms	Time spent between voicing initiation and the beginning of the release
Release	RLS	ms	Release duration
Stop	STP	ms	Stop duration
Oral airflow 1	OA1	cm^3/s	Steady absolute oral airflow amplitude of phone 1
Oral airflow 2	OA2	cm^3/s	Steady absolute oral airflow amplitude of phone 2
Oral airflow 3	OA3	cm^3/s	Steady absolute oral airflow amplitude of phone 3
Amplitude 12	A12	%	Relative oral airflow amplitude between phones 1 and 2
Amplitude 23	A23	%	Relative oral airflow amplitude between phones 2 and 3
Mean oral airflow	MOA	%	Mean relative oral airflow amplitude between A12 and A23

word that follows the stop has various initial phonemic contexts.

○ Twelve (12) words with stops in final position are produced in a frame sentence where the word that follows the stop has the following initial segmental context (vowels were divided into two groups according to their height: group 1—/i, ɨ, u, e, o/, close and close-mid vowels; group 2—/ɛ, ɔ, ɐ, a/), open-mid and open vowels): vowel from group 1 (/i, ɨ, u, e, o/) followed by a lateral and a tap; vowel from group 2 (/ɛ, ɔ, ɐ, a/) followed by a lateral and a tap.

○ Six (6) words with stops in final position are produced in a frame sentence where the word that follows the stops has the following initial segmental context: vowel from group 1 (/i, ɨ, u, e, o/) followed by voiced velar stop; vowel from group 2 (/ɛ, ɔ, ɐ, a/) followed by voiced velar stop.

○ Six (6) words with stops in final position are produced in a frame sentence where the word that follows the stop has one of the following initial segmental contexts: vowel from group 1 (/i, ɨ, u, e, o/) followed by nasal stop; vowel from group 2 (/ɛ, ɔ, ɐ, a/) followed by nasal stop.

○ Six (6) words with stops in final position are produced in a frame sentence where the word that follows the stop has one of the following initial segmental contexts: vowel from group 1 (/i, ɨ, u, e, o/) followed by voiced postalveolar fricative; vowel from group 2 (/ɛ, ɔ, ɐ, a/) followed by voiced postalveolar fricative.

• Environment 3—nine (9) words with stops are embedded in the frame sentences <Diga X por favor> previously used, to facilitate comparisons [33].

○ The words X were [33]: <bala>; <juba>; <cabe>; <dava>; <nada>; <pode>; <gato>; <paga>; <pague>.

In environment 2, the carrier sentences had the form <Diga X Y por favor>, where X was one of the nine words, produced initially without a frame sentenced, and Y was a sequence starting with a word that had an initial phone chosen to represent one of the possible consonantal (taps, laterals, stops and nasals) or vocalic (close, open front and back vowels) real EP contexts.

The nine isolated words containing the EP voiced stops /b, d, g/ in word-initial, word-medial and word-final positions included vocalic contexts representing the different vowel heights used in EP: open, open-mid, close-mid and close.

Recordings were made in a quiet room using a circumferentially vented adult mask (Glottal Enterprises, USA) and a PT-2 pressure transducer (Glottal Enterprises, USA) for measuring the airflow at the mouth. An EGG signal was also collected using an EGG processor (model EG2-PCX, Glottal Enterprises, USA) and two channel 35 mm diameter electrodes (Glottal Enterprises, USA). The oral airflow and EGG signals were recorded with a MS 110 electronics unit (Glottal Enterprises, USA), connected via an audio interface (iMic, Griffin, USA) to a notebook running Waveview Pro Version 2.2.6 (16 bits, 44.1 kHz sampling frequency). Airflow calibration and zero-setting of signals were undertaken before each recording session using a Glottal Enterprises FC-1 airflow calibrator and Waveview Pro Version 2.2.6 standard procedures, e.g., a calibration airflow generated from 140 cc of air injected into the calibrator over 0.8–1.5 s.

2.2 Data annotation

The time waveforms of all the words were manually annotated using Praat Version 5.0.43 [7] to detect the start of the phone or silence (phone1) preceding the stop, the start and end of the stop and the end of the vowel after the stop, using the following specific criteria:

• The start of phone1 (preceding the stop) was defined by the presence of periodicity in the EGG signal and oral airflow waveform and checked against the spectrogram for a discernible second formant (F2).

○ When phone1 was silence, the start was defined at 100 ms prior to the start of phone2.

• The start of phone2 (the stop) was considered to occur when there was either a decrease in airflow amplitude of at least 50% of the maximum peak-to-peak relative of phone1 (see Fig. 1), or, when the stop was preceded by silence (see Fig. 2), prevoicing was discernible (presence of periodicity in the EGG signal and in the oral airflow waveform checked against the spectrogram for a discernible F2).

• For phone3, the start was defined as the point where at least one of the following criteria was satisfied:

○ Aspiration noise ceased.

○ Voicing restarted (as observed in the EGG signal).

○ There was an increase in oral airflow amplitude to at least 75% of the maximum peak-to-peak amplitude of phone3 (see Figs. 1 and 2).

• The end of phone3 was established by listening to the time-derivative of the oral airflow and checking the spectrogram for the presence of F2.

A stop burst was considered discernible when either a sudden peak (rise) in the oral airflow waveform or a vertical bar in its spectrogram could be observed (see Fig. 2). When multiple bursts were discernible, the one with the highest intensity was chosen, as it is believed to correspond to the actual start of the release [33].
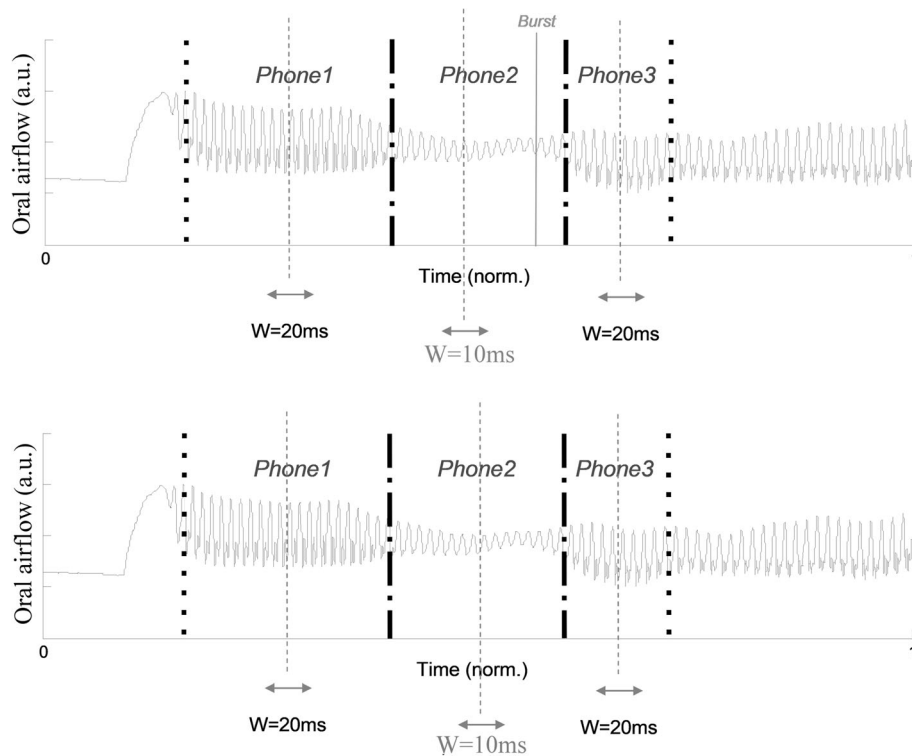
Fig. 1 Oral airflow signal. The vertical dotted lines represent the start of phone1 and the end of phone3, i.e. the speech units surrounding the target phone (which was always a stop). The vertical dash-dotted lines represent the start and end of phone2, the stop. The bidirectional arrows represent the location of the analysis windows. The strategy used to assign the vertical dashed line is illustrated for the tokens with a discernible burst (top) and without a discernible burst (bottom): *w*hen the burst was not discernible, this annotation corresponded to the middle of the stop. The x and y-axes were normalised

The acoustic signature of a stop is not always apparent due to intergestural overlap, which sometimes results in the absence of a clear stop release or burst [19]. To overcome this difficulty, criteria (described above) to annotate the stop release and burst based on the expected oral airflow signal (shown in Fig. 1) was developed. During the closure interval of voiced stops (characterised by regular oscillations in the oral airflow), the vocal folds remain in their position for voice during the entire closure interval; however, the oscillations die out as the back-pressure in the oro-pharyngeal cavity builds, and acts to oppose the lung pressure. This is the most prevalent voicing pattern in phrase-medial stops, as recently observed in acoustic signals [14]. The subsequent increase of the oral flow signal during the release has been reported in previous studies [11].

The criteria used to annotate the onset of stop closure was: The instant in time when the ripple observed on oral airflow waveform (corresponding to formant oscillations mainly at F1) is no longer visible and a small amplitude of periodic EGG signal is produced with no contact between the vocal folds as they are still vibrating apart. This pattern in the oral airflow waveform is shown in Fig. 1. In Fig. 2, the stop is in initial word (isolated) position so the start of the stop was signalled by an amplitude in the periodic airflow signal greater than zero (still without visible formant oscillations).

The release starts (third solid vertical line in Fig. 1) and end time following a closure (fourth solid vertical line at the bottom of Fig. 1) were identified by (as exemplified in Fig. 2—in some cases, this is visible in the oral airflow spectrogram, not in the oral airflow waveform): release (or burst) start time—an increase of the oral airflow signal; release end time—a decrease of the oral airflow signal amplitude and the start of the next phone.

Both the oral airflow and EGG signals were used to annotate the events in the VOT domain, resolving the issue of determining a threshold for the duration of voicing during closure [3].

