The role of vocal tract and subglottal resonances in producing vocal instabilities

Laura Wade, Noel Hanna, John Smith,[a)] and Joe Wolfe
School of Physics, University of New South Wales, Sydney, New South Wales 2052, Australia

(Received 9 May 2016; revised 6 February 2017; accepted 6 February 2017; published online 7 March 2017)

During speech and singing, the vibrating vocal folds are acoustically loaded by resonant ducts upstream (the trachea) and downstream (the vocal tract). Some models suggest that the vocal fold vibration (at frequency f_o) is more stable at frequencies below that of a vocal tract resonance, so that the downstream load is inertive (mass-like). If so, vocal fold vibration might become unstable when f_o and resonance frequencies "cross over" and the load varies rapidly in phase and magnitude. In one experiment, singers produced a slow diphthong at constant pitch, thus shifting the first tract resonance R1 across fixed f_o. In another, pitch glides took f_o across the tract and subglottal resonances. Few instabilities occurred when singers could change lip geometry and thus alter R1. This suggests that avoiding resonance crossings can aid vibrational stability. In experiments in which R1 was constrained using a mouth ring, instabilities occurred at frequencies above R1. When subjects sang into an acoustically infinite pipe, which provided a purely resistive load at the lips, R1 was eliminated. Here, instabilities were reduced and concentrated near the lower limit of the head voice.
 [http://dx.doi.org/10.1121/1.4976954]

[ZZ] Pages: 1546–1559

I. INTRODUCTION

During singing or speech, the vibrating vocal folds are loaded by both an upstream and a downstream impedance which, several models suggest, can play an important role in the self-sustained process that converts some of the steady airflow into an acoustic current with frequency f_o (e.g., Fletcher, 1993). Some models (e.g., Titze, 1988, 2008) suggest that vocal fold vibration is "enhanced," or at least more stable, if at least one of these two loads is inertive. In this scenario, vocal fold vibration could become less stable or even unstable when the impedance of a load is compliant, rather than inertive. Considering each resonance in isolation, the impedance of each of the loads (upstream and downstream) will change from inertive to compliant when f_o changes from below a resonance frequency to above that frequency. Thus, an instability might occur if f_o changes from below f_{R1} to above f_{R1}, where f_{R1} denotes the frequency of the first resonance R1 of the supraglottal tract (Titze *et al.*, 2015). Similarly, an instability could also occur if a transition occurs from below f_{Sg1} to above f_{Sg1}, where f_{Sg1} denotes the frequency of the first resonance of the subglottal tract. Instabilities might also occur if a harmonic falls near the fundamental or higher resonances of the tract. Since the acoustic reactance (imaginary component of acoustic impedance) is generally small other than in close proximity to the tract resonances, another potential hypothesis is that an increase in the magnitude of the reactance close to the resonance, rather than its sign, may trigger instabilities. These hypotheses about instability being associated with tract resonances are not easy to test because a knowledge of both f_{Sg1} and f_{R1} is required.

Normal adult speech is unsuitable for testing the importance of the acoustic load(s) in maintaining stable vocal fold vibration because f_o usually lies in the range of 80–250 Hz. This is well below f_{Sg1} (~600 Hz, e.g., Lulich *et al.*, 2012) and it is also below or at the lower end of the range of f_{R1} (typically 250–1100 Hz). Thus f_o crosses f_{Sg1} or f_{R1} only occasionally during speech. One solution to improve the overlap involves increasing the range of f_o by using singing instead of speech. The soprano voice is particularly suitable as the typical vocal range covers about 250–1000 Hz, which approximately coincides with the range of f_{R1}.

Soprano voices raise a new problem: using the voice itself to estimate f_{R1}, e.g., by linear prediction or inverse filtering, becomes increasingly inaccurate as f_o increases above 300 Hz (e.g., Monsen and Engebretson, 1983; Arroabarren and Carlosena, 2006). One possible alternative is to use the broadband signal produced in vocal fry and to measure the lowest formant frequency $F1$; this provides a reliable estimate of f_{R1} in the frequency range associated with speech (Swerdlin *et al.*, 2010). However, the reliability of using vocal fry to estimate f_{R1} for singing depends upon singers not altering their vocal tract geometries between the vocal fry measurement and the singing experiment. This is complicated because f_{R1} itself can be a function of f_o, such as when resonance tuning occurs (Joliveau *et al.*, 2004a; Henrich *et al.*, 2011).

Rather than increasing f_o so that it approaches f_{R1}, a different approach is to modify the resonant properties of the tract so that f_{R1} begins to approach f_o. The "ecological" way to change f_{R1} is to ask subjects to articulate different vowels (Rothenberg, 1981; Titze *et al.*, 2008; Zañartu *et al.*, 2011). Another technique involves lowering f_{R1} by effectively extending the tract by adding extension tubes of different length at the mouth (Sondhi, 1975). A third technique involves altering

[a)]Electronic mail: john.smith@unsw.edu.au

 0001-4966/2017/141(3)/1546/14/$30.00

the speed of sound within the tract by injecting different gas mixtures (Sundberg, 1981).

One possible approach to detect instabilities is to vary f_o smoothly (strictly a *portamento* although often called a *glissando*) and to examine what happens when f_o is in the frequency ranges that are thought to correspond to f_{R1} and f_{Sg1}. This is called a pitch glide. Alternatively, f_o could be kept constant while f_{R1} is altered continuously between the values for two vowels: a vowel glide. Both of these approaches have been used by Titze *et al.* (2008). In each case, f_{R1} was assumed to be equal to $F1$ found by an intermediate period of vocal fry during each pitch glide. For male voices an instability was found to occur in 54% of pitch glides if a crossover of f_o and $F1$ was predicted, compared with 35% where a crossover was not predicted. For females, however, there was no statistically significant difference whether crossovers were predicted or not. Titze and colleagues also found that 31% of instabilities occurred when f_o was within 100 Hz of $F1$. Their analysis depended upon the assumption that $F1$ measured during vocal fry was similar to f_{R1}, the resonance frequency of the supraglottal tract during singing, and that consequently the reactance of the supraglottal tract would switch between inertive and compliant when $f_o \sim F1$. While $F1 \sim f_{R1}$ has been shown to be a reasonable assumption for speech (Swerdlin *et al.*, 2010), sopranos (both trained and untrained) are known to employ resonance (formant) tuning at high pitch; they vary the mouth geometry so that f_{R1} matches f_o once f_o exceeds the value of f_{R1} in normal speech for that vowel (Sundberg, 1975; Sundberg and Skoog, 1997; Joliveau *et al.*, 2004a). The tendency of women singers to avoid combinations with $f_o > f_{R1}$ might consequently minimise the number of instabilities produced in pitch or vowel glides and partly explain this difference between male and female voices.

Zañartu *et al.* (2011) extended this approach by studying the motion of the vocal folds using laryngeal high-speed videoendoscopy during pitch glides. Vocal fold contact was monitored using an electroglottograph (EGG). The neck acceleration, radiated sound and oral volume velocity were also measured, and inverse filtering of the radiated sound was used to determine $F1$. Of three subjects studied with this setup, only one subject gave consistent results and was selected for a case study; the other two only sometimes produced instabilities. In that study the pitch glides and vowel glides as suggested by Titze *et al.* (2008) were executed where the mouth geometry was determined by the subject. The consistent subject gave instabilities of two sorts, which were attributed to source and acoustic load. Pitch glides on two vowels showed that about 33% of pitch instabilities occurred within one bandwidth of $F1$.

The largest and most dramatic change in the supraglottal impedance will occur when f_o crosses f_{R1}; the downstream load at f_o will then rapidly change between inertive and compliant. At frequencies sufficiently far from $f_o = f_{R1}$, the acoustic load will be a low magnitude impedance. It is thus possible that an instability might be triggered by an increase in the reactive component (or the increase in overall magnitude) as f_o approaches f_{R1}, and that the abrupt change produced by a crossing is not necessary. Thus, the load will become increasingly inertive as an increasing f_o approaches f_{R1}, and increasingly compliant as a decreasing f_o approaches f_{R1}. Previous studies have considered the bandwidth of the resonances and attributed instabilities that occurred within the bandwidth of a resonance as being a consequence of that resonance (Titze *et al.*, 2008; Zañartu *et al.*, 2011).

The aim of the present paper is to use several techniques to examine the possible role of resonances in causing vocal fold instability in sopranos. The study was conducted in three parts.

(i) *Unconstrained*. Vowel glides (smoothly varying f_{R1}) at constant f_o, and pitch glides (smoothly varying f_o) with a nominal vowel (i.e., f_{R1} would be constant in the absence of resonance tuning). This is essentially a repetition of the experimental approach of Titze *et al.* (2008) to allow comparison.

(ii) *Constrained*. Pitch glides with the mouth aperture constrained to a fixed shape using a small ring held between the lips of subjects. Because the lip aperture and jaw height are the chief control parameters used to vary f_{R1}, this reduces substantially the ability of subjects to vary f_{R1}.

(iii) *Resonance-free*. Pitch glides with subjects singing into an acoustically infinite pipe that presents a purely resistive load at the lips: this produces a very low reflection coefficient at the lips and effectively removes R1 without affecting the subglottal resonances.

The *unconstrained* measurements follow the experimental approach of Titze *et al.* (2008), with the addition of two refinements: the subglottal resonances of each subject were measured using an accelerometer and her vocal fold behaviour was monitored using an EGG. The experimental conditions also allowed comparisons to be made with f_{R1} measured previously under comparable conditions. The *constrained* measurements impose a condition that makes it much more difficult for the soprano to vary f_{R1}, because the lip aperture and jaw height are the principal parameters by which f_{R1} is varied. The *resonance-free* measurements aim to minimise the reactive component of the downstream acoustic load "seen" at the larynx. However, it is likely that the impedance seen through the (relatively narrow) glottis remains slightly inertive over much of the range, rather than purely resistive.

