The acoustics of registers and resonances in singing

John Smith (1), Nathalie Henrich (2), Maëva Garnier (2) and Joe Wolfe (1)

(1) School of Physics, University of New South Wales, Sydney NSW 2052, Australia

(2) CNRS GIPSA Lab, Grenoble, France

ABSTRACT

The wide range of the singing voice, from below C2 (65 Hz) to above F6 (1397 Hz), requires a number of strategies that can involve different mechanisms of laryngeal vibration and various adjustments of the vocal tract resonances. The adjustments are made because a vocal tract resonance can boost the radiation of a voice harmonic when it falls close to a resonance frequency. Here we report how singers with different voice categories tune their vocal tract resonances. For the lower voices, the lowest resonance *R*1 is sometimes tuned to a high harmonic, while high voices consistently tune *R*1 to the fundamental over a range of about C5 to C6 (523 to 1046 Hz). The second resonance, *R*2, can be simultaneously tuned to the second harmonic in the pitch range C5 to F5 (523 to 700 Hz). At the very highest pitches, sopranos can no longer increase *R*1 sufficiently and must then switch to adjusting *R*2 so its frequency is close to that of the fundamental.

INTRODUCTION

Singing and speech share much hardware and software, and are thought to have influenced each other, or perhaps evolved from a common (cultural) ancestor (see e.g., Wolfe, 2002; Mithen, 2007). Nevertheless, a number of significant differences mean that singing requires a highly developed set of physical and mental abilities.

In many singing styles, especially Western art music, singers are expected to be able to control pitch and loudness independently. A crescendo or diminuendo on a single note requires precise, compensating adjustments of vocal fold tension and sub-glottal pressure.

The pitch must be controlled with precision, and subtle, sometimes rapid variations are required. A large range of pitch is often required. While artificial musical instruments may achieve multi-octave ranges by using different strings or different vibration modes, the voice has a number of registers or mechanisms, which involve different manners of vibration of the vocal folds. Controlling or disguising the registers and the transitions between them is, for many singers, an important skill.

There is rarely any correlation between the pitch and the phoneme to be produced (some exceptions will be discussed below). Because different phonemes require different vocal tract shapes, they have different radiation efficiencies and present different radiation loads to the vocal folds. This can complicate the constraints mentioned above, as well as the frequent expectation that the timbre of the voice should vary smoothly across the pitch range. Finally, singers are often required to produce a sound level that, without amplification, can compete with powerful accompaniment and still be heard in a large theatre.

This paper describes how the acoustical properties of the vocal system are adjusted by singers to achieve some of these goals.

THE SOURCE-FILTER MODEL

In a simplified model, widely used in speech science, the vocal folds and the glottis (the aperture between them) are considered as a source of acoustic current with frequency components spanning a large range. The vocal tract is considered as an acoustical duct whose resonances give rise to broad spectral peaks or formants in the sound radiated from the mouth. Varying articulations change the frequencies of this filter and thus produce time-varying formants which, together with the changes in amplitude, carry the phonetic information. In the simplest version of this model, the source and the filter are considered to be independent (Fant, 1970): the effects on the source of the variations in the acoustic load and the effects on the tract resonances of different source mechanisms are commonly ignored.

The filter is expected to affect the source, however. A number of analytical and numerical models of the vocal folds predict that the frequency and stability of autonomous vibration depend on, among other variables, the magnitude and phase of the acoustical loads at the glottis (Titze, 1988; Fletcher, 1993). The evidence for this is necessarily indirect, because the relevant properties of the source cannot usually be measured during phonation.

The source can affect the filter, too. For example, the average glottis size is greater in whispering than in normal speech, and this different end effect on the vocal duct has been shown to change the frequencies of resonances (Swerdlin *et al.,* 2010).

LARYNGEAL MECHANISMS

Singers sometimes sing notes lower than C2 (65 Hz) or higher than F6 (1397 Hz). Several different modes of vocal fold vibration are required to cover this wide range (Henrich, 2006). Men usually have longer, more massive vocal folds and consequently lower voices than women or children. Individual singers often have ranges exceeding two octaves.