2.3 Aerodynamic measures
Matlab 7.5.0 (R2007b) and Praat Version 5.0.4 scripts were used to extract the following metrics, based on average values calculated from 10 to 20 ms windows centred within the phones: absolute (cm^3/s) and relative (%) oral airflow of voiced stops and vowels. Oral airflow-based parameters have been shown to be useful and reliable when one is trying to understand the production
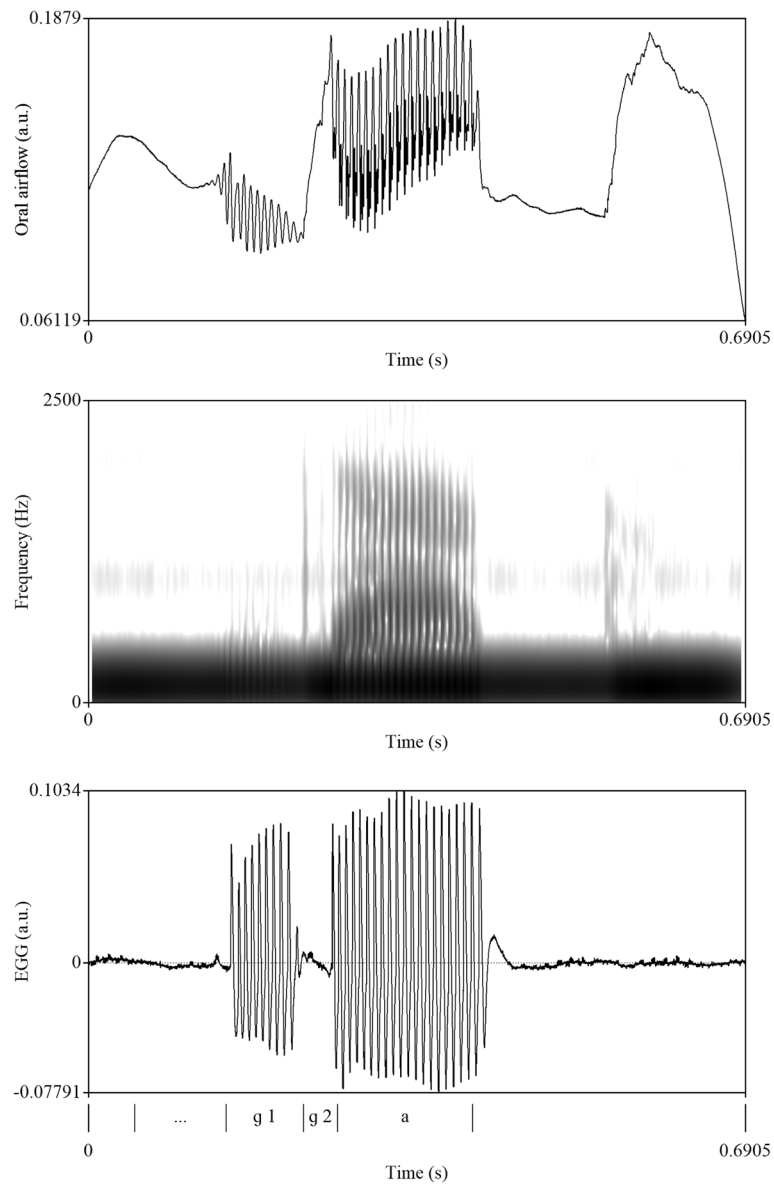
Fig. 2 Oral airflow and EGG signals of SP1's production of the word <gato>. From top to bottom: oral airflow waveform; spectrogram of oral airflow signal; EGG signal waveform; annotation (phone 1 is silence [...]; phone 2 [g] is divided into two intervals—closure (g1) and release (g2); phone 3 [a]). The stop has a discernible burst (visible in the oral airflow waveform and spectrogram)

mechanisms of voiced obstruents [42, 43]. Thus, the choice of parameters in the current paper.

Previous aerodynamic studies used peak oral airflow measures to extract information from the release of the stop [11, 16]. In this paper, relative measures (see Eqs. 1 and 2) that can be used to relate the data from different speakers were analysed. The slope of the stop release (SLP) was calculated from linear regression in Matlab 7.5.0 (R2007b), using all airflow signal points from the start to the end of the release (see Fig. 1). Analysis windows were also defined at three different production

stages: stop closure, and steady state of phones preceding and following target stop. Absolute mean oral flow values and amplitude of oscillations were extracted from these windows for all recordings and speakers, and relative vowel-stop and stop-vowel amplitudes were based on average values calculated from 20 ms windows centred in phone1 and phone3 (see Fig. 1), and a window of 10 ms centred in the stops (phone 2) without a discernible burst, and centred in the closure interval for the other stops.

The relative vowel-stop (phone1–phone2) amplitude values and the relative stop-vowel (phone2–phone3)

amplitude values were calculated using the following formulas ([42], p. 628):

$$\text{phone } (1\text{–}2)(\%) = \frac{\left[\text{Mean (phone1)}_{W_{20\,\text{ms}}} - \text{Mean (phone2)}_{W_{10\,\text{ms}}}\right] \times 100}{\text{Mean (phone1)}_{W_{20\,\text{ms}}}}$$

$$(1)$$

$$\text{phone } (2\text{–}3)(\%) = \frac{\left[\text{Mean (phone3)}_{W_{20\,\text{ms}}} - \text{Mean (phone2)}_{W_{10\,\text{ms}}}\right] \times 100}{\text{Mean (phone3)}_{W_{20\,\text{ms}}}}$$

$$(2)$$

The oral airflow amplitude-based parameters include the comprehensive set shown in Table 2.

2.4 Voicing classification

A decrease in the amplitude of the oral airflow waveforms, during the production of obstruents, when compared to the amplitude of the previous and following phone (see Fig. 1) has been previously observed [42, 43]. The terms *weak voicing* [17] and *slack voicing* [26] have been used by various authors to contrast laryngeal activity during stop production with that typically observed in unreduced vowels (characterised as having strong voicing). The stop's voicing was thus classified into two categories (weak and strong: weak voicing is characterised by a mean ratio [mean of A12 (%) and A23 (%), i.e. MOA(%)] of average oral airflow in the stop to that in the preceding and following phone that is more than 70%.

Figure 3 shows examples of strong voicing at 56% and weak voicing at 82% for the stop [b].

The empirical definition of the 70% threshold was based on the study by Pinho et al. ([42], pp. 635–636) "considering the mean value of all the relative measures of oral airflow amplitude from all the tokens for all speakers". The threshold value comes from computing the mean relative value found across all speakers for each VCV sequence for both phone (1–2)% and phone (2–3)%.

2.5 Statistical analysis

Data analysis considered the variables SLP (%), VOT (ms), release duration—RLS (ms), stop duration—STP (ms), amplitude of the oral airflow signal (cm^3/s) at phones 1, 2 and 3 (OA1, OA2 and OA3) and relative amplitude (%) phone1–2 (A12) and phone2–3 (A23). Inferential analysis was conducted with place of articulation (PLA) and vowel context (VOW) as factors, and the results obtained are presented at the .05 significance level:

- Place of articulation—PLA
  - ○ Bilabial /b/
  - ○ Dental /d/
  - ○ Velar /ɡ/
- Vowel context—VOW
  - ○ Vowel context 1—close and close-mid vowels (higher vowel /i, ɨ, u, e, o/)
  - ○ Vowel context 2—open-mid and open vowels (lower vowel /ɛ, ɔ, ɐ, a/)

Data normality analysis, comparisons between groups and correlation analysis were carried out using R (RStudio, Version 1.0.143). To infer about the normality of the distributions underlying the data, the Kolmogorov-Smirnov test with the Lilliefors' significance correction was run for each independent group. It was concluded that no comparison could be done in a parametric context, as the assumption of normality could not be considered to hold for any set of independent groups of data, at the .05 significance level. As such, comparisons between medians of any two independent groups were made using the Mann-Whitney U test (i.e. whenever the factor considered was VOW, the vowel context), and comparisons between any three or more independent groups were made using the Kruskal-Wallis test (i.e. whenever the considered variables were analysed against the factor PLA, place of articulation).

For the analysis of the correlations between all possible pairs of variables, firstly considering all stops together, and then for each place of articulation, the Pearson correlation coefficient was calculated. However, having concluded that it is not plausible that data comes from a bivariate normal distribution, statistically significant correlations were reported at the .05 significance level, considering the Spearman rank-based correlation test, based on the Spearman's rank correlation coefficient, which was also calculated.

Mixed effects models of the mean oral airflow (MOA) were developed using the lmer function in the lme4 package in R [5]. A mixed effects model was considered with VOT, SLP and the factors PLA and VOW as the fixed effects (without interaction terms). As there are multiple measures per speaker and there is significant individual variation in MOA (as can be seen in Fig. 4), speakers were treated as random effects. Given that MOA variation between stops is much less than between speakers (see Fig. 4), only speakers were considered as random effects, with random intercepts.

Other specifications for the mixed effects model were tested, but the likelihood ratio tests performance pointed towards the proposed model. In fact, running the analysis of variance (ANOVA) methodology with different models as arguments returns some model comparison statistics such as the chi-square statistic representing the difference in deviance between successive models and the p values based on likelihood ratio test comparisons. When considering different specifications for the mixed effects model, no significant drop in deviance was observed. Therefore, the model considered here is the
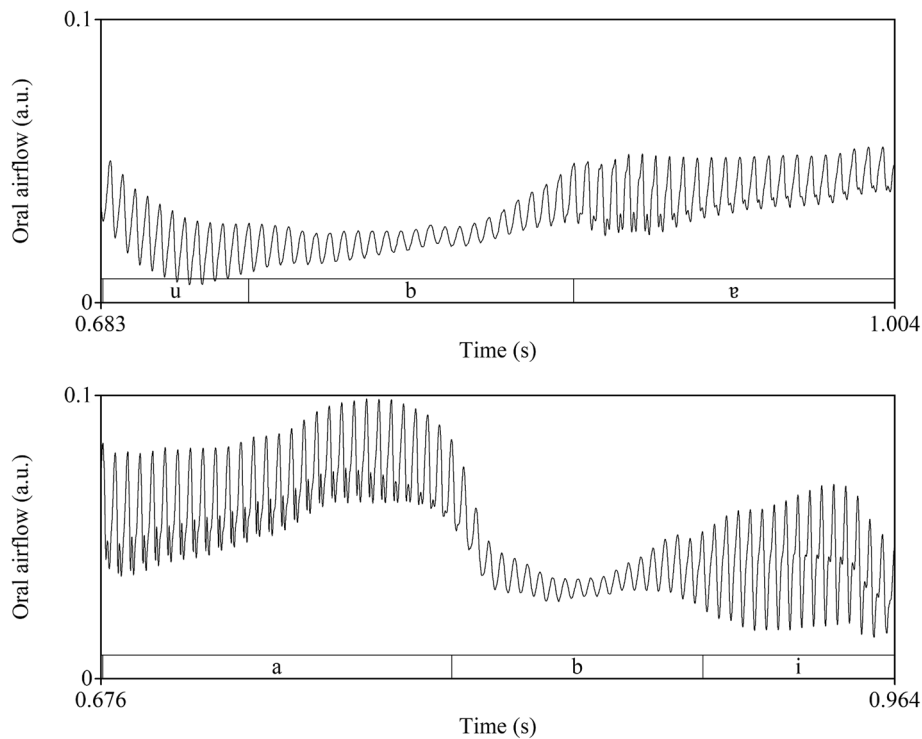
Fig. 3 Oral airflow waveform of two SP3's productions of strong voiced [b] (top) and weakly voiced [b] (bottom)

model that corresponds to the R implementation formula given by Eq. 3:

$$\text{lmer} (\text{MOA} \sim \text{factor(PLA)} + \text{factor(VOW)} + \text{VOT} + \text{SLP} + (1|\text{speaker}))$$
(3)

3 Results

3.1 Oral airflow waveforms

All stops presented periodicity in the oral airflow waveform during closure, i.e. according to our data, there was only one voicing distribution/shape [14]: Continuous (weak) voicing throughout the whole closure (as shown in Figs. 5, 6 and 7).

However, the burst was only discernible for 55% of the tokens, a phenomenon previously observed in EP [33], resulting in low oral pressures which facilitate vocal fold oscillation [14]. The SLP, VOT and RLS parameters (described below) were only calculated for tokens with a discernible burst and the others were excluded. Figures 5, 6 and 7 indicate, with vertical dashed lines, the burst when this is discernible in either the oral airflow waveform or its spectrogram, for all the tokens of one of the speakers.