II. MATERIALS AND METHODS

A. Experimental apparatus

1. Measurement of subglottal resonances

The subglottal resonances f_{Sg1} and f_{Sg2} were measured using the method of Lulich *et al.* (2012)—see Fig. 1(a). An accelerometer (Hot Spot, K&K Sound Systems, Coos Bay, OR) with a diameter of 12 mm and thickness of 0.8 mm was attached to the neck just below the thyroid cartilage using a piece of very thin, transparent medical dressing (Nexcare Tegaderm Transparent dressing, 3M Health Care, St. Paul, MN) with a mass of approximately 0.5 mg. The frequency response of this low-mass accelerometer with and without the adhesive dressing was checked using an accelerometer with known characteristics (4394, Bruel & Kjaer, Denmark);

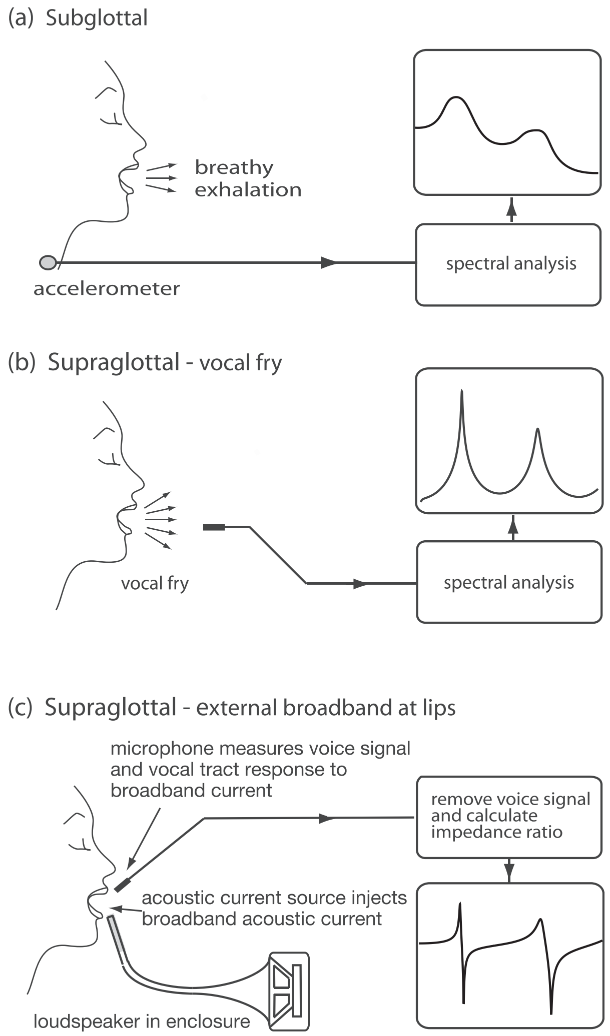

FIG. 1. Schematic diagram (not to scale) showing the three different techniques used in this paper to measure the vocal tract resonances.

above a weak resonance around 300 Hz, the frequency response in the range of interest for subglottal resonances (350–2000 Hz) was within ±1.2 dB. The spectrograms of the recorded signal from the accelerometer were analysed using PRAAT 6.0.17 (Boersma and Weenink, 2016). The sound outside the mouth was also recorded; this helped discern whether the supraglottal tract was responsible for any of the enhanced frequency bands observed in the accelerometer signal.

Although the spectrograms of the accelerometer signal were examined during each of the different vocal exercises described later [vowel glides at constant pitch, pitch glides on a neutral (Australian English) "er" vowel, and during vocal fry], the clearest results were found when singers performed a separate breathy exhalation, thereby exciting both supra- and sub-glottal tracts with a broad band signal.

2. Measurement of supraglottal resonances

In the present study, the reported supraglottal resonances were measured using two different techniques.

The first method used vocal fry, also called creak phonation or mechanism 0, in which the vocal folds undergo non-periodic motion (Hollien and Michel, 1968; Gobl, 1989). Spectral analysis of the sound then allows the formant frequency $F1$ to be estimated—see Fig. 1(b). The values of $F1$ thus measured using vocal fry have been found to be slightly higher, by an average of 50 Hz, than the resonance frequencies f_{R1} measured by broadband excitation at the lips (Epps *et al.*, 1997) when measurements are made during the same vocal gesture in speech (Swerdlin *et al.*, 2010). (Of course, the vocal fry measurement cannot be made whilst simultaneously singing.)

Vocal fry has been used to estimate the resonances of the vocal tract during singing in several studies (e.g., Miller *et al.*, 1997). There is a potential problem, however, because the resonance frequencies during singing at high pitch are known to increase (resonance tuning). Consequently, the values obtained from vocal fry are not necessarily identical to those produced during singing.

The second method used external broadband at the lips. This involves injecting a broadband current in parallel with the external radiation field. The loudspeaker generated a broadband current from 100 Hz to 4 kHz using a signal synthesised from sine waves spaced at 2.69 Hz. See Fig. 1(c) and Epps *et al.* (1997) for details. This method was used for the detailed measurements of f_{R1} as a function of f_o with lips constrained by a mouth ring—see Fig. 10(b). It was also used for the previously measured values of f_{R1} quoted in this paper.

3. Measurement of vocal fold contact

A dual-channel electroglottograph (model EG-2, Glottal Enterprises, Syracuse, NY) was used to monitor the vocal fold vibration throughout each experiment using an oscilloscope. Analysis of the recorded waveform allowed the identification of instabilities associated with laryngeal behaviour.

4. The "infinite" pipe

The acoustically "infinite" pipe used in this project is 197 m long with 26.2 mm internal diameter, and installed in the ceiling space of the building. The first section of 30 m is straight, and subsequent curves have a minimum radius of about 7 m. The pipe is sufficiently long so that visco-thermal losses at the walls reduce the amplitude of reflections from the other open end by at least 80 dB at the frequencies used in this experiment. Consequently, no acoustic resonances occur in the pipe and the input acoustic impedance is purely resistive with a characteristic impedance of 0.77 MPa s m^{-3}. When a subject seals her mouth around one end of this pipe, with the tongue kept low in the mouth as instructed, there are no large discontinuities in the vocal tract or at the lips, thus no strong reflections should occur, and so the vocal folds should then experience this approximately resistive load. (Calculations are reported below.) The measured DC resistance of the pipe is low (around 0.7 to 0.8 MPa s m^{-3}) and does not cause inconvenience to the subjects during exhalation and singing. In related experiments (Wade *et al.*, 2016), subjects similarly reported no inconvenience when singing sustained notes at a wide variety of pitches into the infinite pipe, however, they experienced various instabilities when R1 was suddenly introduced by rapidly pulling a plug from the side of the infinite pipe close to the lips.

5. Recordings

For most measurements the sound was recorded approximately 10 cm away from the mouth using a condenser microphone (RODE NT3, Sydney, Australia). For experiments when the infinite pipe was used, a 1/4 in. pressure-field microphone (4944A, Brüel & Kjær, Denmark) was located flush with the internal surface of the pipe 12 cm from the lips. The sound and the EGG signals were sampled at 44.1 kHz with 16-bit resolution using a FireWire audio interface (MOTU 828, Cambridge, MA). They were analysed using PRAAT 6.0.17 (Boersma and Weenink, 2016).

B. Subjects

Eight female singers volunteered for this study. All spoke Australian English, and had varying amounts of formal vocal training and singing experience as summarised in Table I. The experimental protocol was approved by the University's Human Ethics Committee.

C. Experimental protocol

Measurements were made in two sessions. In the first session, the subglottal resonances were measured using a breathy exhalation—see Fig. 1(a). A series of measurements (not reported herein) using an impedance head were then made. The impedance head was placed between the subject's lips and had the same outside diameter as the mouth ring and infinite pipe. This gave subjects the opportunity to become familiar with singing whilst their lips were constrained to the same fixed diameter as the mouth ring and infinite pipe.

In the second session the pitch and vowel glides were measured. Usually each glide was attempted three to four times in each pitch range; this ensured that at least two glides satisfactorily covered the desired range and could be included in the analysis. The values of $F1$ were measured using vocal fry. For all the pitch glide experiments, the target starting and finishing notes were prompted using a synthesised piano note at the required pitch. The vowel glides typically lasted 5 s, whereas the pitch glides typically lasted 15 s including a 3 s intermediate period of vocal fry. The speed of the pitch glides was thus around three semitones per second; this probably allowed singers sufficient time to adjust their resonances if they so desired.

The subjects stood for all measurements. However, the infinite pipe is installed 2.6 m above the laboratory floor. So, for the resonance-free experiments, the subjects were required to stand on a support of adjustable height, which was in turn placed on a platform 0.75 m above the laboratory floor. The support height was adjusted to bring each subject's mouth to the level of the pipe. No subject reported that this caused any difficulty.

TABLE I. Details of the experimental subjects.

Subjects	Age	Singing experience
S1	29	10 years formal training, 10 years choral experience
S2	22	16 years formal training, formal teaching qualifications
S3	22	4 years choral experience
S4	23	6 years formal training, choral and opera experience
S5	19	6 years formal training, choral experience
S6	23	2 years formal training, 6 years choral experience
S7	25	8 years formal training, 19 years choral singing
S8	20	2 years formal training, 6 years choral experience

D. Vocal glide experiments

The vocal glide experiments involved three different configurations—see Fig. 2. The "unconstrained" configuration shown in Fig. 2(a) allowed the subject to sing normally, including possible variations in their lip aperture geometry. The "constrained" configuration shown in Fig. 2(b) used a mouth ring that limits the ability of a subject to vary lip aperture and jaw height, which are the normal ways of varying f_{R1}. The tongue, which can also vary f_{R1}, was not constrained but subjects were asked to keep the tongue low in the mouth. This configuration is expected to reduce the possibility of resonance tuning, and thus to make f_{R1} largely independent of f_o. The effectiveness of this constraint was tested in a separate experiment. The "resonance-free" configuration shown in Fig. 2(c) replaced the radiation field at the mouth with the purely resistive load of an acoustically infinite pipe; R1 should now be effectively removed. (Supporting calculations are presented below.) The mouth ring and the end of the infinite pipe were thoroughly disinfected with ethanol after each experimental session.