In M0 (mechanism zero, also called the creak voice or vocal fry), the vocal folds vibrate aperiodically. This produces a broadband spectrum with no clear pitch.

In mechanism M1 (also called modal voice and, in women, the chest voice), the vocalis muscle, which contributes much of the mass of the vocal folds, vibrates with the vocal folds. M1 is usually used for speech. However, M0 may occur at the ends of phrases, and is used more extensively by some speakers. Phonation in M1 is typically very periodic and the relatively rapid closure of the vocal folds give a rapid reduction in the glottal current, which results in a spectrum rich in harmonics.

In mechanism M2 (called the falsetto voice for men and the head voice for women), only the ligament and mucosa covering the vocalis muscle vibrate. M2 may be used for excited speech and indeed some speakers may use it more extensively in normal conversation. M2 is typically highly periodic, but its higher harmonics are usually rather weaker than those of M1.

The highest singing register, used mainly by sopranos specialising in the highest ranges and typically starting somewhere near C6 (1046 Hz), is known as the whistle register. The physiology and acoustics of this register are not well understood, and so the name 'mechanism M3' is not unanimously assigned.

VOCAL TRACT RESONANCES

One may consider the tract as one or two acoustical ducts leading from the glottis to the nose (for nasal phonemes), the mouth (for most phonemes), or to both for partially nasalised speech. We concentrate on non-nasalised vowels, which have a particularly great importance in singing because they are used to sustain notes.

The first resonance, *R*1, has a range from about 300 to 1000 Hz, and so overlaps most of the typical soprano range, and a substantial fraction of all other vocal ranges, with the exception of low basses. *R2* ranges from about 800 to 2200 Hz, and so includes the very high range used by some sopranos.

In both speech and song, the pair of resonances (*R*1,*R*2) produce formants (*F*1,*F*2) and the different vowels are associated with particular values of (*F*1,*F*2).

Usually, the frequencies of vocal tract resonances are estimated from the formants in the speech or song produced. This has the advantage of simplicity, but the disadvantage that its precision is limited to about half the frequency f_0 of the note sung. This approach thus becomes increasingly unreliable in the lower octave of the soprano voice (261 - 523 Hz) and essentially useless for the upper octave (523 - 1046 Hz) (Monson and Engebretson, 1983).

To overcome this problem, we developed the technique of broadband excitation at the mouth (Epps *et al*., 1997; Dowd *et al*., 1998). This has the disadvantage of requiring singers or speakers to perform with a device positioned at the lower lip, but has the advantage of measuring the resonant frequencies with greatly increased precision for high f_0.

This technique is shown schematically in Figure 1. A broadband source of acoustic current is synthesised and produced through a narrow tube at the subject's lower lip.

In a preliminary calibration step, the pressure p_{closed} is measured at the lips with the lips closed. During singing or speech, p_{open} is measured and the pressure ratio $\gamma = p_{\text{open}} / p_{\text{closed}}$ is calculated and displayed as a function of frequency.

RESONANCE TUNING

Figure 2 shows an example of the use of the technique, and also the effect of resonance tuning. Here, an alto sang the vowel in ‘heard’ at two different pitches. The quasi-continuous line on the graph is the measured pressure ratio γ; maxima in this curve indicate the frequencies of the vocal tract resonance. The sharp peaks superposed on the curve are the harmonics of the voice. In the graph at left, she sang the note B3 (nominally 247 Hz) and, in this case, none of the low harmonics fall very close to the resonances.

At right, she sang the same vowel at D4 (294 Hz) and, for this note, the second and fifth harmonics fell close to the first two resonances. In shorthand, we write these two examples of resonance tuning as $R1{:}2f_0$ and $R2{:}5f_0$.