The small oral airflow oscillations we observed during closure (shown in Figs. 5, 6 and 7) "do not generally produce significant acoustic excitation" ([42], p. 633), which would have resulted in what Abramson and Whalen [3]

designate as voiceless closure, i.e. had we based our measurements in acoustic data, they would have resulted in short-lag (positive) VOT values. This view is founded in previous acoustic waveform and spectrographic data shown in Fig. 8, where one can observe in what ways Portuguese voiced stops resemble those found in several Germanic languages [40]. For example, the German [b, g] and Portuguese [d] are devoiced as can be seen from the aperiodic waveform and the lack of a voice bar during closure, and the German [d] and Portuguese [g] present multiple bursts, as seen from the "vertical smudge" in the spectrogram. Further cross-language (Germanic versus Romance) comparisons can be found in Pape and Jesus [40].

3.2 Slopes of the stops' releases

The values for the mean slope of the stop release (SLP) and its standard deviation, and the median, are shown in Table 3 for each of the four speakers. Comparing the SLP values for the three different places of articulation, significant differences (t = 8.635, df =2, p = .013, Kruskal-Wallis test) were found. Running a post-hoc analysis of pairwise comparisons through the Tukey and Kramer (Nemenyi) test, with the Tukey distance approximation for independent samples, significant differences were only found between [d] and [g] (p =. 025). There was no significant difference between vowel contexts.
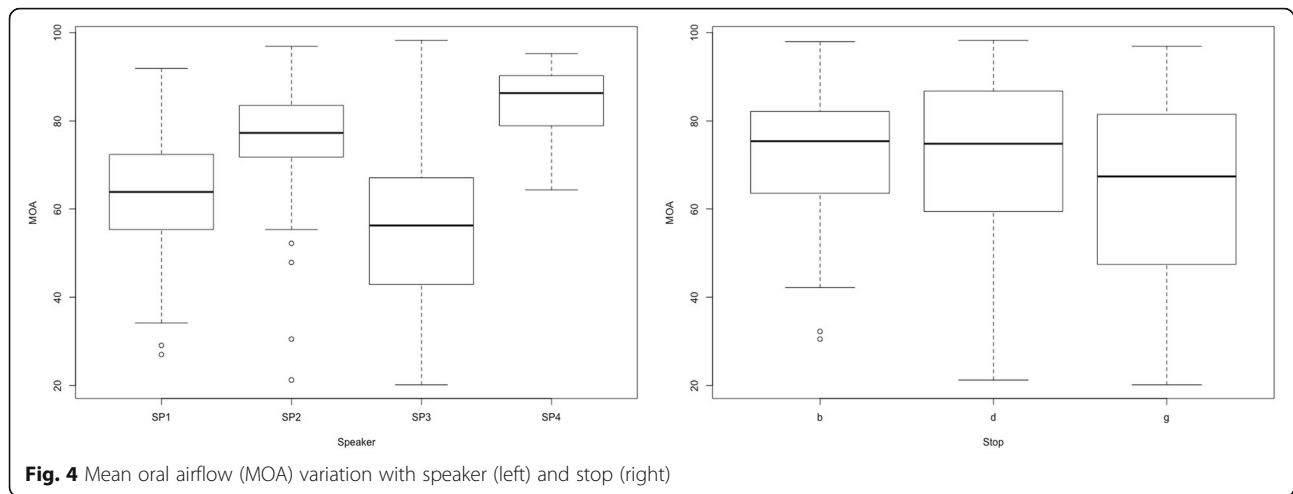
Fig. 4 Mean oral airflow (MOA) variation with speaker (left) and stop (right)

3.3 Voice onset time, release and stop durations

The voice onset time (VOT) values (shown in Table 4) were found to be significantly different (t = 8.504, df = 2, p = .014, Kruskal-Wallis test) for the different places of articulation. Although having observed that VOT values for velar stops (median VOT = − 61.0 ms) were higher than for bilabial stops (median VOT = − 53.5 ms) and dental stops presented the lowest VOT values (median VOT = − 47.0 ms), significant differences were only found between dental and velar stops (p = .019), according to the post-hoc analysis of pairwise comparisons through the Tukey and Kramer (Nemenyi) test, with the Tukey distance approximation for independent samples. There were no significant differences between medians of the VOT values comparing the two vowel contexts.

The values for the mean release duration (RLS) and stop duration (STP) and their standard deviation, and median values, are shown in Tables 5 and 6, respectively.
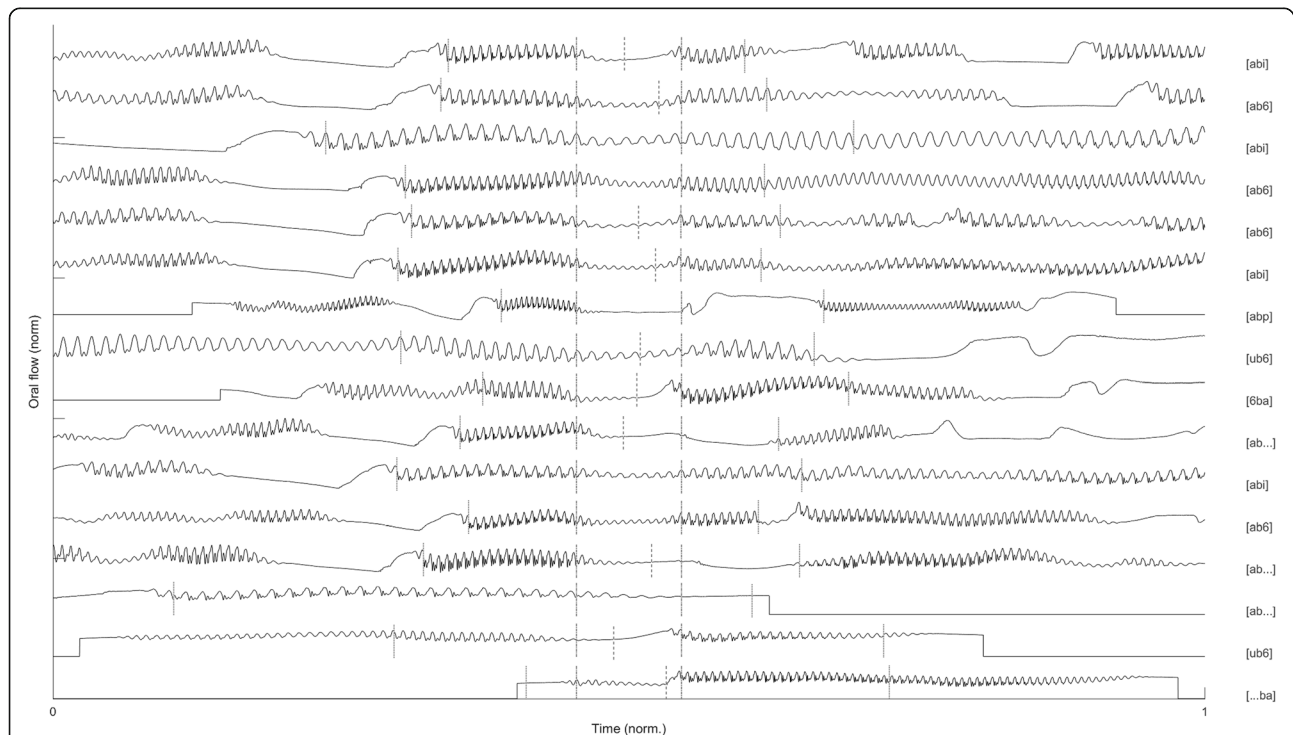

Fig. 5 Oral airflow waveforms of words with stop [b] (16 files), produced by male speaker SP1. The vertical dotted lines indicate the start of phone1 and the end of phone3. The vertical dash-dotted lines represent the phone1–phone2 and phone2–phone3 boundaries where phone2 is the stop. The vertical dashed lines indicate the burst when this is discernible in either the oral airflow waveform or its spectrogram. The x- and y-axes were normalised

Fig. 6 Oral airflow waveforms of words with stop [d] (16 files), produced by male speaker SP1. The vertical dotted lines indicate the start of phone1 and the end of phone3. The vertical dash-dotted lines represent the phone1–phone2 and phone2–phone3 boundaries where phone2 is the stop. The vertical dashed lines indicate the burst when this is discernible in either the oral airflow waveform or its spectrogram. The x- and y-axes were normalised
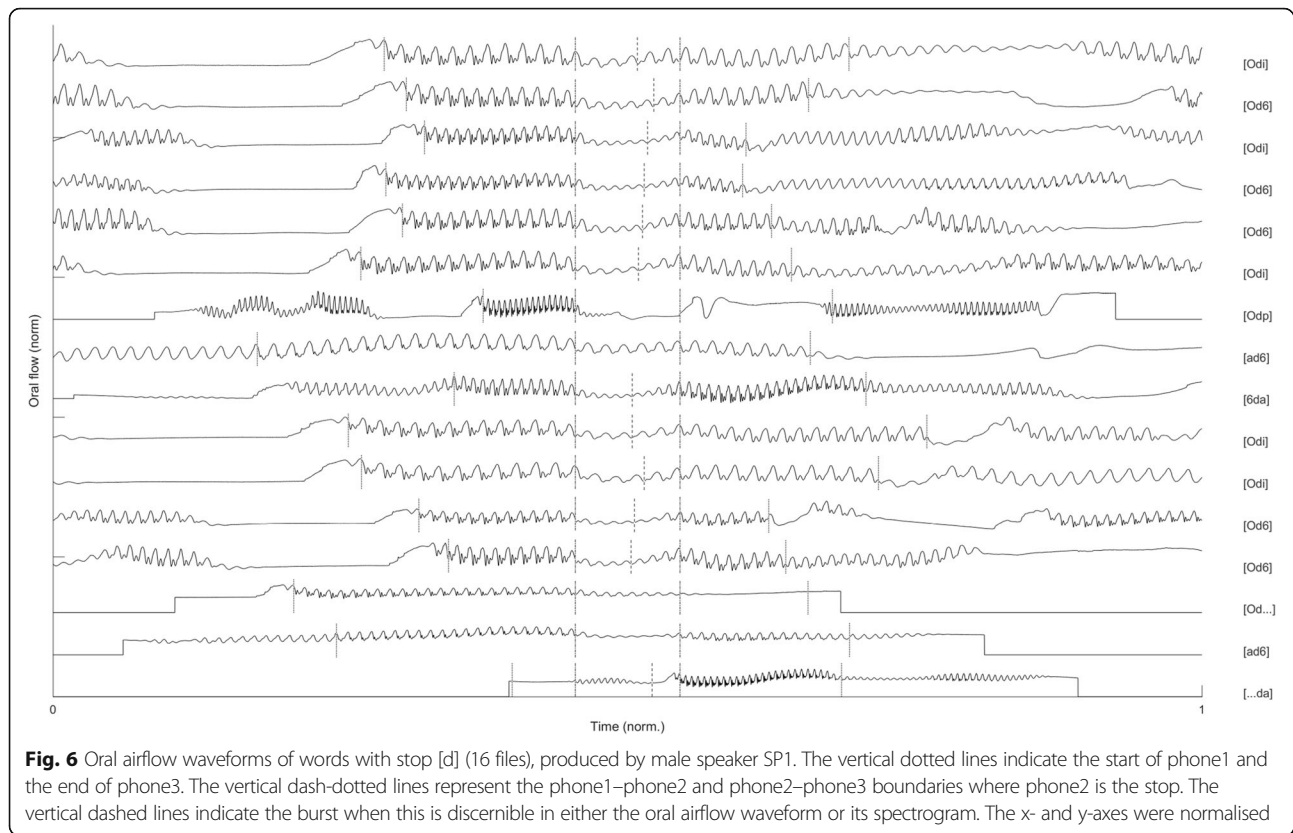