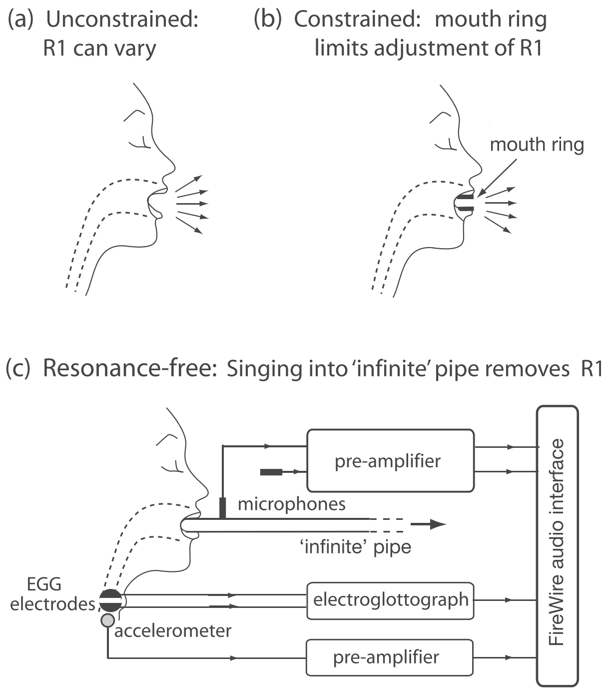

FIG. 2. Schematic diagram (not to scale) showing the three different configurations used for the vocal glide experiments. For brevity, the electronic equipment is only shown in (c), but was also used in configurations (a) and (b).

1. Unconstrained vowel glide at constant pitch

These vowel glides were made at two different nominal pitches (C5 and C6) with the mouth unconstrained; i.e., using the configuration in Fig. 2(a). The vowels chosen were in the carrier words "who'd," "heed," and "hard," which are typically pronounced /u/, /i/, and /ɑ/ in Australia (Delbridge, 1985).

The vowel glides change the resonances of the supraglottal tract, but are not expected to affect the subglottal resonances. The pitches were chosen to be the same as those used by Titze *et al.* (2008). Four subjects were unable to produce a steady vowel glide at C6 as it was at the extreme of their range, so they were asked to sing at B5 or A5 instead. This was taken into account when calculating the expected crossings presented in Table IV.

2. Unconstrained pitch glide with requested constant vowel

The protocol followed Titze *et al.* (2008) and involved singing from high pitch to low, then low to high, with an intermediate period of vocal fry. S4 could not produce vocal fry during any of the glides. Four of the subjects (including S4) found it difficult to start the pitch glide at C6; these subjects started at low pitch and then increased it. If a subject found it difficult to perform an intermediate period of vocal fry, or started the glide at low pitch, their vocal fry was measured separately after each glide. At least two vocal fry phonations were recorded for each subject. This exercise was performed on the same three vowels used for the vowel glides: the vowels in "who'd," "heed," and "hard," which were sustained for the duration of the pitch glide. The two pitch ranges used were C5 to F3 (523 to 175 Hz), and C6 to F4 (1047 to 349 Hz). Each pitch range was usually repeated four times.

3. Constrained pitch glide with mouth ring

Subjects repeated the C5 to F3, and C6 to F4 pitch glides with a short (10 mm) plastic ring of inner diameter 26.2 mm and outer diameter 32 mm held between the lips to keep the mouth geometry constant. The inner and outer diameters of this ring matched those of the infinite pipe. The subjects were instructed to produce a neutral, relaxed (Australian English) "er" vowel as in "heard" with a low and relaxed tongue. They were reminded to be conscious of the position of their tongue at all times.

Adding the mouth ring changes the shape of the supraglottal tract and thus the resonance frequencies. Figure 3(a) shows a vocal tract radius profile taken from an MRI study of a man phonating with a mouth ring of the same dimensions in place (pale) (Hanna *et al.*, 2016a), and with a simulated neutral lip position (dark). Figure 3(b) uses the empirically derived non-rigid lossy parameters of the vocal tract model from Hanna *et al.* (2012, 2016b) to calculate impedance curves downstream from the glottis for those two vocal tract configurations, and additionally, for when the mouth ring is replaced with the 197 m infinite pipe of the same diameter (behaving as a rigid pipe with standard values

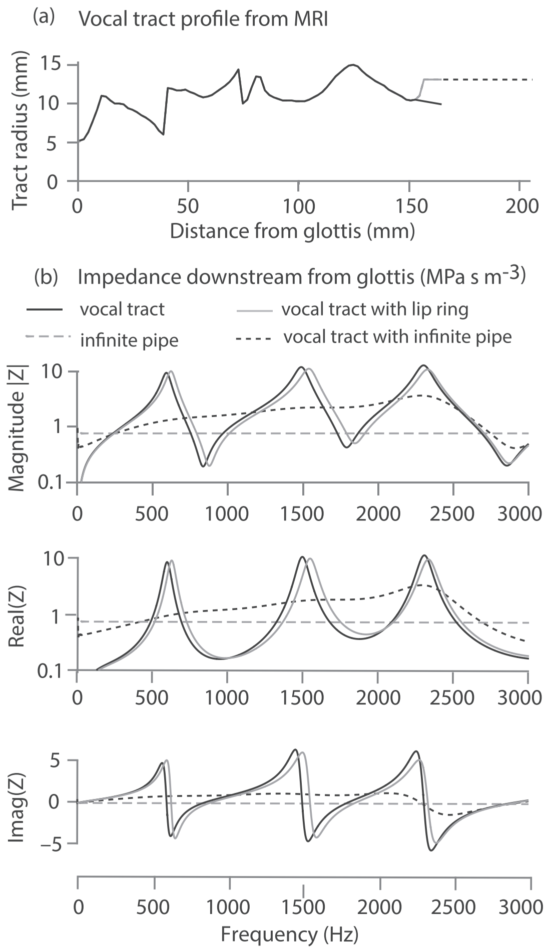

FIG. 3. The calculated effects of a mouth ring and an infinite pipe on the impedance Z downstream from the glottis. The mouth ring or infinite pipe had an internal diameter of 26.2 mm and was inserted 10 mm into the vocal tract. Calculations assumed the vocal tract profile shown in (a) with the magnitude of the (complex) visco-thermal loss factor α increased by a factor of 5 within the tract (Hanna *et al.*, 2016a).

of visco-thermal losses). This shows that the mouth ring increases the frequency of the resonances slightly, whereas the infinite pipe substantially reduces the magnitude of the impedance of the resonances, such that the first resonance is effectively removed.

4. Resonance-free pitch glide with resistive load

For these measurements the mouth ring was replaced by the infinite pipe. Due to the requirement that each subject's mouth be sealed around the end, their lip geometry was then fixed for all measurements. The inner and outer diameter of the infinite pipe also matched those of the mouth ring; subjects should thus have the same lip configuration when singing into the mouth ring or infinite pipe. In both cases, subjects were instructed to keep their tongue low in the mouth and produce a neutral vowel as in the Australian English pronunciation of "heard." The acoustic response of this shape approximates that of a cylindrical pipe, closed at the glottis end and open at the mouth end (Hanna *et al.*, 2016b).

The calculated impedance curves in Fig. 3(b) show that the presence at the lips of the purely resistive load of the infinite pipe can completely remove R1, i.e., there is no peak in absolute magnitude and the imaginary part of the impedance does not change sign, even when there are considerable constrictions in the vocal tract. The vocal tract profiles of the subjects in this study are unknown, so it is possible that considerable constrictions could be present. However, a duct resonance requires multiple reflections with coefficients of reflection very close to one. At the open lips, the coefficient of reflection is very close to one, because the radiation impedance at the lips is very much less than the characteristic impedance of the tract. A constriction in the tract, such as is provided by raising the tongue, provides an increased inertance at that region and while this substantially lowers the frequency of resonances that do not have a flow node in that region, its reflection coefficient is still substantially less than one, so (as shown) it does not produce a resonance.

E. Objectively quantifying instabilities

In order to quantify the stability of phonation, several types of behaviour were noted: (i) rapid changes in f_o, (ii) vibrato, (iii) additional broadband noise in the signal, (iv) period doubling, (v) bifurcations, (vi) abrupt changes in EGG amplitude, and (vii) changes in phase of harmonics in the EGG (i.e., a change in the shape of the EGG time domain waveform).

Since the EGG signal is sensitive to the degree of contact with the skin and the vertical position of the larynx, changes in the EGG signal do not necessarily imply changes in vocal fold motion. Similarly, the sound signal is filtered by the vocal tract and radiation impedance, and therefore is not simply related to the vocal fold oscillation. For these reasons, both EGG and sound signals were used for different purposes.

First, f_o was estimated from the sound signal using the standard PRAAT (version 6.0.17) autocorrelation algorithm with a time window $\Delta t = 0.05$ s and $100 \leq f_o \leq 1200$ Hz. The f_o contours were then extracted and processed with custom MATLAB software to calculate the rate of change $df_o/dt = 12\log_2(f_o(n)/f_o(n-1))/\Delta t$ between successive samples in semitones per second.

A 0.5 s moving average was then subtracted from df_o/dt to give a signal centred around 0 for both ascending and descending pitch glides. Finally, the absolute magnitude of the signal was displayed as a function of f_o on an axis from 0 to 48 semitones/s. (To give an indication of this scale: a vibrato rate of ~7 Hz and amplitude of 0.5 semitone gives a maximum df_o/dt of ~20 semitones/s.) On their own initiative, subjects often added vibrato at the beginning and end of their pitch glides, so increases in variation are often visible at the end of the range. Rates higher than this were only observed with period doubling, where f_o drops an octave (12 semitones) abruptly. When period doubling events were detected a "D" is displayed instead of df_o/dt on the figures.

Next, the EGG signal was used to compute the "sample entropy," a measurement of the complexity of a time-varying signal, using MATLAB code provided by Selamtzis (Selamtzis and Ternström, 2014). For this study, the amplitude and phase of the f_o and $2f_o$ components of the EGG signal were calculated over a window of ten periods of oscillation and compared with the following window overlapping by nine cycles. The matching tolerances (see Selamtzis and Ternström, 2014 for details) were set to 0.4 for amplitude and 1 for phase (to determine rapid changes in amplitude), and 1 for amplitude and 0.4 for phase (to determine rapid changes in phase). The resulting "sample entropy" is close to 0 for near stationary sequences and close to 1 for abrupt changes in the EGG signal. The largest values, those with normalised amplitude larger than 0.9, corresponding to the most abrupt changes, were used to indicate a discontinuity in amplitude (filled circles on the figures), or a discontinuity in phase (open circles). If discontinuities in both amplitude and phase were present at the same frequency, only a filled circle was shown.