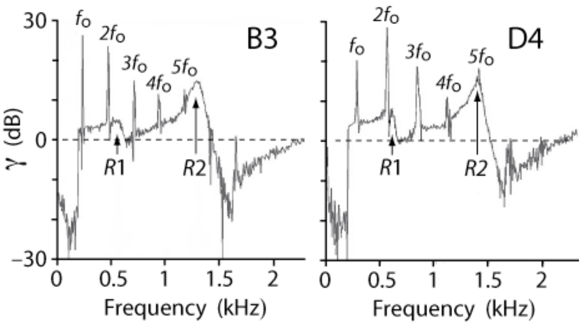


Figure 2. The effect of resonance tuning on the relative magnitude of harmonics in the voice of an alto singing the vowel in 'heard'. From Henrich *et al.* 2011.

Tuning one or more vocal tract resonances at or near the frequency of one of the harmonics of the voice has several potential advantages.

First, the harmonic falling near a resonance is likely to be radiated more strongly from the mouth, providing a helpful power boost to a singer working to be heard over a loud accompaniment. This is particularly important for high voices, where the large spacing between harmonics can mean that one or more tract resonances may lie far from any harmonic.

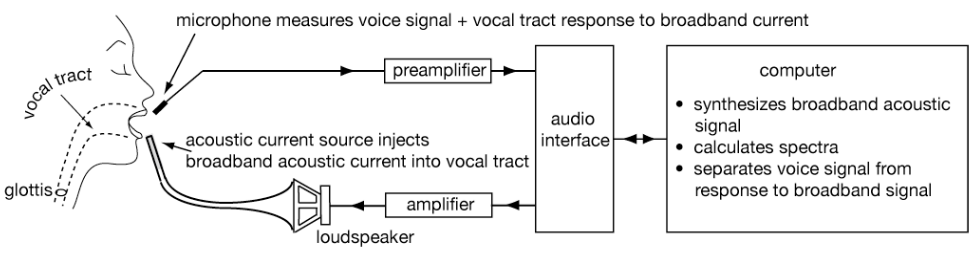


Figure 1. The technique used for real-time masurement of vocal tract resonance.

Second, systematic tuning of a resonance to a harmonic over a range of several notes will give a relatively continuous variation in timbre across the range.

Third, it is possible that the smooth variation in the magnitude and phase of the acoustic impedance loading the glottal source will be less likely to lead to instabilities in the vocal mechanism.

We turn now to the various voice ranges to discuss the possible and observed types of resonance tuning.

BASS AND BARITONE VOICE

Figure 3 illustrates the possible resonances for the bass male voice. In this and subsequent figures, shading is used to show the typical range of the first two tract resonance, diagonal lines $R\text{i} = nf_0$ show the possible tuning relationships between the resonances and the harmonics nf_0. The lines are dashed outside the vocal range, and potentially useful tuning regimes are indicated by double-headed arrows.

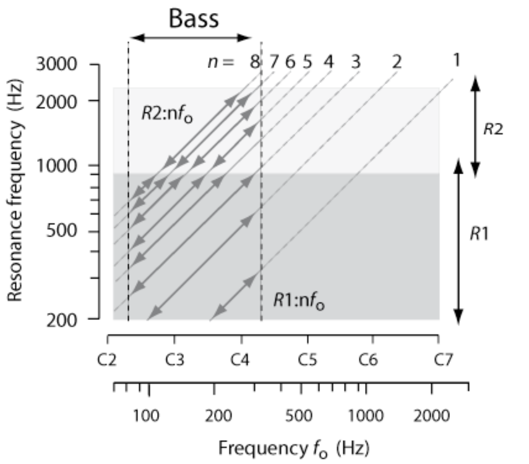


Figure 3. The wide variety of possible resonance tunings available for the bass voice. From Henrich *et al.* 2011.

The 'standard' range expected of basses is from about E2 to E4 (82 to 330 Hz). Over much of this range, several harmonics of the voice fall in the range of $R1$ and $R2$ so, in principle, a range of tuning strategies is possible.