Fig. 7 Oral airflow waveforms of words with stop [g] (16 files), produced by male speaker SP1. The dotted vertical lines indicate the start of phone1 and the end of phone3. The vertical dash-dotted lines represent the phone1–phone2 and phone2–phone3 boundaries where phone2 is the stop. The vertical dashed lines indicate the burst when this is discernible in either the oral airflow waveform or its spectrogram. The x- and y-axes were normalised
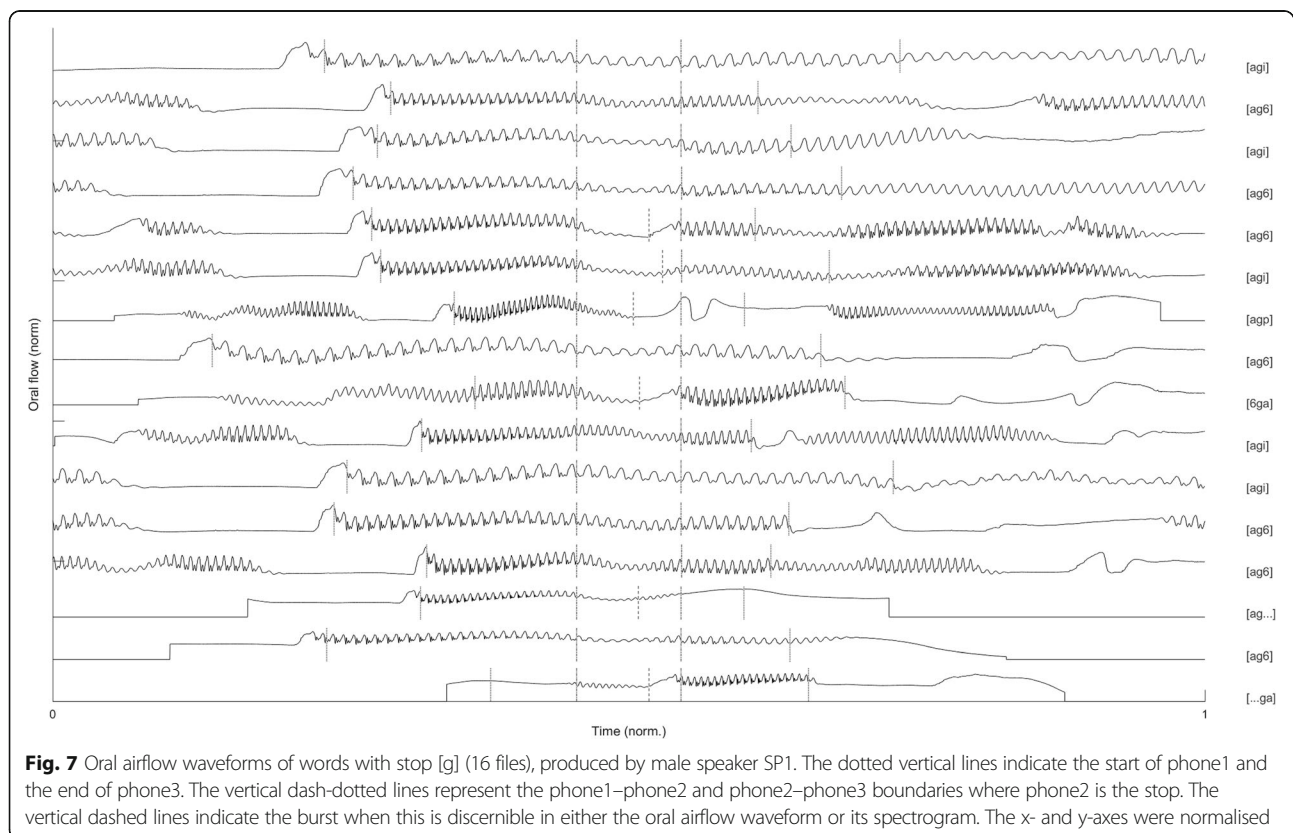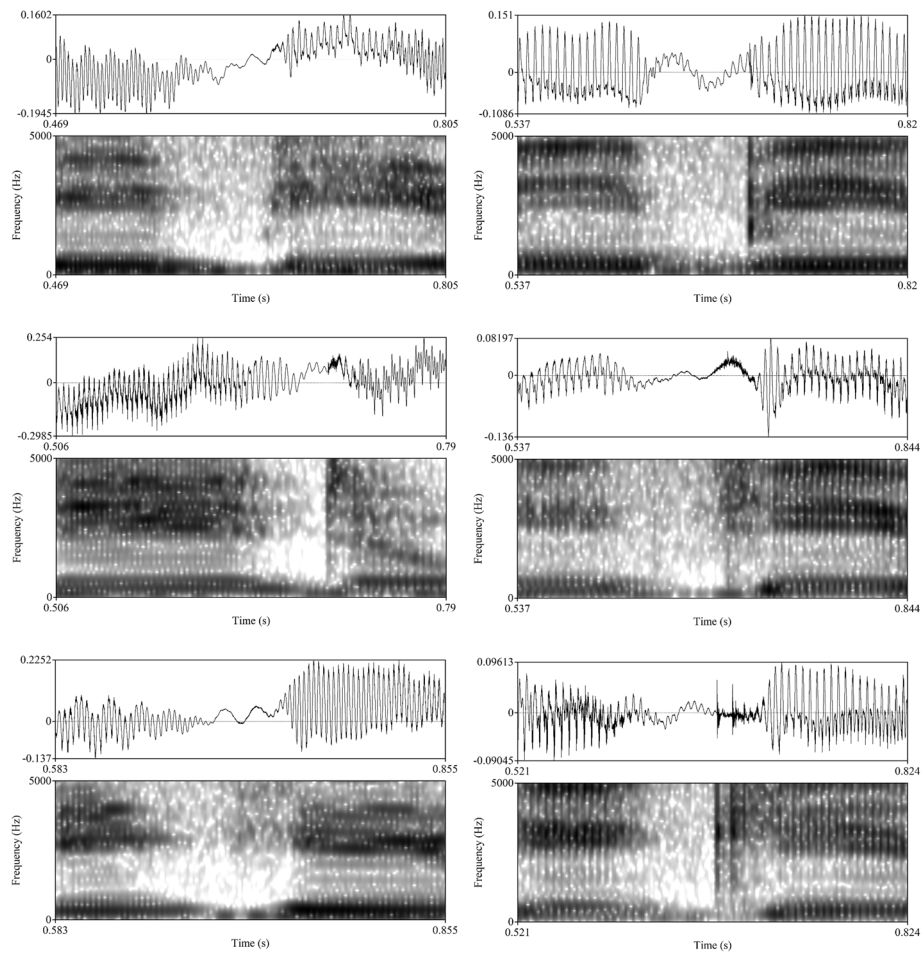
Fig. 8 Waveforms and spectrograms of acoustic data collected by Pape and Jesus [40] for a German (left) and a Portuguese (right) female speaker: speech signals corresponding to [ibi] (top), [idi] (middle) and [igi] (bottom) sequences from the previously recorded [40] CVCV (consonant, vowel, consonant, vowel) items in the context of a frame sentence. The consonant and vowel were pairwise identical (e.g., <bibi>, <didi> and <gigi>), sentence stress was on the CVCV pseudoword, and lexical stress was set to the first syllable of the CVCV pseudoword

Table 3 Mean (standard deviation) SLP, number of tokens and median SLP

	[b]	[d]	[g]	
♂SP1	72.7 (13.8)	77.6 (6.9)	78.7 (4.1)	Mean (standard deviation) SLP (%)
	10	12	6	Number of tokens
	77.7	78.8	80.5	Median SLP (%)
♂SP2	83.7 (4.4)	84.9 (2.2)	79.4 (5.2)	Mean (standard deviation) SLP (%)
	5	9	6	Number of tokens
	82.1	84.9	80.1	Median SLP (%)
♀SP3	64.7 (9.5)	72.3 (11.5)	62.3 (10.8)	Mean (standard deviation) SLP (%)
	11	16	11	Number of tokens
	64.9	76.6	61.1	Median SLP (%)
♀SP4	70.1 (9.4)	76.8 (8.0)	61.2 (15.6)	Mean (standard deviation) SLP (%)
	8	13	5	Number of tokens
	69.7	77.9	62.3	Median SLP (%)

Table 4 Mean (standard deviation) VOT values, number of tokens and median VOT

	[b]	[d]	[g]	
♂SP1	− 50.4 (16.0)	− 40.6 (17.9)	− 68.3 (10.3)	Mean (standard deviation) VOT (ms)
	10	12	6	Number of tokens
	− 47.5	− 37.0	− 71.0	Median VOT (ms)
♂SP2	− 74.8 (29.0)	− 60.2 (31.0)	− 92.5 (79.9)	Mean (standard deviation) VOT (ms)
	5	9	2	Number of tokens
	− 83.0	− 47.0	− 92.5	Median VOT (ms)
♀SP3	− 66.5 (29.0)	− 51.0 (18.0)	− 57.0 (15.3)	Mean (standard deviation) VOT (ms)
	11	15	11	Number of tokens
	− 59.0	− 46.0	− 58.0	Median VOT (ms)
♀SP4	− 62.1 (22.6)	− 64.9 (23.5)	− 83.0 (31.0)	Mean (standard deviation) VOT (ms)
	8	13	4	Number of tokens
	− 51.0	− 65.0	− 83.0	Median VOT (ms)

No significant differences were found between places of articulation or vowel contexts.

3.4 Amplitude of the oral flow waveforms

Absolute oral airflow values, shown in Table 7, were not significantly different for place of articulation or vowel context.

The relative amplitudes of the oral airflow waveforms phone(1–2)% (shown in Table 8) were not significantly different considering place of articulation and vowel context. Phone(2–3)% relative values were also not significantly different for place of articulation and vowel context. Therefore, factors PLA (place of articulation) and VOW (vowel context) do not seem to affect amplitudes of oral airflow at phones 1, 2 and 3, before or after the normalisation procedure (i.e. considering the absolute oral airflow values, OA1, OA2 and OA3, or the relative oral airflow values, A12 and A23).

3.5 Correlation analysis

For the correlation analysis, the Pearson's correlation coefficient was calculated in order to evaluate the degree of linear dependence between any two given variables. However, as the normality assumption could not be verified for any of the variables considered, the parametric test of significance based on the Pearson's correlation coefficient is not appropriate and, instead, the Spearman correlation coefficient and the correspondent rank-based correlation nonparametric test are presented.

Significant correlations (details shown in Table 9) were found between SLP and RLS; SLP and STP; SLP and A23; VOT and STP, when all stops were analysed together. There were also significant correlations between the values of the steady absolute oral airflow amplitude of phones 1 and 3, OA1/OA3 (number of tokens: 157; Pearson's correlation coefficient 0.833; Spearman's correlation coefficient: 0.870; p = .000).