Previously, Selamtzis and Ternström used "sample entropy" to identify mechanism changes in male singers. However, this approach has yet to be validated for detecting mechanism changes in women singers and detection thresholds are expected to be different; so putative changes in mechanism are not proposed here. Instead, the algorithm was used to calculate the amplitude changes and the phase changes separately.

Figure 4 illustrates the analysis of the f_o contour, df_o/dt, the microphone and EGG spectrograms, EGG amplitude and "sample entropy" (amplitude in dark, phase in pale).

III. RESULTS AND DISCUSSION

A. Subglottal resonances

The measured subglottal resonances frequencies were similar for all subjects; f_{Sg1} was 630 Hz and f_{Sg2} was 1635 Hz, with ranges of ± 50 and ± 100 Hz, respectively. These values are comparable with other measured values (e.g., Chi and Sonderegger, 2007; Lulich, 2010; Lulich *et al.*, 2012). As expected, they did not vary significantly with f_o or the vowel. The values of f_{Sg1} measured on individual singers are indicated by grey dots in Figs. 7–9, 11, and 12. For the experiments described here, it is sufficient to know that f_{Sg2} was well above the highest glide frequency of 1050 Hz, and the measurements found no interaction between f_{Sg2} and $2f_o$ (see Fig. 7).

B. Supraglottal resonances

Table II shows the values of $F1$ measured using vocal fry and compares them with other (previous) measurements made on Australian women using vocal fry, or values of f_{R1} measured using external broadband during speech. Table III compares the values of f_{R1} for two vowels measured during singing in previous experiments with those measured using vocal fry in this study. The values in Table III are consistent with previous observations that resonance tuning commences once f_o exceeds the value of f_{R1} in speech. A conclusion is that the values of $F1$ measured using a period of vocal fry do not reliably indicate the f_{R1} for singing, particularly at high f_o, unless the singers were reliably capable of maintaining the same vocal tract configuration for both.

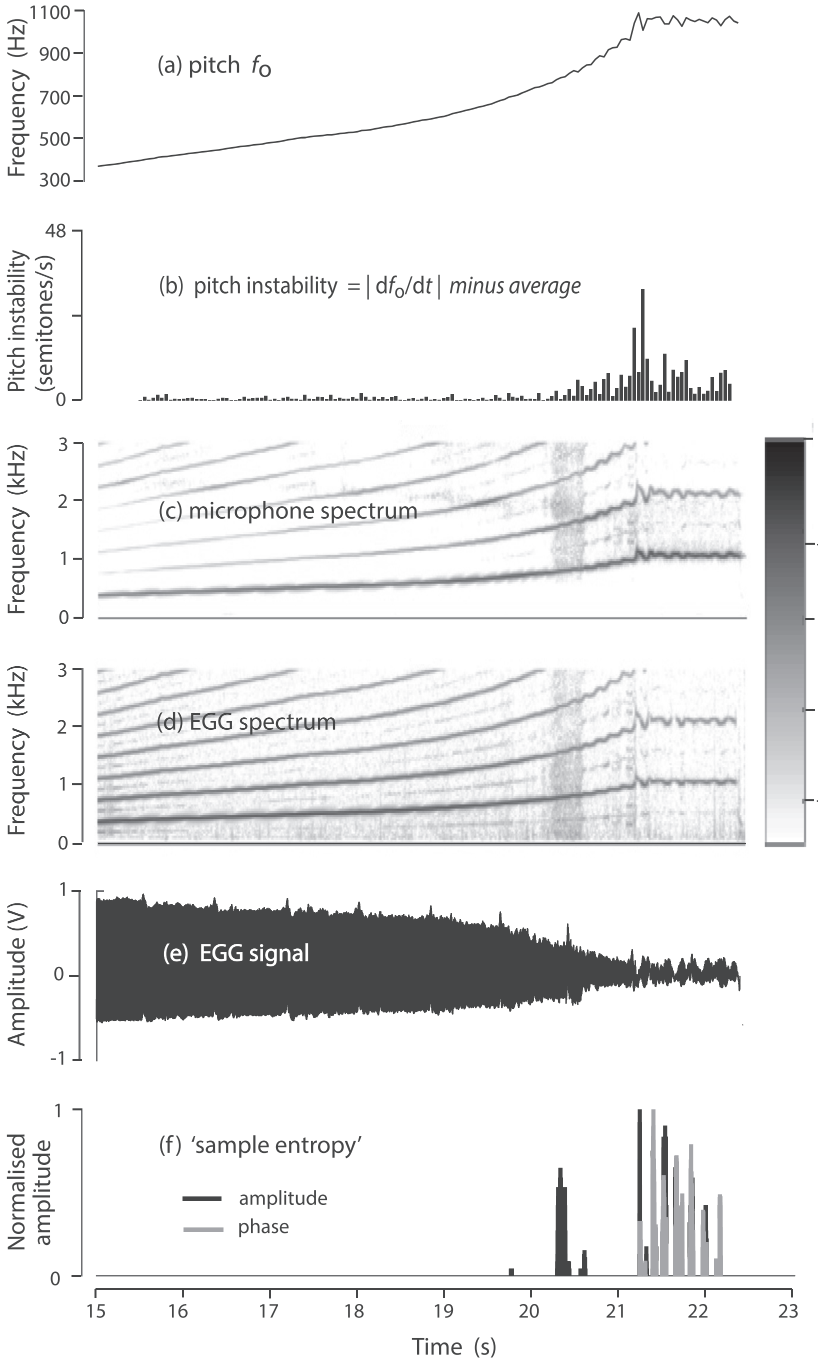

FIG. 4. Example showing the calculations of pitch instability during an unconstrained high pitch glide from F4 to C6 by singer S3 on the vowel in "heed." (a) and (b) show the variations in pitch f_o and pitch instabilities ($|df_o/dt|$ minus the 0.5 s average) calculated using PRAAT from the measured EEG signal shown in (e). (d) shows the EGG spectrum and (c) shows the microphone spectrum measured simultaneously. (f) displays the calculated "sample entropy" showing discontinuities in amplitude and phase.

Performing vocal fry into the infinite pipe showed that there was no first formant corresponding to $F1$ and the frequency and bandwidth of the second formant were increased when compared to the same measurements with the mouth ring in place, as shown in the sound signal in Fig. 5.

C. Unconstrained vowel glides at constant pitch

1. Vowel glide: Who'd to hard to who'd

First, the possible consequences of resonance tuning during the vowel glides needs to be considered. The resonances measured using vocal fry presented in Table II suggest that $F1$ would vary from about 405 to 710 Hz during the vowel glide in the absence of resonance tuning, i.e., as expected for the unconstrained experimental configuration [Fig. 2(a)] if the singers are able to hold $F1$ constant. In this situation, $F1$ would usually cross f_o for C5 (523 Hz) and rarely cross for C6 (1047 Hz) as illustrated in Fig. 6(a). In contrast, Fig. 6(b) shows other measurements made during normal singing, where the singers were asked to sing a given vowel, but not asked to keep the tract in the same configuration. Here, data for the note C5 (see Table III) indicate that f_{R1} would vary instead from 535 to 655 Hz as a consequence of resonance tuning and consequently there would be no

TABLE II. Values of $F1$ and $F2$ in Hz measured using vocal fry for Australian women compared with values of f_{R1} and f_{R2} measured during speech in earlier studies. Values expressed as mean ± standard deviation.

	Vowel					
	heard	who'd	heed	hard	Technique	Reference
f_{R1}	600 ± 40	390 ± 80	350 ± 60	740 ± 140	ext broadband	Donaldson *et al.*, 2003
f_{R1}	625 ± 60	435 ± 70	—	780 ± 95	ext broadband	Swerdlin *et al.*, 2010
f_{R1}	580 ± 50	410 ± 80	—	825 ± 110	ext broadband	Henrich *et al.*, 2011
F1	665 ± 75	480 ± 55	—	800 ± 60	vocal fry	Swerdlin *et al.*, 2010
F1	620 ± 25	405 ± 45	395 ± 50	710 ± 85	vocal fry	This study
f_{R2}	1620 ± 160	1670 ± 250	2490 ± 390	1330 ± 70	ext broadband	Donaldson *et al.*, 2003
f_{R2}	1555 ± 100	1480 ± 355	—	1370 ± 45	ext broadband	Swerdlin *et al.*, 2010
f_{R2}	1545 ± 270	1410 ± 400	—	1265 ± 130	ext broadband	Henrich *et al.*, 2011
F2	1550 ± 120	1440 ± 325	—	1365 ± 70	vocal fry	Swerdlin *et al.*, 2010
F2	1125 ± 120	910 ± 95	2145 ± 180	1210 ± 110	vocal fry	This study

crossing as shown in Fig. 6(b). Similarly for the note C6, resonance tuning means that f_{R1} could actually vary from 890 to 1030 Hz and again there would be no crossing.

For each of the measured vowel glides, the spectrograms of the EGG signal and the sound were examined for instabilities in pitch. The results presented in Table IV show no agreement between the number of observed pitch instabilities and the number expected if they arose from f_{R1} crossing f_o in the case where f_{R1} is estimated using $F1$ for "who'd." It also shows no agreement with the expected absence of instabilities if resonance tuning occurred.

2. Vowel glide: Who'd to heed to who'd

The resonances measured using vocal fry presented in Table I suggest that $F1$ would vary from about 405 to 395 Hz in the absence of resonance tuning. There would then be no crossover for either C5 or C6. No measured values of f_{R1} are available for singing "heed" at C5 and C6. However, resonance tuning has been measured for another close vowel ("who'd") so it is likely that it also occurs for "heed." In this case $F1$ would not cross f_o. The results shown in Table IV are again inconsistent with the instabilities predicted in the absence or presence of resonance tuning.

3. Vowel glide: Heed to hard to heed

Table II shows that f_{R1} would vary from about 395 to 710 Hz in the absence of resonance tuning and so f_{R1} would cross f_o for C5; however, if resonance tuning occurred there would be no crossover. For C6 there would be no crossover in either case. (The exception was S8, who could only reliably produce this vowel glide at 700 Hz, and then crossings would be expected.)

TABLE III. The formant frequencies in Hz for two of the vowels measured using vocal fry in this study (eight subjects), compared with values of f_{R1} measured in previous studies during singing at C5 and C6 [14 subjects (who'd) and 26 subjects (hard)—data combined from Joliveau *et al.* (2004b) and Henrich *et al.* (2011)]. Values expressed as mean ± standard deviation (number of samples).