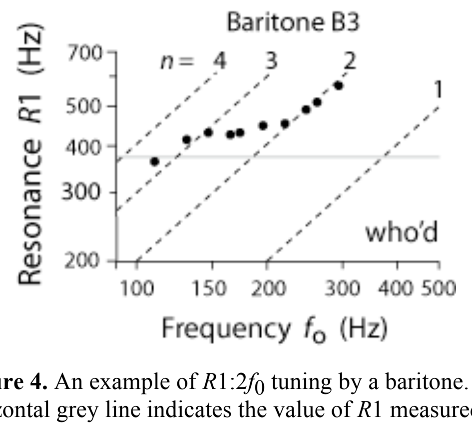


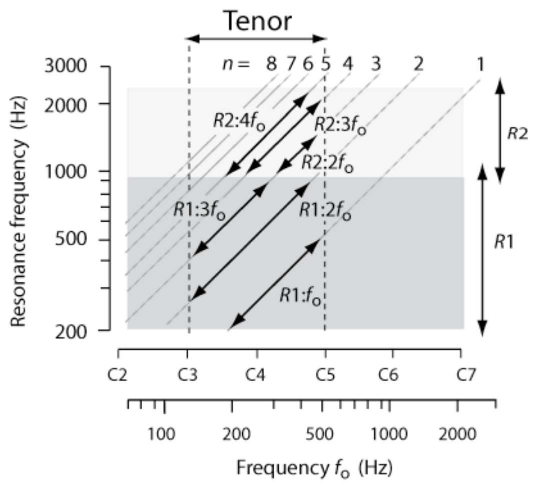
Figure 4. An example of $R1{:}2f_0$ tuning by a baritone. The horizontal grey line indicates the value of $R1$ measured for speech for that singer. From Henrich *et al.* 2011.

In practice however, the relatively broad bandwidth of the resonances and the close spacing of the harmonics mean that at least one harmonic will fall near a resonance even in the absence of a deliberate tuning strategy.

A similar argument could be made for baritones. However, their range is a little higher, their harmonics are therefore a little more widely spaced, and the range of the second harmonic extends further into the range of $R1$. Figure 4 shows an example of $R1{:}2f_0$ tuning measured for a baritone

TENOR VOICE

Tenors, particularly those who sing up to C5 (the 'high C' of tenors at 523 Hz) have a substantial range in which f_0 enters the range of $R1$. Some of the possible resonance tunings are shown in Figure 5.



. **Figure 5.** Possible resonance tunings for the tenor voice. From Henrich *et al.* 2011.

Figure 6 shows some examples. The vowel in the word 'who'd' has a very low value of $R1$ and so overlap of f_0 and $R1$ is greatest for this vowel. For the vowels with larger mouth opening and therefore higher $R1$, $R1{:}2f_0$ tuning is again observed. Figure 6 shows an example where a tenor sequentially uses $R1{:}4f_0$, $R1{:}3f_0$ and $R1{:}2f_0$ tuning as f_0 increases

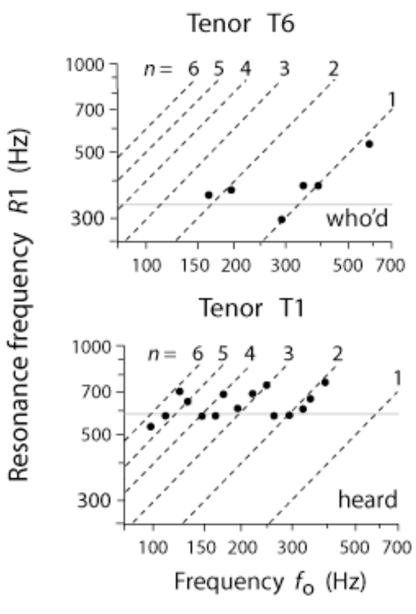


Figure 6. Examples of $R1{:}f_0$, $R1{:}2f_0$, $R1{:}3f_0$ and $R1{:}4f_0$ tuning by tenors. Horizontal grey lines indicate the values of $R1$ measured for speech for the same singer. From Henrich *et al.* 2011.