When individual analysis of correlations between parameters was developed for each place of articulation, and according to Table 10, significant results were found for:

○ SLP/RLS—places of articulation [b] (number of tokens: 34; Pearson's correlation coefficient − 0.610; Spearman's correlation coefficient: − 0.684; p = .000) and [d] (number of tokens: 49; Pearson's correlation coefficient − 0.600; Spearman's correlation coefficient: − 0.381; p =. 007)
○ VOT/RLS—place of articulation [b] (number of tokens: 34; Pearson's correlation coefficient 0.385; Spearman's correlation coefficient: 0.352; p = .041)
○ VOT/STP—places of articulation [b] (number of tokens: 34; Pearson's correlation coefficient − 0.805; Spearman's correlation coefficient: − 0.713; p = .000), [d] (number of tokens: 49; Pearson's correlation coefficient − 0.906; Spearman's correlation coefficient: − 0.907; p = .000) and [g] (number of tokens: 23; Pearson's correlation coefficient − 0.877; Spearman's correlation coefficient: − 0.866; p = .000)
○ VOT/A12—places of articulation [d] (number of tokens: 38; Pearson's correlation coefficient − 0.347; Spearman's correlation coefficient: − 0.351; p = .031) and [g] (number of tokens: 16; Pearson's correlation coefficient − 0.528; Spearman's correlation coefficient: − 0.621; p = .010)
○ VOT/A23—Place of articulation [g] (number of tokens: 17; Pearson's correlation coefficient − 0.398; Spearman's correlation coefficient: − 0.585; p = .014)
○ VOT/MOA—place of articulation [g] (number of tokens: 19; Pearson's correlation coefficient − 0.516; Spearman's correlation coefficient: − 0.601; p =. 007)
○ STP/RLS: places of articulation [d] (number of tokens: 49; Pearson's correlation coefficient 0.587; Spearman's correlation coefficient: 0.421; p = .003) and

Table 5 Mean (standard deviation) RLS, number of tokens and median RLS

	[b]	[d]	[g]	
♂SP1	31.9 (15.6)	23.2 (6.5)	37.5 (16.6)	Mean (standard deviation) RLS (ms)
	10	12	6	Number of tokens
	26.5	22.0	35.5	Median RLS (ms)
♂SP2	27.4 (10.7)	25.3 (7.3)	29.5 (2.1)	Mean (standard deviation) RLS (ms)
	5	9	2	Number of tokens
	31.0	28.0	29.5	Median RLS (ms)
♀SP3	39.2 (18.4)	39.3 (15.4)	40.6 (17.2)	Mean (standard deviation) RLS (ms)
	11	15	11	Number of tokens
	40.0	36.0	38.0	Median RLS (ms)
♀SP4	34.8 (11.3)	26.0 (7.9)	24.8 (7.3)	Mean (standard deviation) RLS (ms)
	8	13	14	Number of tokens
	38.5	26.0	27.5	Median RLS (ms)

[g] (number of tokens: 23; Pearson's correlation coefficient 0.570; Spearman's correlation coefficient: 0.617; p = .002)

○ STP/A12—place of articulation [d] (number of tokens: 48; Pearson's correlation coefficient 0.243; Spearman's correlation coefficient: 0.366; p = .011)

○ STP/MOA—place of articulation [d] (number of tokens: 53; Pearson's correlation coefficient 0.098; Spearman's correlation coefficient: 0.276; p = .045)

Key findings were shorter releases (lower RLS values) resulted in steeper slopes (higher SLP values), which were also significantly correlated to higher relative oral airflow values (A23); steeper slopes (SLP) were correlated to shorter releases (RLS) in bilabial and dental stops.

A particularly striking (high correlation coefficient values) result was that for VOT/STP correlations. Higher A12 values were correlated to shorter dental and velar stops VOT, and VOT/MOA were significantly (negatively) correlated, for velar stops.

Finally, significant positive correlations were found between STP and RLS of dental and velar stops.

3.6 Mixed effects models of the mean oral airflow

Considering that in this study there are only four speakers, we are in the context of repeated measures experiments. Instead of computing means for each speaker, as would be the case in a traditional analysis of variance (ANOVA) approach, there is no doubt that a better insight of the data can be obtained considering a mixed effects model, where, besides including all data points produced by a single speaker, there is the possibility of accounting for both by-speaker and by-item variance. As explained in Section 2.5., several specifications for the mixed effects model were tested, but the one being considered here corresponds to the model described by Eq. 3.

Table 6 Mean (standard deviation) STP, number of tokens and median STP

	[b]	[d]	[g]	
♂SP1	82.1 (25.1)	67.9 (27.9)	81.2 (27.8)	Mean (standard deviation) STP (ms)
	16	16	16	Number of tokens
	83.5	63.5	82.0	Median STP (ms)
♂SP2	81.0 (26.7)	79.3 (31.3)	82.9 (33.7)	Mean (standard deviation) STP (ms)
	16	16	16	Number of tokens
	81.5	68.5	74.5	Median STP (ms)
♀SP3	105.8 (30.8)	90.0 (29.6)	98.8 (26.9)	Mean (standard deviation) STP (ms)
	16	16	16	Number of tokens
	99.0	82.0	89.5	Median STP (ms)
♀SP4	83.9 (22.2)	90.3 (19.1)	88.1 (24.2)	Mean (standard deviation) STP (ms)
	16	16	16	Number of tokens
	78.5	92.5	84.0	Median STP (ms)

Table 7 Mean (standard deviation) absolute values (OA), number of tokens and median absolute values of oral airflow. Columns designated as (1) present values for a 20 ms window centred in the phone preceding the target stop; (2) a 10 ms window centred in the middle of the closure; (3) a 20 ms window centred in the phone after the target stop

Phone	[b]			[d]			[g]			
	1	2	3	1	2	3	1	2	3	
♂SP1	244.2 (37.0)	74.2 (27.2)	238.1 (31.4)	230.6 (24.0)	83.4 (31.4)	220.1 (23.8)	227.0 (27.7)	101.1 (34.2)	227.7 (31.8)	Mean (standard deviation) OA (cm^3/s)
	15	12	12	15	15	14	14	14	13	Number of tokens
	244.6	74.8	233.6	228.5	75.9	217.4	229.1	103.2	221.3	Median OA (cm^3/s)
♂SP2	490.4 (76.4)	107.9 (59.4)	390.5 (124.2)	475.4 (82.0)	123.9 (46.8)	401.6 (87.3)	453.4 (101.8)	69.1 (32.4)	402.7 (105.6)	Mean (standard deviation) OA (cm^3/s)
	14	14	14	13	13	14	15	7	15	Number of tokens
	491.6	100.8	402.4	461.1	105.5	409.2	458.7	72.6	387.4	Median OA (cm^3/s)
♀SP3	107.5 (17.2)	41.5 (17.1)	119.3 (43.5)	125.1 (28.6)	47.8 (29.2)	127.9 (38.8)	104.0 (22.4)	61.6 (23.9)	119.3 (32.8)	Mean (standard deviation) OA (cm^3/s)
	15	16	12	15	9	16	15	14	16	Number of tokens
	107.5	39.5	118.2	130.4	44.6	127.6	102.9	70.8	120.2	Median OA (cm^3/s)
♀SP4	285.3 (51.1)	51.0 (12.4)	280.9 (62.6)	296.6 (88.7)	29.8 (11.0)	258.3 (83.1)	271.9 (49.5)	38.8 (21.7)	248.4 (72.8)	Mean (standard deviation) OA (cm^3/s)
	15	16	15	15	16	15	15	16	16	Number of tokens
	305.7	48.5	286.6	295.7	28.8	268.2	269.1	33.5	255.4	Median OA (cm^3/s)

The model was fitted by restricted maximum likelihood (REML) and the value of the REML criterion at convergence, which is the equivalent of the deviance for models fitted by maximum likelihood (ML), was 620.7. The estimate of the standard deviation of the random effects for the intercept was 13.64, as seen in Table 11. The "Residual" standard deviation stands for the estimation of the remaining variability that is not due to individual by-speaker variation, and since it is lower than the variability explained by the inclusion of

the random factor for Speaker, we can conclude that the random factor makes a good contribution to the model.

Regarding the structural part of the model, the estimated coefficients for the fixed effects are given in Table 12, where the confidence intervals via Wald approximations are also provided (Lower Bound (LB) and Upper Bound (UB) of the 95% Confidence Interval (CI)).

The estimated equation for the structural part of the model can be written as

Table 8 Mean (standard deviation) relative values of oral airflow (A12 and A23), number of tokens and median relative values of oral airflow

Oral airflow (%)

Phones	[b]		[d]		[g]		
	1-2	2-3	1-2	2-3	1-2	2-3	
♂SP1	71.6 (10.4)	69.0 (10.2)	65.4 (12.5)	62.2 (15.5)	55.0 (16.7)	52.4 (13.6)	Mean (standard deviation) rel. A (%)
	11	11	14	14	13	12	Number of tokens
	70.2	67.8	71.2	69.7	58.4	58.4	Median rel. A (%)
♂SP2	77.5 (10.0)	66.2 (28.4)	74.5 (10.2)	66.4 (17.3)	85.3 (6.8)	78.6 (11.9)	Mean (standard deviation) rel. A (%)
	12	13	11	12	7	7	Number of tokens
	77.0	74.5	77.6	74.0	82.9	79.8	Median rel. A (%)
♀SP3	60.9 (18.2)	64.7 (17.8)	54.0 (15.2)	64.5 (17.3)	41.8 (22.7)	50.7 (15.8)	Mean (standard deviation) rel. A (%)
	15	12	8	9	13	14	Number of tokens
	62.1	63.8	57.0	62.3	38.7	51.1	Median rel. A (%)
♀SP4	81.5 (4.3)	81.5 (5.6)	88.4 (6.8)	86.3 (9.4)	85.7 (7.2)	81.6 (13.4)	Mean (standard deviation) rel. A (%)
	15	15	15	15	15	16	Number of tokens
	81.4	81.9	90.1	88.9	85.7	87.1	Median rel. A (%)

Table 9 Results from the correlation analysis: number of tokens; Pearson's correlation coefficient; Spearman's correlation coefficient, *p* value of the Spearman rank-based correlation test

	RLS	STP	A12	A23	MOA	
SLP	106	112	81	85	93	Number of tokens
	− 0.574	− 0.269	0.268	0.232	0.181	Pearson's correlation coef.
	− 0.546	− 0.202	0.216	0.225	0.171	Spearman's correlation coef.
	.000*	.033*	.053	0.039*	.101	*p* value
VOT	106	106	80	83	91	Number of tokens
	− 0.015	− 0.872	− 0.235	− 0.042	− 0.114	Pearson's correlation coef.
	− 0.019	− 0.867	− 0.217	− 0.079	− 0.140	Spearman's correlation coef.
	.850	.000*	.053	.477	.186	*p* value

*Correlation is significant at the .05 significance level

$$\text{Estimated MOA} = 78.239 - 4.525 \, \text{Factor (PLA)[d]} - 13.834 \text{Factor(PLA)[g]}$$
$$- 1.176 \, \text{Factor (VOW)2} - 0.015 \text{VOT} - 0.059 \text{SLP}$$
$$(4)$$

Factor(PLA)[d] and Factor(PLA)[g] are the dummy variables associated with the categories of the place of articulation (meaning that if Factor(PLA)[d] = Factor(PLA)[g] = 0 then PLA = bilabial; if Factor(PLA)[d] = 1 and Factor(PLA)[g] = 0, then PLA = dental; if Factor(PLA)[d] = 0 and Factor(PLA)[g] = 1, then PLA = velar); Factor(VOW)2 is the dummy variable associated with the vowel context (if Factor(VOW)2 = 0 then VOW = vowel context 1; if Factor(VOW)2 = 1 then VOW = vowel context 2).