	Vocal fry	C5 ($f_o = 523$ Hz)	C6 ($f_o = 1047$ Hz)
who'd	405 ± 45	535 ± 40 (17)	890 ± 145 (7)
hard	710 ± 85	655 ± 100 (52)	1030 ± 55 (46)

4. Vowel glide: Instabilities and resonance frequencies

Although the expected number of vowel glides where $F1$ determined from vocal fry would cross f_o varied from 0 to 28, Table IV shows that the recorded number of pitch instabilities was similar for the six different vowel glides. Indeed, instabilities were only recorded in 47% of the 61 glides where a crossing was predicted using $F1$ measured using vocal fry. This result is comparable with that of Titze and colleagues (2008) for female singers where it was found that instabilities only occurred in 34% of the vowel glides where vocal fry measurements suggested a crossover might occur.

The data for vowel glides are thus inconclusive as the glides for different vowels and pitches all show a similar rate of instabilities. If f_{R1} was similar to the value of $F1$ measured using vocal fry, the number of expected resonance crossings would be between 0 and 100%. If resonance tuning occurred and the singers kept $f_{R1} > f_o$, the number of expected crossing would be very small and consequently the number of expected pitch instabilities would then also be correspondingly small. However, the results show a similar rate of pitch instabilities for each of the vowel glides, even when crossings are not expected; this suggests that in this case, similar to previous reported results (Titze *et al.*, 2008; Zañartu *et al.*, 2011), the instabilities were not necessarily associated with changes in the acoustic load. One difficulty is that only the end points of vowel glides at high pitch can be determined, the actual trajectory on the (f_{R2}, f_{R1}) plane and how it varies with f_o is not known.

It is also possible that instabilities might be caused by f_{R2} crossing f_o or $2f_o$, however Table IV again shows no strong agreement between the observed number of pitch instabilities and the predicted number of crossings.

D. Unconstrained pitch glides with normal singing

To assess the stability of the pitch glides, the rate of change of f_o, and discontinuities in the amplitude and phase of f_o and $2f_o$ were calculated as described in Sec. II E. The upper parts of Figs. 7 and 8, 9, 11, and 12 show the pitch

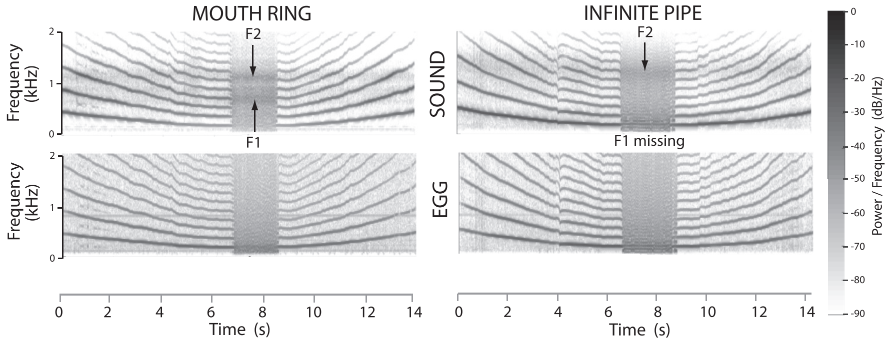

FIG. 5. Spectrograms of the microphone sound and EGG signals measured simultaneously when singer S4 performed pitch glides with a mouth ring (left) and the infinite pipe (right). The intermediate period of vocal fry indicates that R1 is absent when phonating into the infinite pipe.

instability as a function of f_o when each singer sang the pitch glide using the nominated vowel. The data from the C5 to F3 (nominal 523 to 175 Hz), and C6 to F4 (nominal 1047 to 349 Hz) glides are shown separately; data from the

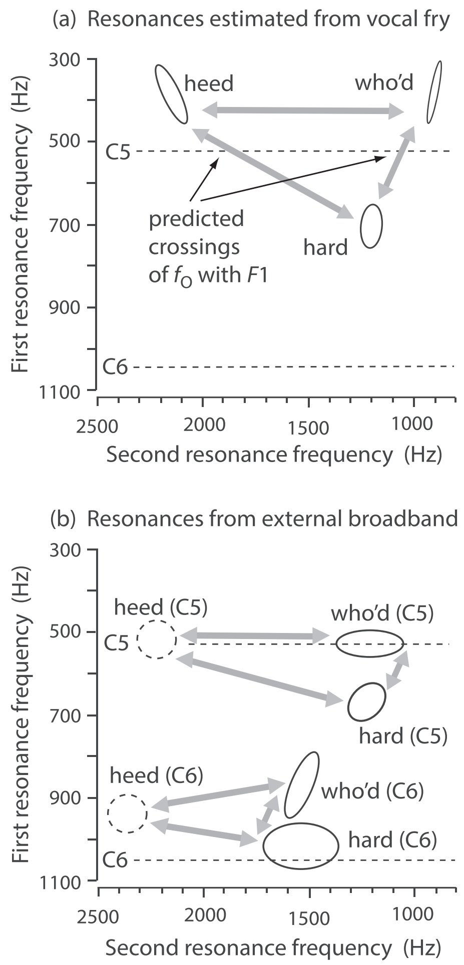

FIG. 6. Possible trajectories on the (R2, R1) plane for vowel glides at two constant pitches, C5 and C6. (a) shows $F1$ measured using vocal fry for the eight singers in this experiment. Each ellipse is centred on the mean value of ($F2$, $F1$), the slope of the major axis indicates the regression of $F1$ on $F2$, and the semi-axes indicate the standard deviations in those directions. (b) shows the values of f_{R1} measured during singing using external broadband in previous experiments in this laboratory (data combined from Joliveau *et al.*, 2004b, Garnier *et al.*, 2010, and Henrich *et al.*, 2011). The ellipses display the data for f_{R1} and f_{R2}. Values of (f_{R2}, f_{R1}) are not available for singing "heed" at high pitch; an estimate of those ellipses is indicated by dashed lines.

repetitions (at least two increasing and two decreasing pitch glides) within each range have been overlaid. These data are not averaged, so that instabilities occurring even once will remain visible.

Across all the pitch glides, large instabilities in pitch and discontinuities in harmonic phase and amplitude (evident as peaks in the "sample entropy") occur at both the low and high frequency ends. At the low end of the glides, the subjects were instructed to transition into vocal fry phonation. At the high end of the glides, the subjects tended (without instruction) to introduce vibrato around the target note. Both of these behaviours increase the pitch instability and "sample entropy."

Soprano parts are rarely written below C4, and most sopranos cover the range of C4 and above using laryngeal mechanism M2 (the "head voice"). So, notes below C4 (262 Hz) are unfamiliar territory, and it is therefore expected that the singers would experience instability in this region. Furthermore, singers would be expected to cross their "break," i.e., change between vocal fold mechanism M1 and M2 (the *primo passagio* between the "chest voice" and "head voice"). These breaks have not been distinguished from the other instabilities, since it was not possible to identify them unambiguously.

Due to poor contact with the EGG electrodes, the EGG signal for S5 was recorded with a low signal to noise ratio (SNR), so the phase of the harmonics of the voice were frequently uncorrelated. For this reason, the phase data from the "sample entropy" calculations are often not shown for S5. Occasional low SNR is responsible for several of the phase discontinuities displayed in the data from other subjects, however since in some cases this may be due to laryngeal movement rather than loss of contact, they have not been removed.

The values of f_{Sg1} and $F1$ reported in Secs. III A and III B are indicated for each singer, respectively, by gray circles and downward pointing triangles. The lower parts of Figs. 7, 8, and 10 display, using the same frequency axis, the values of f_{R1} measured during singing on similar subjects in previous experiments in this laboratory. For comparison, the values of $F1$ or f_{R1} measured for speech in several studies are indicated by the horizontal dashed lines.

1. Who'd

The upper part of Fig. 7 shows the pitch instability as a function of f_o when each singer sang the pitch glide using the

TABLE IV. Comparison between the number of pitch instabilities occurring during vowel glides and the number of times a resonance crosses f_o, assuming that f_{R1} and f_{R2} are equal to the values of $F1$ and $F2$, respectively, as measured using vocal fry.

				No. of expected crossings using vocal fry values		
Pitch	Vowels	No. of glides	No. of measured pitch instabilities	f_{R1} crosses f_o	f_{R2} crosses f_o	f_{R2} crosses $2f_o$
C5	who'd – hard	28	12	28	0	14
C6	who'd – hard	23	16	3	18	0
C5	who'd – heed	28	12	0	0	20
C6	who'd – heed	25	15	0	14	19
C5	heed – hard	27	15	27	0	0
C6	heed – hard	26	15	3	0	20

vowel in "who'd." Many instabilities are seen to occur towards or beyond the normal limits of the soprano vocal range; typically below about C4 (262 Hz) to around A5 to B5 (880–988 Hz). Instabilities do not appear to be caused by the subglottal impedance passing through a maximum and changing between inertive and compliant, as the instabilities are not clustered around the subglottal resonance frequencies. Similarly the absence of instabilities clustering around the values of $F1$ indicates either that changes in the supraglottal load are not important, or that the values of $F1$ measured using vocal fry are not appropriate when singing.

To examine this further, the lower part of Fig. 7 displays, using the same frequency axis, the values of f_{R1} (the first tract resonance) measured during singing in previous experiments in this laboratory. No clustering of instabilities is apparent in the region where f_o crosses the resonance values measured for speech. Instead, this range (around 400–450 Hz), which lies comfortably away from either extreme of the soprano range, showed stable phonation for all subjects except S5. Phase discontinuities were also observed in S1 and S7. In the lower part of Fig. 7, the range over which $f_{R1} \approx f_o$ indicates resonance tuning. Below this range $f_{R1} > f_o$. The error bars show that f_o will probably cross f_{R1} only at frequencies around or above 450 Hz. The occurrence of instabilities in this region for most singers in the low pitch glide is likely due to vibrato around the final target note, since the same region (450–600 Hz) is stable in the high pitch glide. This is consistent with the predictions assuming the presence of resonance tuning. Some pitch instabilities occur above 600 Hz in S2, S5, and S6, and amplitude and phase discontinuities are observed in S1, 2, 3, 5, and 7. All singers were able to sing beyond the frequency of instability, showing that the instability was not simply a consequence of approaching the limits of their range. For many sopranos, notes of A5 (880 Hz) and above approach the upper limit of the range over which they practice resonance tuning of f_{R1} to f_o (see Joliveau *et al.*, 2004a).