Running the ANOVA methodology for likelihood ratio test comparisons with a null model without the PLA factor and the proposed model as arguments shows that there is a significant drop in deviance when the PLA factor is considered ($X^2(2)$ = 10.071, *p* = .006). Taking into account the results presented in Table 12, it can be said that the place of articulation seems to affect MOA (which is itself a percentage) lowering it by about 4.5 ± 3.7% (standard errors), when the place of articulation is dental, and lowering it about 13.8 ± 4.4% (standard errors), when the place of articulation is velar. However, considering the confidence intervals via Wald approximations, it can be concluded that the only significant

difference between places of articulation is between bilabial and velar positions. No other factors seem to have significant effects on MOA, according to the same ANOVA methodology, a conclusion that agrees with the confidence intervals via Wald approximations, given in Table 12.

The confidence intervals via Wald approximations are also provided for the parameters of the stochastic part that includes only two terms: the difference between the individual's intercept and that of the population average (for which CI = (6.134; 29.391)) and the term allowing for random scatter of the individual's measures around their particular intercept or baseline (for which CI = (10.697; 14.747)).

The assumptions of normality regarding the residuals of the model were validated: the scatter plot representing standard fitted values vs. residuals for the estimated model suggests no obvious deviations from homoscedasticity, and a small bias; the normal QQ-plot for the residuals of the estimated model suggests a deviation from normality that is not significant.

3.7 Voicing classification

Results of voicing classification (shown in Table 13), based on the mean relative amplitude of the oral airflow signal [MOA (%)], revealed that 57% of stops were weakly voiced ([b]: 62%; [d]: 60%; [g]: 47%).

The mean of phone(1−2)% and phone(2−3)% relative amplitudes of the oral flow were not significantly different between places of articulation or vowel contexts.

Table 10 Summary of results from the correlation analysis when different places of articulation were considered separately: identification of the place of articulation for which the correlation is significant at the .05 significance level, considering the Spearman rank-based correlation test

	RLS	STP	A12	A23	MOA
SLP	[b]; [d]				
VOT	[b]	[b]; [d]; [g]	[d]; [g]	[g]	[g]
STP	[d]; [g]	−	[d]		[d]

Table 11 Random effects results

Groups	Variance	Std. Dev.
Speaker (Intercept)	186	13.64
Residual	166	12.89

Table 12 Fixed effects results

	Estimate	Std. Error	*t* value	2.5% LB of CI	97.5% UB of CI
Intercept *	78.239	15.349	5.097	48.952	107.261
Factor (PLA)[d]	− 4.525	3.657	− 1.238	− 11.489	2.555
Factor (PLA)[g] *	− 13.834	4.396	− 3.147	− 22.355	− 5.470
Factor (VOW)2	− 1.176	3.391	− 0.347	− 7.646	5.374
VOT	− 0.015	0.072	− 0.204	− 0.154	0.123
SLP	− 0.059	0.186	− 0.319	− 0.410	0.306

*Significant coefficients considering the confidence intervals via Wald approximation, at the .05 confidence level

Weak or strong voicing during voiced stop closures is speaker-specific as shown in Table 13.

4 Discussion
Results presented in this paper provide new evidence towards the view that stop voicing is not phonologically active in EP as previously suggested for this language, for German and English [6, 15, 40]. Only around 40% of this study's stops were strongly voiced (supporting H1.1), a percentage which is even lower than the one previously reported (around 60%) for German and English [6, 15].

The decrease in the amplitude of oral airflow (i.e. the amplitude of voicing) relative to the adjacent vowel (supporting H1.2), observed in EP stops is very similar to what has been previously reported for German, and suggests that the laryngeal feature of contrast in EP is not [*voice*], as recently shown for German and English [6]. The laryngeal contrast in EP could be between stops with no laryngeal specification and those specified as [*spread glottis*], evidenced by the low oral airflow amplitude oscillations (observed in more than 95% of the tokens) which are very likely generated by mucosal oscillations for a spread glottis, as initially hypothesised—H1.3 [18].

4.1 The aerodynamics of stops
The SLP results (significantly different values for the three places of articulation, supporting H2; t = 8.635, df = 2, p = .013, Kruskal-Wallis test) could be related to those previously presented, i.e. the idea that glottal area and resistance affect peak oral airflow as supported by findings of higher peak oral airflow for men than women, adults than children and voiceless than voiced consonants [22, 29]. Greater burst energy for a stop can be expected in two cases [11]: when there is a relatively smaller amount of linguopalatal contact, resulting in a fast release as opposed to a larger contact area; when the airflow is greater at the release, which is presumably due to a greater air pressure behind the constriction immediately before the release. These are

aerodynamic and voicing mechanisms that account for the special behaviour of EP stops (evidence supporting hypothesis H2).

It has been hypothesised (H1.3) that the oscillations seen in the oral airflow would produce no perceptible acoustic excitation, and its low average amplitude and absence of high frequency content, is likely to be generated by mucosal oscillations for a *spread glottis*. This hypothesis is firmly grounded in the aerodynamic mechanisms of voiced obstruent production described above. Analysis of our raw absolute oral airflow data during stop closures (see average values for each speaker in Table 7) led us to conclude that only 2/162 (1%) tokens were produced with an average glottal airflow (median OA between 72.6 cm^3/s and 105.5 cm^3/s for male speakers and between 28.8 cm^3/s and 70.8 cm^3/s for female speakers) above the minimal glottal airflow required to initiate phonation (supporting H1.2), thus

Table 13 Stops classified (based on oral airflow) as having weak and strong voicing, applying the 70% threshold. Stops are classified using mean phone1-2–phone2-3 values

70% threshold classification	[b] (%)	[d] (%)	[g] (%)	[b, d, g] (%)
♂SP1				
Weak voicing	50	47	7	34
Strong voicing	50	53	93	66
♂SP2				
Weak voicing	79	62	100	76
Strong voicing	21	38	0	24
♀SP3				
Weak voicing	19	22	14	18
Strong voicing	81	78	86	82
♀SP4				
Weak voicing	100	94	88	94
Strong voicing	0	6	13	6
♂♀All				
Weak voicing	62	60	47	57
Strong voicing	38	40	53	43

resulting in what has been previously classified as a *devoiced* stop [38]. We have, nevertheless, adopted a new and more general criterion of classification, denoting changes in laryngeal articulation during stop closures as weak voicing. Even so, 57% of stops in our database ([b] – 62 %; [d] – 60 %; [g] – 47 %) were produced with *weak voicing* (i.e. *devoiced*), supporting H1.1.

Aerodynamic variables (i.e. oral and glottal airflow) have an impact on the biomechanics of vocal fold vibration (i.e. onset and offset of vibration, opening and closing times of the vibratory cycle) and the parameters (SLP, VOT, RLS, STP, OA1, OA2, OA3, A12, A23 and MOA) obtained from real speech can be used to understand the effect of variability in stop production.

A negative slope in the amplitude of voicing during closure has been observed for several languages, but it is not possible to compare our absolute oral airflow values with those reported in the literature because authors typically report peak values [11, 16]. The most recent aerodynamic voicing contrast study [49] unfortunately does not report any oral airflow values despite having clearly collected detailed data at multiple time points for this variable.

The English and German acoustic data previously analysed [6] showed that, on average, 35–60% of all closure intervals measured in these studies was *weakly voiced* (*devoiced*). Data presented in this paper has a very different nature: The relative airflow amplitudes were analysed to derive a new classification of voicing. However, it is plausible to assume that the tokens classified as having *strong voicing* in this paper would have been labelled as *voiced* by Beckman et al. [6], so having 43% of the tokens in this study with *weak voicing* is comparable to 35–60% reported for Germanic languages. This is further supported by a previous acoustic phonetics study [40] that shows that German and Portuguese have very similar voicing profiles (going back to our first hypotheses).

4.2 Contextual effects on stops' production
Here, we discuss evidence supporting hypothesis H2, i.e. that place and vowel related variations are mediated by EP language-specific phonology.

European Portuguese results presented in this paper have shown that negative VOT values (duration of voicing during closure) were significantly different for place of articulation (supporting H2), with longer VOT values for velar stops than for bilabial stops and dental stops presenting the shortest average VOT values. An articulatory and aerodynamic-based account of VOT place-related variations [18] would have predicted that $VOT_{[g]} > VOT_{[d]} > VOT_{[b]}$, but studies such as ours, that have looked at the influence of place on closure

voicing have gradually supported a novel hypothesis that "language-specific rules mediate the phonetic outputs" ([2], p. 69).

Previous English [15, 28] and German [25] studies [five British English (BE) speakers [15], six AE speakers [28] and six German speakers [25]] have reported the following distinct orders: AE [28] – $VOT_{[b]} > VOT_{[d]} > VOT_{[g]}$; BE [15] – $VOT_{[g]} = VOT_{[d]} > VOT_{[b]}$; German [25] – $VOT_{[g]} > VOT_{[d]} > VOT_{[b]}$. As previously reported for French, EP "variations in VOT cannot be systematically correlated to variations in closure duration or to the duration of abduction gesture" ([2], p. 75), a claim that is supported by the fact that there were no significant differences between places of articulation for the values of RLS and STP (not supporting research hypothesis H2), and the following order, as observed in the present study: $VOT_{[g]} > VOT_{[b]} > VOT_{[d]}$.

Evidence about contextual effects for stops in running speech is contradictory and study dependent, so the fact that the contextual vowels did not affect the oral airflow measurements we made (there was no support for hypothesis H2), is in line with the lack of a systematic influence found in the literature [18, 21, 22, 37]. Despite the reported vowel-dependent results regarding obstruent voicing for some languages, there are no clear patterns that have been observed in previous studies of EP stops [1, 18, 33, 40]. Nevertheless, previous studies have reported low/high vowel contextual differences, so our corpus included vocalic contexts representing the different vowel heights used in EP. The vowels in the initial syllable of the first word in the sequence Y were divided into two groups according to their height: group 1—/i, ɨ, u, e, o/; group 2—/ɛ, ɔ, ɐ, a/. A rich variety of phonetic contexts using real EP words was selected to study the most relevant phoneme variants, and fully describe the aerodynamic properties of EP stops. Words were chosen and sentences were built following language-specific phonological rules. The corpus was designed to include as many words and cross-word contexts that could elicit devoicing (or help maintain voicing) in stops.