2. Hard

The vowel in "hard" is an open vowel, with a high value of f_{R1} in speech—see Table II. Pitch glides for this vowel are shown in Fig. 8. The absence of significant pitch instabilities during the C6 to F4 pitch glides in all except S5 is consistent with the suggestion that resonance tuning is maintaining f_{R1} above f_o and consequently the supraglottal load will be inertive. Some phase and amplitude discontinuities are observed in these high pitch glides but generally do not occur close to f_{Sg1} and the $F1$ from vocal fry. Again, pitch instabilities and discontinuities are observed at the top of the pitch glides, consistent with vibrato. In the low glides, pitch instabilities, and period doubling in S3, S4, and S7, occur near or below the lower limit of the M2 ("head voice") range.

3. Heed

For pitch glides on the vowel in "heed," Fig. 9 shows that, for the low pitch glides, there are no pitch instabilities in the 300–500 Hz range, which includes $F1$. In the high pitch glides, S5 and S6 show pitch instabilities close to $F1$. The region above $F1$ is stable in pitch for all except S5 until around 900 Hz. Although values of f_{R1} during singing are not available for "heed," it is expected that resonance tuning will occur in a similar fashion to the other close vowel studied ("who'd"), and that there will be an upper limit to this resonance tuning when it becomes uncomfortable or anatomically impossible to continue raising f_{R1} (perhaps around 900 Hz). In that situation, the supraglottal load may become compliant. Again, there is no sign of the clustering of instabilities around f_{Sg1}, and the majority of pitch instabilities occur below the normal range of M2 ("head voice").

E. Constrained pitch glides with mouth ring

Inspection of the pitch glides performed by the panel of subjects on the vowels in "who'd," "hard," and "heed" (Figs. 7–9) showed no significant correlation between the frequencies where pitch instabilities occurred and the measured values of f_{Sg1} or $F1$. Even if the measured bandwidths of R1 or Sg1 (both around 75 Hz in women—see Hanna, 2014; Hanna *et al.*, 2016b) are considered, no significant correlation is apparent. One explanation for this was the likely use of at least some degree of resonance tuning, i.e., subjects can smoothly increase f_{R1} to keep f_{R1} above or around f_o. Consequently, an additional experiment was included to limit resonance tuning by constraining the lip aperture using a small mouth ring—see Fig. 2(b). In this condition, the subject can still change f_{R1} by changing the shape of the tongue and other internal articulators. However, since f_{R1} is strongly dependent on lip aperture, the range of f_{R1} should be limited sufficiently to reduce the possible range of resonance tuning. (Further, because singers are unfamiliar with singing with a

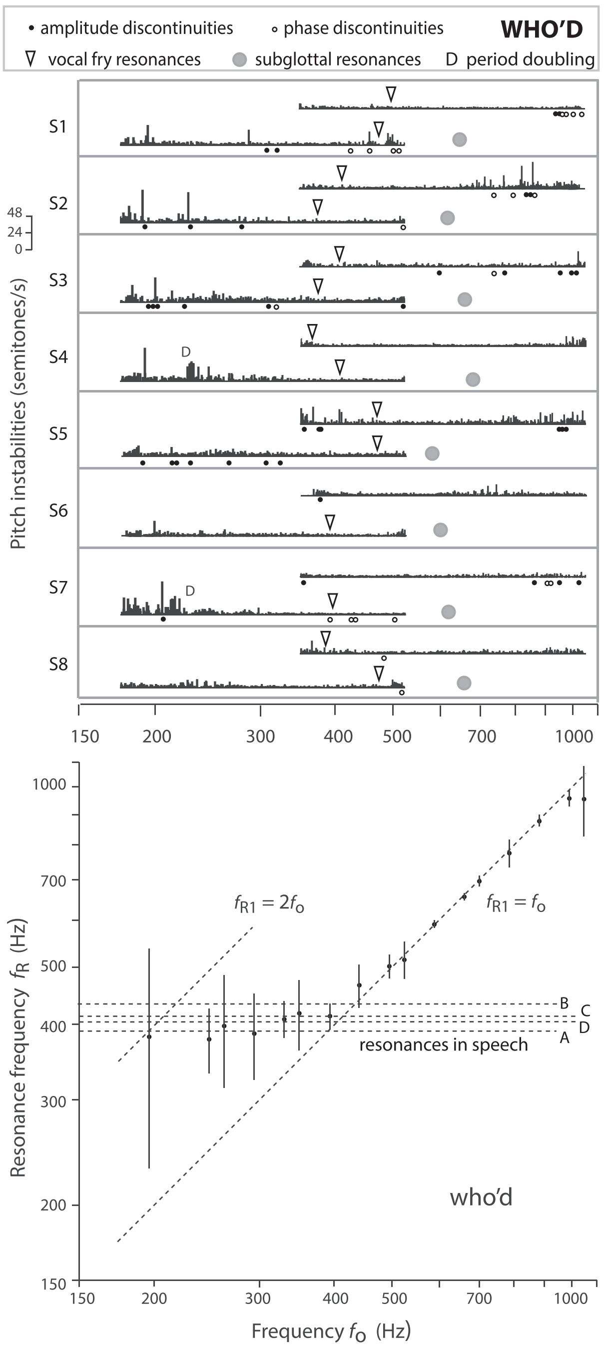

FIG. 7. The relationships between pitch instabilities, pitch frequency and resonance frequencies for the vowel in "who'd" in normal singing. The range of the glides between C5 and F3, and between C6 and F4 are shown separately. The upper part of the figure displays the pitch instability (absolute value of df_o/dt in semitones/s, minus the steady average) as a function of f_o for the two glide ranges for each singer. The triangles and grey dots indicate the measured average values of $F1$ and f_{Sg1}, respectively, for each singer. The lower part shows previously measured values and error bars for f_{R1} from 248 measurements on 19 singers (Joliveau *et al.*, 2004b; Henrich *et al.*, 2011). The horizontal dashed lines indicate the average values for speech measured by an external broadband signal at the lips in previous studies; A (Donaldson *et al.*, 2003), B (Swerdlin *et al.*, 2010), C (Henrich *et al.*, 2011), and as estimated by vocal fry in this study (D). The long diagonal dashed line indicates the relationship $f_o = f_{R1}$ and also indicates the overall frequency range covered by the glides (F3 to C6). The shorter diagonal dashed line indicates the relationship $f_o = 2f_{R1}$. Error bars indicate the standard deviation.

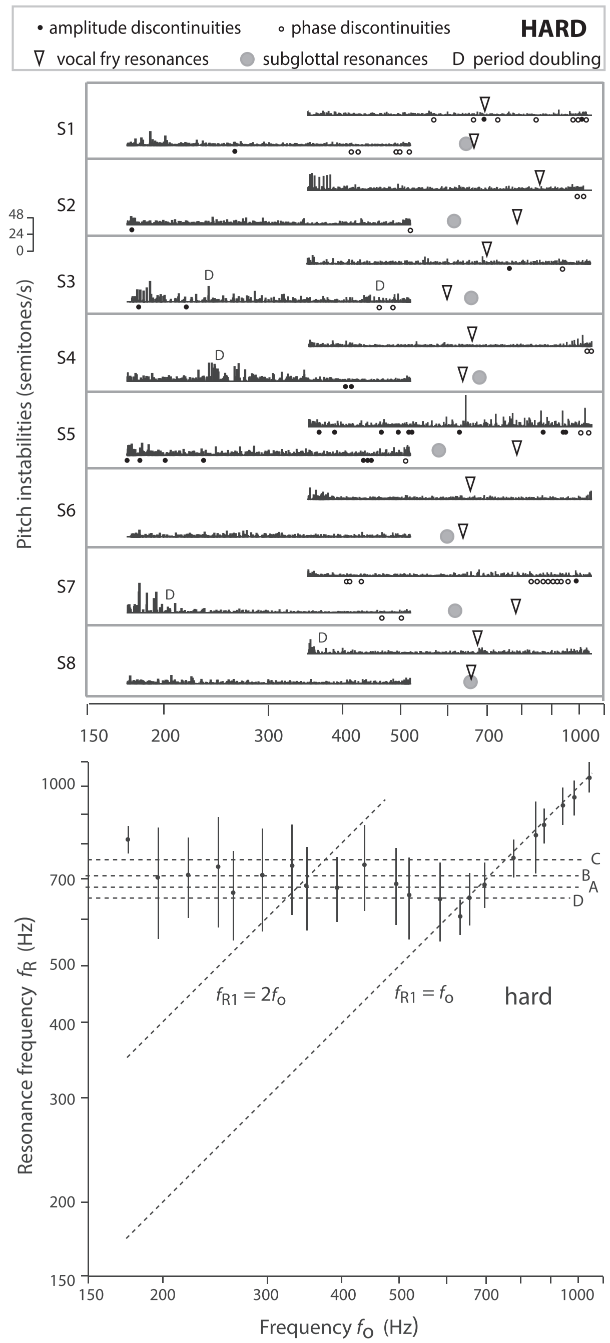

FIG. 8. The relationships between pitch instabilities, pitch frequency and resonance frequencies for the vowel in "hard" in normal singing. The data points and error bars for f_{R1} are calculated from 765 previous measurements on 31 singers (Joliveau *et al.*, 2004b; Garnier *et al.*, 2010; Henrich *et al.*, 2011). Otherwise the legend for Fig. 7 applies.

mouth ring, they will not have learned how to tune resonances while using one.)

To confirm that the constrained lip aperture produced by the mouth ring did indeed limit the extent of adjustment of R1 and accuracy of any resonance tuning, a detailed study of f_{R1} as a function of f_o was made using external broadband at the lips for one singer using a mouth ring—results are shown in the lower part of Fig. 10.