Martins's et al. [34] EP results regarding contextual effects were also not consistent across all stops. We believe this could be attributed to the instructions given to the speakers on how to produce the tokens, resulting in overarticulation during VC and CV transitions. One also has to be cautious about interpreting their results because the authors recognise "... the secondary role of coarticulation..." ([34], p. 930) in their study. Martins et al. [34] specifically instructed their speakers to carefully articulate each syllable "resulting in overarticulation during VC and CV transitions". We have not done so: We just asked our speakers to read each sentence as naturally as possible (this is

mentioned in our Method section), producing segmental environments that occur in EP, which was one of the aims of this corpus design.

Some of the carrier phrases elicited "word-final" stops before vowel-initial words. Especially when they appear before non-stress-initial words (e.g. [i.lɨ.gal ˈmẽt] or [ɐˈli]), we would expect the target consonants to be resyllabified in syllable-initial position. Regardless of the degree of resyllabification, word-final stops will be heavily coarticulated with the following word-initial consonant unless there was a strong prosodic boundary inserted after the elicitation item. There have, indeed, been a large number of tokens where resyllabification occurred. Resyllabification in EP occurs across the intonational phrase, i.e. it is not bound by the phonological word [53]. This produced segmental environments that occur in EP, which was one of the aims of this corpus design (including relative phone2/phone3 airflow derived parameters calculated both within word and cross-word when resyllabification occurred). We also included in our corpora isolated words and words produced in the carrier sentence <Diga X por favor>, where the phoneme following the final word stops was /p/ [33].

Words were chosen and sentences were built following language-specific phonological rules, for example: vowels /ɐ/ and /u/ can occur in the tonic syllable; vowels /ɐ/, /ɨ/ and /u/ can occur before and after the tonic syllable; the stops can all occur in initial and medial positions. Phonetically, any stop can be found in word final position as a consequence of deletion of unstressed vowels.

A fundamental issue with the variant vocalic environments created by the asymmetrical choice of stimuli is the impact that these differences had on the methodology. The major metric reported in this paper depends critically on the oral airflow amplitude. Yet we should expect this value to vary intrinsically between vowels of different qualities, all other factors being equal. Some stops are produced in the context of low vowels, produced with a lowered jaw and presumably a greater mean oral aperture, while others a produced adjacent to mid-high vowels, produced with a more constricted palatal gesture. Neither was backness controlled for, a factor which will have a large impact on pharyngeal constriction degree, and therefore also potentially oral airflow amplitude.

There were no significant differences for phone(2–3)% relative amplitudes of the oral flow considering, place of articulation and vowel context (i.e. results did not support H2), which is inconsistent with traditional accounts of the relationship between devoicing and place of articulation [38]. However, some of the patterns reported before were also observed for our tokens, but the results were not statistically significant:

the place of articulation affected the values of MOA; the duration of voicing during closure was longer for velar stops than for bilabial stops and dental stops presented the shortest values.

Generalised linear mixed effects models were used to test for the fixed effects of VOT, SLP and the factors PLA and VOW (without interaction terms) on the mean oral airflow. By-speaker variation, considered as a random effect with random intercept, was found to explain a considerable part of the variability in mean oral airflow. No significant heteroscedasticity nor deviations from normality were found in the analysis of the residuals of the proposed model.

5 Conclusions

One of the key differences of the results presented in this paper, from previous aerodynamic studies of stops, is the use of real words in grammatically feasible carrier sentences. The differences between the use of nonsense words and real words in obstruent production studies have been clearly shown: As one "moves" from less realistic to more realistic conditions, many of the patterns in data are less distinctive. In this work, we used a representative variety of phonetic environments resulting in phonetic phenomena that realistically occur in EP.

This paper contributes new aerodynamic evidence towards capturing underlying laryngeal settings and phonetic properties of voicing contrast for an understudied language (EP) based on previous theoretical and experimental grounding on voicing in obstruents.

The empirical evidence presented in this paper (based on the oral airflow signal) consubstantiates the claim that voicing found in Portuguese stops is the result of the same phonetic process observed for German and English. It could be hypothesised from the Portuguese results that the high percentages of *weakly voiced* stops ($> 50\%$) are a consequence of passive voicing, and that low amplitude oscillations of oral airflow during closure supports the view that the feature of contrast in Portuguese is *privative* [*spread glottis*].

A fundamental issue with the current work is the limited number of speakers and spoken exemplars. In addition, the variant vocalic environments (three very different segmental environments specified in Section 2.1) created by the asymmetrical choice of stimuli had an impact on the methodology.

Aerodynamic variables measurable from real speech condition the mechanics of vocal fold vibration (i.e. onset and offset of vibration, opening and closing quotients of the vibratory cycle), so future work could incorporate these in realistic speech models. They can also be used to understand the effect of variability in stop production on the performance of stop detectors.

6 Appendix

Table 14 Corpus of voiced stops: words without a frame sentence

Stop	Word	IPA
/b/	<bala>	['ba.lɐ]
	<juba>	['ʒu.bɐ]
	<cabe>	['ka.b]
/d/	<dava>	['da.vɐ]
	<nada>	['na.dɐ]
	<pode>	['pɔ.d]
/g/	<gato>	[' ga.tu]
	<paga>	['pa.gɐ]
	<pague>	['pa.g]

Table 15 Corpus of voiced stops: words in frame sentences

Stop	Sentence	IPA
/b/	<Diga cabe algo por favor>	['di.gɐ 'cab 'al.gu 'pur fe'voɾ]
	<Diga cabe areia por favor>	['di.gɐ 'cab e'ɾej.ɐ 'pur fe'voɾ]
	<Diga cabe ilegalmente por favor>	['di.gɐ 'cab i.ɫi.gal'mẽt 'pur fe'voɾ]
	<Diga cabe hirto por favor>	['di.gɐ 'cab 'iɾ.tu 'pur fe'voɾ]
/d/	<Diga pode alucinar por favor>	['di.gɐ 'pɔd ɐ.lu.si'naɾ 'pur fe'voɾ]
	<Diga pode arcar por favor>	['di.gɐ 'pɔd ɐɾ'kaɾ 'pur fe'voɾ]
	<Diga pode iluminar por favor>	['di.gɐ 'pɔd i.lu.mi'naɾ 'pur fe'voɾ]
	<Diga pode ironizar por favor>	['di.gɐ 'pɔd i.ɾu.ni'zaɾ 'pur fe'voɾ]
/g/	<Diga pague ali por favor>	['di.gɐ 'pag e'li 'pur fe'voɾ]
	<Diga pague artesão por favor>	['di.gɐ 'pag ɐɾ.t'zẽw 'pur fe'voɾ]
	<Diga pague ilegalmente por favor>	['di.gɐ 'pag i.ɫi.gal'mẽt 'pur fe'voɾ]
	<Diga pague irmã por favor>	['di.gɐ 'pag iɾ'mẽ 'pur fe'voɾ]
/b/	<Diga cabe igualar por favor>	['di.gɐ 'kab i.guɐ'laɾ 'pur fe'voɾ]
	<Diga cabe agradar por favor>	['di.gɐ 'kab ɐ.gɾɐ'daɾ 'pur fe'voɾ]
/d/	<Diga pode igualar por favor>	['di.gɐ 'pɔd i.guɐ'laɾ 'pur fe'voɾ]
	<Diga pode agradecer por favor>	['di.gɐ 'pɔd ɐ.gɾɐ.d'seɾ 'pur fe'voɾ]
/g/	<Diga pague igual por favor>	['di.gɐ 'pag i'gual 'pur fe'voɾ]
	<Diga pague agora por favor>	['di.gɐ 'pag e'gɔ.ɾɐ 'pur fe'voɾ]
/b/	<Diga cabe aninhado por favor>	['di.gɐ 'kab ɐ.ni'ɲa.du 'pur fe'voɾ]
	<Diga cabe imóvel por favor>	['di.gɐ 'kab i'mɔ.vɛl 'pur fe'voɾ]
/d/	<Diga pode amar por favor>	['di.gɐ 'pɔd e'maɾ 'pur fe'voɾ]
	<Diga pode imaginar por favor>	['di.gɐ 'pɔd i.mɐ.ʒi'naɾ 'pur fe'voɾ]
/g/	<Diga pague amanhã por favor>	['di.gɐ 'pag a.ma'ɲẽ 'pur fe'voɾ]
	<Diga pague inicialmente por favor>	['di.gɐ 'pag i.ni.sial'mẽt 'pur fe'voɾ]
/b/	<Diga cabe agitar por favor>	['di.gɐ 'kab ɐ.ʒi'taɾ 'pur fe'voɾ]
	<Diga cabe ejetar por favor>	['di.gɐ 'kab i.ʒɛ'taɾ 'pur fe'voɾ]
/d/	<Diga pode ajudar por favor>	['di.gɐ 'pɔd ɐ.ʒu'daɾ 'pur fe'voɾ]
	<Diga pode higienizar por favor>	['di.gɐ 'pɔd i.ʒiɛ.ni'zaɾ 'pur fe'voɾ]
/g/	<Diga pague ajudando por favor>	['di.gɐ 'pag ɐ.ʒu'dã.du 'pur fe'voɾ]
	<Diga pague Egipto por favor>	['di.gɐ 'pag i'ʒi.tu 'pur fe'voɾ]
/b/	<Diga bala por favor>	['di.gɐ 'ba.lɐ 'pur fe'voɾ]

Table 15 Corpus of voiced stops: words in frame sentences *(Continued)*

Stop	Sentence	IPA
	<Diga juba por favor>	[ˈdi.ɡe ˈʒu.bɐ ˈpuɾ feˈvoɾ]
	<Diga cabe por favor>	[ˈdi.ɡe ˈkab ˈpuɾ feˈvoɾ]
/d/	<Diga dava por favor>	[ˈdi.ɡe ˈda.ve ˈpuɾ feˈvoɾ]
	<Diga nada por favor>	[ˈdi.ɡe ˈna.de ˈpuɾ feˈvoɾ]
	<Diga pode por favor>	[ˈdi.ɡe ˈpɔd ˈpuɾ feˈvoɾ]
/g/	<Diga gato por favor>	[ˈdi.ɡe ˈɡa.tu ˈpuɾ feˈvoɾ]
	<Diga paga por favor>	[ˈdi.ɡe ˈpa.ɡe ˈpuɾ feˈvoɾ]
	<Diga pague por favor>	[ˈdi.ɡe ˈpaɡ ˈpuɾ feˈvoɾ]

Authors' contributions

LJ and MC analysed and interpreted the data. LJ was a major contributor in writing the manuscript. Both authors read and approved the final manuscript.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]School of Health Sciences (ESSUA) and Institute of Electronics and Informatics Engineering of Aveiro (IEETA), University of Aveiro, 3810-193 Aveiro, Portugal. [2]Department of Mathematics (DMat) and Centre of Research and Development in Mathematics and Applications (CIDMA), University of Aveiro, 3810-193 Aveiro, Portugal.