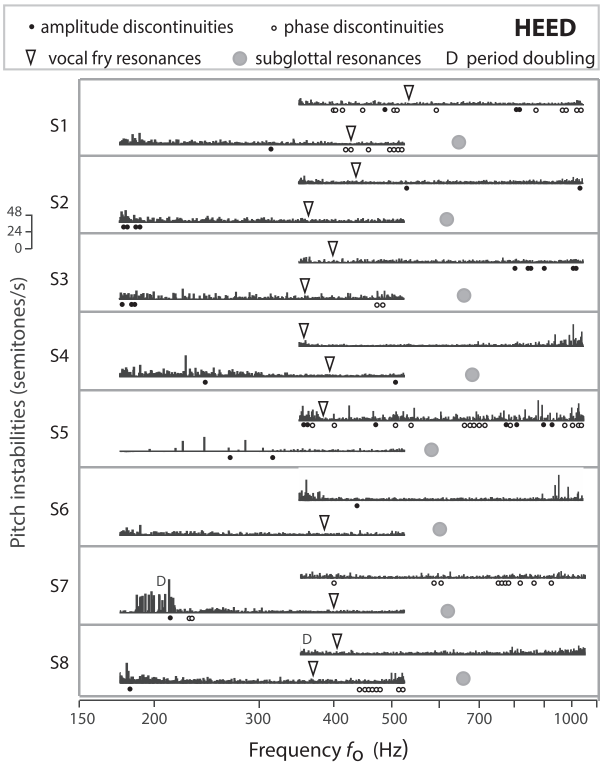

FIG. 9. The relationships between pitch instabilities, pitch frequency and resonance frequencies for the vowel in "heed" in normal singing. Otherwise the legend for Fig. 7 applies.

The lower part of Fig. 10 shows the expected reduction of resonance tuning above 600 Hz; consequently, f_o passes through the resonance f_{R1} at a lower frequency. It is also apparent that f_{R1} is greater than the typical value for the vowel in "heard," presumably because the mouth ring enforces an effective lip aperture that is larger than normal for this vowel.

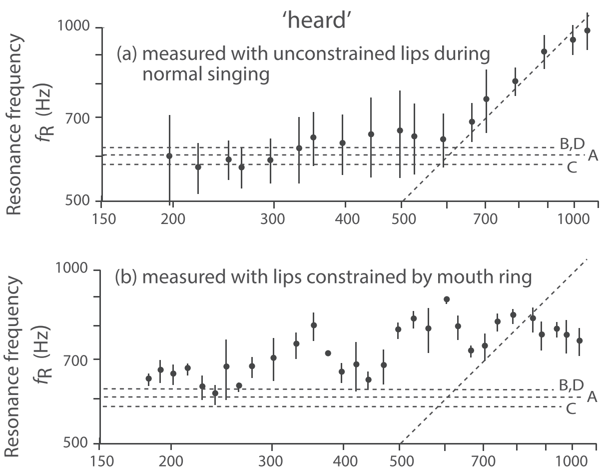

FIG. 10. The effect of a fixed mouth aperture on the measured resonance frequency f_{R1} when subjects sing the vowel in "heard." (a) shows the results of previous measurements at the lips [see Fig. 1(c)] during normal singing (270 measurements on 19 singers from Joliveau *et al.*, 2004b; Henrich *et al.*, 2011). (b) shows the results of 167 measurements on one singer (S6) with a mouth ring to constrain lip aperture. Measurements in both (a) and (b) were made using external broadband at the lips—see Fig. 2(c). Horizontal dashed lines show previous results for unconstrained speech—see Fig. 7 for details.

Figure 11 shows the results of pitch glides with the mouth ring for the panel of subjects. In this configuration, and in the absence of resonance tuning, one would then

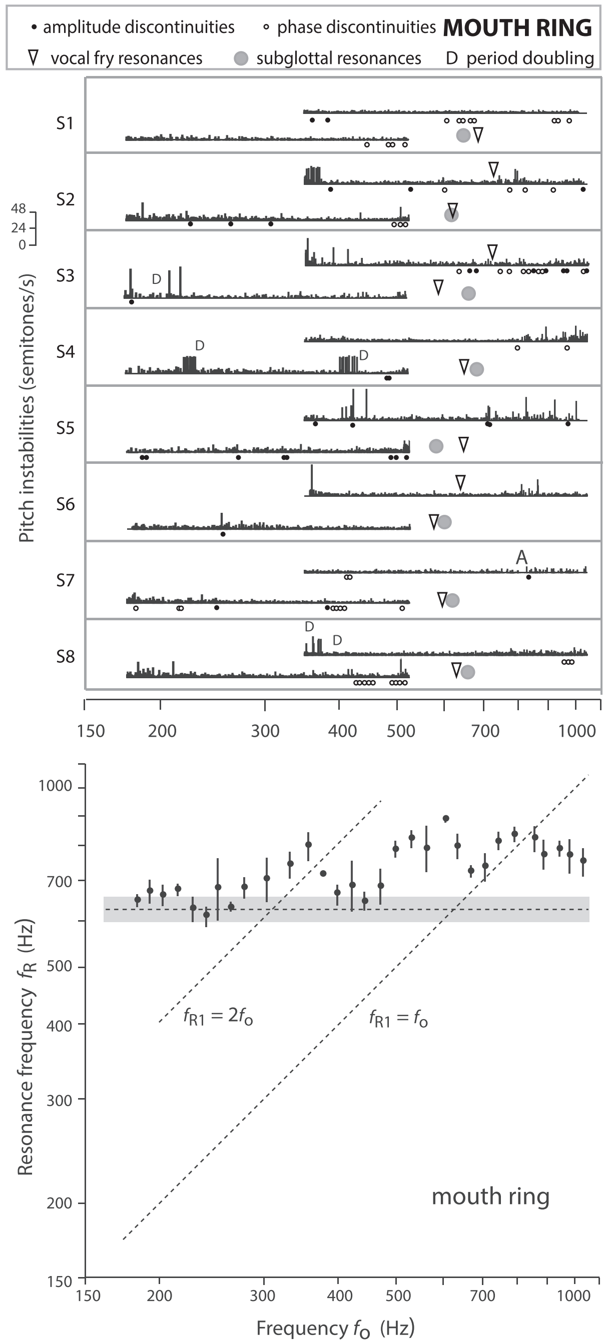

FIG. 11. The relationships between pitch instabilities, pitch frequency and resonance frequencies with the subject instructed to sing the vowel in "heard" (Australian English), but with the lip aperture constrained by a ring. A brief period of aphonia for S7 is indicated by A on the upper figure. The data points and error bars for f_{R1} in the lower figure are calculated from 167 measurements on singer S6. The horizontal dashed line shows the average value of $F1$ measured with the mouth ring using vocal fry; the grey area shows mean ± SD. Otherwise the legend for Fig. 7 applies.

expect f_o to cross f_{R1} at about 800–900 Hz. Indeed, four of the eight singers S2, 4, 5, and 6 exhibited a pitch instability, and S7 exhibited aphonia (just above 800 Hz) in this frequency range. Subjects 3, 4, and 5 also showed rapid pitch changes around 400 Hz, perhaps where $2f_o \sim f_{R1}$.

F. Resonance-free pitch glides singing into the infinite pipe

The experiments with the mouth ring are consistent with vocal tract resonances causing an instability. A more direct method to investigate whether instabilities are associated with vocal tract resonances is to remove the resonances. The mouth ring was therefore replaced with the infinite pipe of the same external and internal diameters. This should present an almost purely resistive load at the lips, and thus an impedance loading the glottis with only relatively small values of inertance or compliance. This should effectively remove R1—see the result of calculations in Fig. 3(b) and the results for intermediate vocal fry shown in Fig. 5. Figure 12 shows the results of pitch glides when the subjects sang into the infinite pipe. A notable feature is that the magnitude of pitch instabilities is reduced compared with the mouth ring for S2, 3, 4, 6, and S7 did not experience aphonia. The instabilities close to $2f_o \sim f_{R1}$ when measured with the mouth ring are also reduced. The removal of R1 now leaves Sg1 as the only resonance covered by the pitch glide, but even then there is no evidence of instabilities around f_{Sg1} as has been reported for *in vitro* laryngeal models (Zhang *et al.*, 2006).

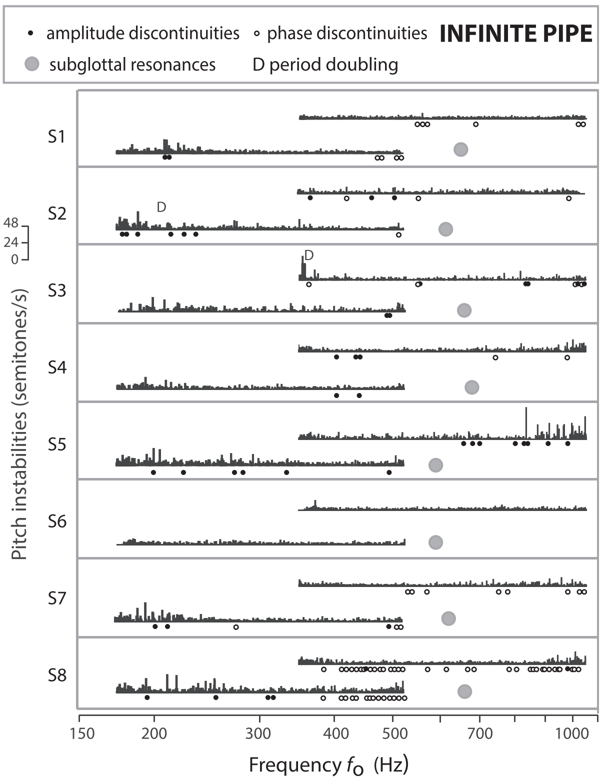

FIG. 12. The relationships between pitch instabilities, pitch frequency and resonance frequencies when singers sang into an infinite pipe. Otherwise the legend for Fig. 7 applies.

At this stage it is useful to consider the instabilities of the individual singers at high pitches above the subglottal resonances (in the approximate region 700 to 1000 Hz). Singer S5 shows several instabilities in most pitch glides; consequently, we will now only consider the other seven singers. For the resonance-free pitch glides using the infinite pipe there were no significant instabilities (Fig. 12); this is consistent with f_o never crossing f_{R1}. The introduction of the mouth ring makes it difficult to alter f_{R1} significantly and consequently f_o might be expected to cross f_{R1} in this frequency range; indeed S2, S4, S6 showed instabilities and S7 had a brief period of aphonia. (see Fig. 11). The vowels in "who'd" and "heed" have low values of f_{R1} in speech (see Table II) and subjects would be expected to reach an upper tuning limit for f_{R1} (Joliveau *et al.*, 2004a) because it is anatomically difficult for subjects to keep their lips in the "who'd" or "heed" position and simultaneously shift the resonance to a very high value. For "who'd," S2 shows significant pitch instability around 800 Hz with S3 and S4 showing smaller instabilities at higher frequencies—see Fig. 7. For "heed," S4 and S6 display instabilities towards the upper limit of their range—see Fig. 9. The vowel in "hard" already has a high value of f_{R1} in speech (see Table II) and subjects should find it relatively easy to adjust to avoid a crossing; indeed Fig. 8 displays no instabilities (except the atypical S5).

The results given in Table V are thus consistent with the generalisation that a pitch instability may result when f_o crosses the resonance f_{R1}, but that the sopranos were usually able to adjust f_{R1} to avoid this.

The "sample entropy" data are more difficult to interpret. As previously mentioned, some of the phase discontinuities result from low SNR in the EGG signal, e.g., S8 singing into the infinite pipe. Similarly, the EGG amplitude may present discontinuities but whether these represent changes in vocal fold motion, laryngeal height, or skin contact is unknown. Regardless, there is no pattern of clustering of these events with the estimates of f_{R1} or measurements of $F1$ and f_{Sg1}. An amplitude discontinuity accompanied by a pitch instability is likely to indicate a mechanism change, particularly below 300 Hz, but this condition is not sufficient to identify these changes unambiguously.

IV. CONCLUSIONS

Pitch instabilities were rarely found at frequencies near those of the measured frequencies of the lowest subglottal

TABLE V. The occurrence of pitch instabilities in the frequency range above 700 Hz during the five different pitch glides. P indicates a pitch instability, and A indicates brief aphonia.

		Singer							
Pitch glide	R1 tuning	S1	S2	S3	S4	S5	S6	S7	S8
who'd	f_{R1} tuning has upper limit	—	P	P	P	P	—	—	—
hard	f_{R1} tuning possible over range	—	—	—	—	P	—	—	—
heed	f_{R1} tuning has upper limit	—	—	—	P	P	P	—	—
mouth ring	limited f_{R1} tuning	—	P	—	P	P	P	A	—
infinite pipe	f_{R1} removed	—	—	—	—	P	—	—	—

resonance. One possibility is that the subglottal resonances have little effect upon phonation. Another is that, as the transition from an inertive to a compliant subglottal load will always have been present at a similar frequency throughout the adult life of each individual soprano, they have learnt a technique to cope satisfactorily with the change in the phase and magnitude of this acoustic load.

The results from the unconstrained pitch and vowel glides where singers were free to adjust their supraglottal tract geometry showed no correlation between pitch instabilities and the resonant frequencies determined using vocal fry. This is consistent with the expectation of resonance tuning, whereby soprano singers increase f_{R1} from its value in speech, the increase being larger for high f_o; consequently the possibility of crossovers is reduced.

Resonance-free experiments, where the lowest acoustic tract resonance was essentially eliminated when singers sang into a purely resistive load, showed a reduced number of instabilities, except at frequencies near or below the lower limit of laryngeal mechanism M2 ("head voice") for sopranos. When a mouth ring was used to constrain their ability to perform resonance tuning, they experienced more instabilities, possibly due to the strong variations in the acoustic impedance with frequency that are associated with the presence of resonances.

ACKNOWLEDGMENTS

We thank our volunteer subjects, Andreas Selamtzis for providing Matlab code and advice for calculating "sample entropy," the Australian Research Council for its support, and thank two anonymous reviewers and the Associate Editor for very helpful comments.

Arroabarren, I., and Carlosena, A. (**2006**). "Inverse filtering in singing voice: A critical analysis," IEEE Trans. Audio Speech Lang. Process. **14**, 1422–1431.

Boersma, P., and Weenink, D. (**2016**). "Praat: Doing phonetics by computer" [Computer program], version 6.0.17, http://www.praat.org/ (Last viewed November 3, 2016).

Chi, X., and Sonderegger, M. (**2007**). "Subglottal coupling and its influence on vowel formants," J. Acoust. Soc. Am. **122**, 1735–1745.

Delbridge, A. (**1985**). *The Macquarie Dictionary* (Macquarie Library, Dee Why, Australia), pp. 810, 820, 825.

Donaldson, T., Wang, D., Smith, J., and Wolfe, J. (**2003**). "Vocal tract resonances: A preliminary study of sex differences for young Australians," Acoust. Austr. **31**, 95–98.

Epps, J., Smith, J. R., and Wolfe, J. (**1997**). "A novel instrument to measure acoustic resonances of the vocal tract during speech," Meas. Sci. Technol. **8**, 1112–1121.

Fletcher, N. H. (**1993**). "Autonomous vibration of simple pressure-controlled valves in gas flows," J. Acoust. Soc. Am. **93**, 2172–2180.

Garnier, M., Henrich, N., Smith, J., and Wolfe, J. (**2010**). "Vocal tract adjustments in the high soprano range," J. Acoust. Soc. Am. **127**, 3771–3780.

Gobl, C. (**1989**). "A preliminary study of acoustic voice quality correlates," STL-QPSR **30**, 9–22.

Hanna, N. (**2014**). "Investigations of the acoustics of the vocal tract and vocal folds *in vivo*, *ex vivo* and *in vitro*," UNSW, Australia and Université de Grenoble, France, https://tel.archives-ouvertes.fr/tel-01174056/ (Last viewed February 5, 2017), p. 60.

Hanna, N., Amatoury, J., Smith, J., and Wolfe, J. (**2016a**). "How long is a vocal tract? Comparison of acoustic impedance spectrometry with magnetic resonance imaging," Proc. Meet. Acoust. **28**, in press.

Hanna, N., Smith, J., and Wolfe, J. (**2012**). "Low frequency response of the vocal tract: Acoustic and mechanical resonances and their losses," in *Proceedings of the Australian Acoustical Society* (T. McGinn, Fremantle, Australia), pp. 317–323.

Hanna, N., Smith, J., and Wolfe, J. (**2016b**). "Frequencies, bandwidths and magnitudes of vocal tract and surrounding tissue resonances, measured through the lips during phonation," J. Acoust. Soc. Am. **139**, 2924–2936.

Henrich, N., Smith, J., and Wolfe, J. (**2011**). "Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and basses," J. Acoust. Soc. Am. **129**, 1024–1035.

Hollien, H., and Michel, J. F. (**1968**). "Vocal fry as a phonational register," J. Speech Hear. Res. **11**, 600–604.

Joliveau, E., Smith, J., and Wolfe, J. (**2004a**). "Tuning of vocal tract resonance by sopranos," Nature **427**, 116.

Joliveau, E., Smith, J., and Wolfe, J. (**2004b**). "Vocal tract resonances in singing: The soprano voice," J. Acoust. Soc. Am. **116**, 2434–2439.

Lulich, S. M. (**2010**). "Subglottal resonances and distinctive features," J. Phonetics **38**, 20–31.

Lulich, S. M., Morton, J. R., Arsikere, H., Sommers, M. S., Leung, G. K., and Alwan, A. (**2012**). "Subglottal resonances of adult male and female native speakers of American English," J. Acoust. Soc. Am. **132**, 2592–2602.

Miller, D. G., Sulter, A. M., Schutte, H. K., and Wolf, R. F. (**1997**). "Comparison of vocal tract formants in singing and non-periodic phonation," J. Voice **11**, 1–11.

Monsen, R. B., and Engebretson, A. M. (**1983**). "The accuracy of formant frequency measurements: A comparison of spectrographic analysis and linear prediction," J. Speech Hear. Res. **26**, 89–97.

Rothenberg, M. (**1981**). "Acoustic interaction between the glottal source and the vocal tract," in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hinano (University of Tokyo Press, Tokyo), pp. 305–328.

Selamtzis, A., and Ternström, S. J. (**2014**). "Analysis of vibratory states in phonation using spectral features of the electroglottographic signal," J. Acoust. Soc. Am. **136**, 2773–2783.

Sondhi, M. M. (**1975**). "Measurement of the glottal waveform," J. Acoust. Soc. Am. **57**, 228–232.

Sundberg, J. (**1975**). "Formant technique in a professional female singer," Acustica **32**, 89–96.

Sundberg, J. (**1981**). "Formants and fundamental frequency control in singing. An experimental study of coupling between vocal tract and voice source," Acustica **49**, 47–54.

Sundberg, J., and Skoog, J. (**1997**). "Dependence of jaw opening on pitch and vowel in singers," J. Voice **11**, 301–306.

Swerdlin, Y., Smith, J., and Wolfe, J. (**2010**). "The effect of whisper and creak vocal mechanisms on vocal tract resonances," J. Acoust. Soc. Am. **127**, 2590–2598.

Titze, I. R. (**1988**). "The physics of small-amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**, 1536–1552.

Titze, I. R. (**2008**). "Nonlinear source-filter coupling in phonation: Theory," J. Acoust. Soc. Am. **123**, 2733–2749.

Titze, I. R., Baken, R. J., Bozeman, K. W., Granqvist, S., Henrich, N., Herbst, C. T., Howard, D. M., Hunter, E. J., Kaelin, D., Kent, R. D., Kreiman, J., Kob, M., Löfqvist, A., McCoy, S., Miller, D. G., Noé, H., Ronald, C., Scherer, R. C., Smith, J. R., Story, B. H., Švec, J. G., Ternström, S., and Wolfe, J. (**2015**). "Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization," J. Acoust. Soc. Am. **137**(5), 3005–3007.

Titze, I. R., Riede, T., and Popolo, P. (**2008**). "Nonlinear source-filter coupling in phonation: Vocal exercises," J. Acoust. Soc. Am. **123**, 1902–1915.

Wade, L., Hanna, N., Smith, J., and Wolfe, J. (**2016**). "Soprano singing, with and without resonances," Proc. Meet. Acoust. **28**, in press.

Zañartu, M., Mehta, D. D., Ho, J. C., Wodicka, G. R., and Hillman, R. E. (**2011**). "Observation and analysis of *in vivo* vocal fold tissue instabilities produced by nonlinear source-filter coupling: A case study," J. Acoust. Soc. Am. **129**, 326–339.

Zhang, Z., Neubauer, J., and Berry, D. A. (**2006**). "The influence of subglottal acoustics on laboratory models of phonation," J. Acoust. Soc. Am. **120**, 1558–1569.