References

1. Abdelli-Beruh, N. B. (2004). The Stop voicing contrast in French sentences: Contextual sensitivity of vowel duration, closure duration, voice onset time, stop release and closure voicing. *Phonetica*, 61(4), 201–219. https://doi.org/10.1159/000084158
2. Abdelli-Beruh, N. B. (2009). Influence of place of articulation on some acoustic correlates of the stop voicing contrast in Parisian French. *J. Phon.*, 37(1), 66–78. https://doi.org/10.1016/j.wocn.2008.09.002
3. Abramson, A. S., & Whalen, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *J. Phon.*, 63, 75–86. https://doi.org/10.1016/j.wocn.2017.05.002
4. Awan, S. N., Novaleski, C. K., & Yingling, J. R. (2013). Test-retest reliability for aerodynamic measures of voice. *J. Voice*, 27(6), 674–684. https://doi.org/10.1016/j.jvoice.2013.07.002
5. Bates, D., Mächler, M., Bolker B., Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015) https://doi.org/10.18637/jss.v067.i01
6. Beckman, J., Jessen, M., & Ringen, C. (2013). Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics*, 49(02), 259–284. https://doi.org/10.1017/S0022226712000424
7. Boersma, P., Praat, a system for doing phonetics by computer. *Glot International* **5**(9/10), 341–345 (2001)
8. Brinca, L., Araújo, L., Nogueira, P., & Gil, C. (2016). Voice onset time characteristics of voiceless stops produced by children with European Portuguese as mother tongue. *Ampersand*, 3, 137–142. https://doi.org/10.1016/j.amper.2016.06.006
9. Brunner, J., Fuchs, S., Perrier, P., Supralaryngeal control in Korean velar stops. *J. Phon.* **39**(2), 178–195 (2011)
10. Bucella, F., Hassid, S., Beeckmans, R., Soquet, A., & Demolin, D. (2000). Pression sous-glottique et débit d'air buccal des voyelles en français. In *XXIIIèmes Journées d'Etude sur la Parole*, (pp. 449–452). Aussois.
11. Cho, T., Jun, S., Ladefoged, P., Acoustic and aerodynamic correlates of Korean stops and fricatives. *J. Phon.* **30**(2), 193–228 (2002)
12. Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *J. Phon.*, 27(2), 207–229. https://doi.org/10.1006/jpho.1999.0094
13. Cho, T., Whalen, D. H., & Docherty, G. (2019). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *J. Phon.*, 72, 52–65. https://doi.org/10.1016/j.wocn.2018.11.002
14. Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *J. Phon.*, 54, 35–50. https://doi.org/10.1016/j.wocn.2015.09.003
15. G.J. Docherty, *The Timing of Voicing in British English Obstruents* (Foris, Berlin, 1992)
16. Emanuel, F. W., & Counihan, D. T. (1970). Some characteristics of oral and nasal air flow during plosive consonant production. The Cleft Palate Journal, 7(1), 249–260. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5266335
17. Ernestus, M., H. Baayen, Paradigmatic effects in auditory word recognition: The case of alternating voice in Dutch. *Lang. Cogn. Process.* **22**(1), 1–24 (2007)

18. A. Esposito, On vowel height and consonantal voicing effects: data from Italian. *Phonetica* **59**(4), 197–231 (2002) https://doi.org/10.1159/000068347

19. Ghosh, P. K., & Narayanan, S. S. (2009). Closure duration analysis of incomplete stop consonants due to stop-stop interaction. *The Journal of the Acoustical Society of America*, *126*(1), EL1–EL7. https://doi.org/10.1121/1.3141876

20. Helgason, P., Ringen, C., Voicing and aspiration in Swedish stops. *J. Phon.* **36**(4), 607–628 (2008)

21. Higgins, M. B., Netsell, R., & Schulte, L. (1994). Aerodynamic and electroglottographic measures of normal voice production: intrasubject variability within and across sessions. Journal of Speech and Hearing Research, 37(1), 38–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8170128

22. Higgins, M. B., Netsell, R., & Schulte, L. (1998). Vowel-related differences in laryngeal articulatory and phonatory function. Journal of Speech, Language, and Hearing Research : JSLHR, 41(4), 712–724. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9712121

23. Hottinger, D. G., Tao, C., & Jiang, J. J. (2007). Comparing Phonation Threshold Flow and Pressure by Abducting Excised Larynges. The Laryngoscope, 117(9). 1695–1699. https://doi.org/10.1097/MLG.0b013e3180959e38.

24. Iskarous, K., Mooshammer, C., Hoole, P., Recasens, D., Shadle, C. H., Saltzman, E., & Whalen, D. H. (2013). The coarticulation/invariance scale: Mutual information as a measure of coarticulation resistance, motor synergy, and articulatory invariance. *The Journal of the Acoustical Society of America*, *134*(2), 1271–1282. https://doi.org/10.1121/1.4812855

25. Jessen, M., *Phonetics and phonology of tense and lax obstruents in German* (John Benjamins Publishing Company, Amsterdam/Philadelphia, 1998)

26. Jessen, M., Roux, J., Voice quality differences associated with stops and clicks in Xhosa. *J. Phon.* **30**, 1–52 (2002)

27. Jiang, J. J., & Tao, C. (2007). The minimum glottal airflow to initiate vocal fold oscillation. *The Journal of the Acoustical Society of America*, *121*(5), 2873–2881. https://doi.org/10.1121/1.2710961

28. Keating, P. (1984). Phonetic and Phonological Representation of Stop Consonant Voicing. *Language*, *60*(2), 286. https://doi.org/10.2307/413642

29. Klatt, D. H., Stevens, K. N., & Mead, J. (1968). Studies of articulatory activity and airflow during speech. *Ann. N. Y. Acad. Sci.*, *155*(1), 42–55. https://doi.org/10.1111/j.1749-6632.1968.tb56748.x

30. Koenig, L.L., Lucero, J.C., Stop consonant voicing and intraoral pressure contours in women and children. *J. Acoust. Soc. Am.* **123**(2), 1077–1088 (2008)

31. Koenig, L.L., Mencl, W.E., Lucero, J.C., Multidimensional analyses of voicing offsets and onsets in female speakers. *J. Acoust. Soc. Am.* **118**(4), 2535–2550 (2005)

32. Lisker, L. (1986). "Voicing" in English: a catalogue of acoustic features signaling /b/ versus /p/ in Trochees. *Lang. Speech*, *29*(1), 3–11. https://doi.org/10.1177/002383098602900102

33. Lousada, M., Jesus, L. M. T., & Hall, A. (2010). Temporal acoustic correlates of the voicing contrast in European Portuguese stops. *J. Int. Phon. Assoc.*, *40*(3), 261–275. https://doi.org/10.1017/S0025100310000186

34. Martins, P., Carbone, I., Pinto, A., Silva, A., Teixeira, A., European Portuguese MRI based speech production studies. *Speech Comm.* **50**(11–12), 925–952 (2008)

35. Möbius, B., Corpus-based investigations on the phonetics of consonant voicing. *Folia Linguistica* **38**(1–2), 5–26 (2004)

36. Mücke, D., Hermes, A., & Cho, T. (2017). Mechanisms of regulation in speech: Linguistic structure and physical control system. *J. Phon.*, *64*, 1–7. https://doi.org/10.1016/j.wocn.2017.05.005

37. Netsell, R., Lotz, W. K., DuChane, A. S., & Barlow, S. M. (1991). Vocal tract aerodynamics during syllable productions: Normative data and theoretical implications. *J. Voice*, *5*(1), 1–9. https://doi.org/10.1016/S0892-1997(05)80157-2

38. Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In N. Y. S. Verlag (Ed.), The Production of Speech (pp. 189–216).

39. Ohala, J.J., Riordan, C., in *Speech communication papers*, ed. by J. Wolf, D. H. Klatt. Passive vocal tract enlargement during voiced stops (Acoustical Society of America, New York, 1979), pp. 89–92

40. Pape, D., & Jesus, L. M. T. (2015). Stop and fricative devoicing in European Portuguese, Italian and German. *Lang. Speech*, *58*(2), 224–246. https://doi.org/10.1177/0023830914530604

41. Pape, D., Mooshammer, C., Hoole, P., Fuchs, S., in *Speech production: models, phonetic processes, and techniques*, ed. by J. Harrington, M. Tabain. Devoicing of word-initial stops: A consequence of the following vowel? (Psychology Press, New York, 2006)

42. Pinho, C. M. R., Jesus, L. M. T., & Barney, A. (2012). Weak voicing in fricative production. *J. Phon.*, *40*(5), 625–638. https://doi.org/10.1016/j.wocn.2012.06.002

43. Pinho, C. M. R., Jesus, L. M. T., & Barney, A. (2013). Aerodynamic measures of speech in unilateral vocal fold paralysis (UVFP) patients. *Logopedics, Phoniatrics, Vocology*, *38*(1), 19–34. https://doi.org/10.3109/14015439.2012.696138

44. Recasens, D., Espinosa, A., An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *J. Acoust. Soc. Am.* **125**, 2288–2298 (2009)

45. Regner, M. F., Tao, C., Zhuang, P., & Jiang, J. J. (2008). Onset and Offset Phonation Threshold Flow in Excised Canine Larynges. The Laryngoscope, 118(7), 1313–1317. https://doi.org/10.1097/MLG.0b013e31816e2ec7.

46. Ringen, C., & van Dommelen, W. A. (2013). Quantity and laryngeal contrasts in Norwegian. *J. Phon.*, *41*(6), 479–490. https://doi.org/10.1016/j.wocn.2013.09.001

47. Rothenberg, M., *Breath-Stream Dynamics of Simple-Released-Plosive Production* (Karger, Basel, 1968)

48. Shadle, C.H., in *The Handbook of Phonetic Sciences*, 2nd edn., ed. by W. J. Hardcastle, J. Laver, F. E. Gibbon. The Aerodynamics of Speech (Blackwell, Chichester, 2010), pp. 39–80

49. Solé, M. J. (2018). Articulatory adjustments in initial voiced stops in Spanish, French and English. *J. Phon.*, *66*, 217–241. https://doi.org/10.1016/j.wocn.2017.10.002

50. Stathopoulos, E. T., & Weismer, G. (1985). Oral airflow and air pressure during speech production: a comparative study of children, youths and adults. Folia Phoniatrica, 37(3–4), 152–159. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4054770

51. Stephens, J. D. W., & Holt, L. L. (2011). A standard set of American-English voiced stop-consonant stimuli from morphed natural speech. *Speech Comm.*, *53*(6), 877–888. https://doi.org/10.1016/j.specom.2011.02.007

52. Storkel, H.L., A corpus of consonant–vowel–consonant real words and nonwords: Comparison of phonotactic probability, neighborhood density, and consonant age of acquisition. *Behav. Res. Methods* **45**(4), 1159–1167 (2013) https://doi.org/10.3758/s13428-012-0309-7

53. Vigário, M., Freitas, M., Frota, S., Grammar and frequency effects in the acquisition of prosodic words in European Portuguese. *Lang. Speech* **49**(2), 175–203 (2006)

54. Winn, M. B., Chatterjee, M., & Idsardi, W. J. (2013). Roles of voice onset time and F0 in stop consonant voicing perception: effects of masking noise and low-pass filtering. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1097–1107. https://doi.org/10.1044/1092-4388(2012/12-0086)

55. Zajac, D., in *The Handbook of Speech Production*, ed. by M. Redford. Velopharyngeal function in speech production: some developmental and Structural considerations (Wiley-Blackwell, Malden, 2015), pp. 109–130

Publisher's Note