ALTO VOICE

$R1{:}f_0$ tuning is, in principle, possible across almost all of the traditional alto range. $R1{:}2f_0$ tuning is also potentially very useful, as is shown in Figure 7.

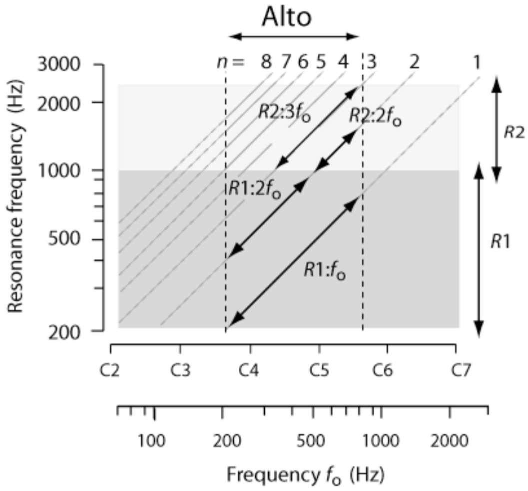


Figure 7. Possible resonance tunings for the alto voice. From Henrich *et al.* 2011.

An example of resonance tuning for an alto was shown in Figure 2. Figure 8 gives an example of one alto using only $R1{:}f_0$ tuning, whereas another switches from $R1{:}2f_0$ to $R1{:}f_0$ tuning for a vowel with lower $R1$ in speech..

Extensive use of $R1{:}2f_0$ tuning by altos occurs in the folk music of some cultures. Because the second harmonic falls in a range where the ear is more sensitive than it is to the first, this can produce very loud voices, as well as an unusual timbre, both of which are observed in at style of Bulgarian women's singing in which $R1{:}2f_0$ tuning by altos is used (Henrich *et al.,* 2007). It can also be used in 'belting', a theatrical singing style (e.g. Bourne and Garnier, 2010).

SOPRANO VOICE

Sopranos have most to gain from resonance tuning, because $R1$ covers almost all of the standard soprano range (C4-C6), as is shown in Figure 9.

Resonance tuning was first suggested, and later confirmed, for the case of sopranos, based upon observations that the mouth opening increased with increasing pitch suggesting that $R1$ was being increased to match f_0 (Lindblom and Sundberg, 1971; Sundberg, 1975; Sundberg and Skoog, 1997). This case was therefore the first to be investigated using broad band excitation (Joliveau *et al.,* 2004a,b). Some results are shown in Figure 10. While singers in other voice categories use several different strategies, the sopranos we have studied, whether trained or not, virtually all use the same strategy over the range from C5 to C6. For this reason, it is appropriate to plot the average values of $R1$ as a function of f_0, as in Figure 10.

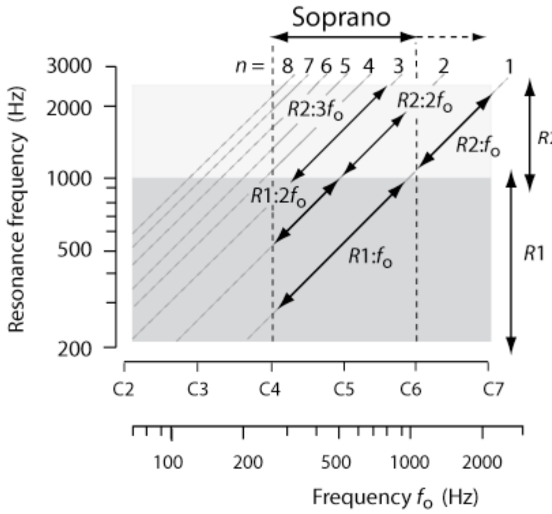


Figure 9. Possible resonance tunings for the soprano voice. From Henrich *et al.* 2011.

.

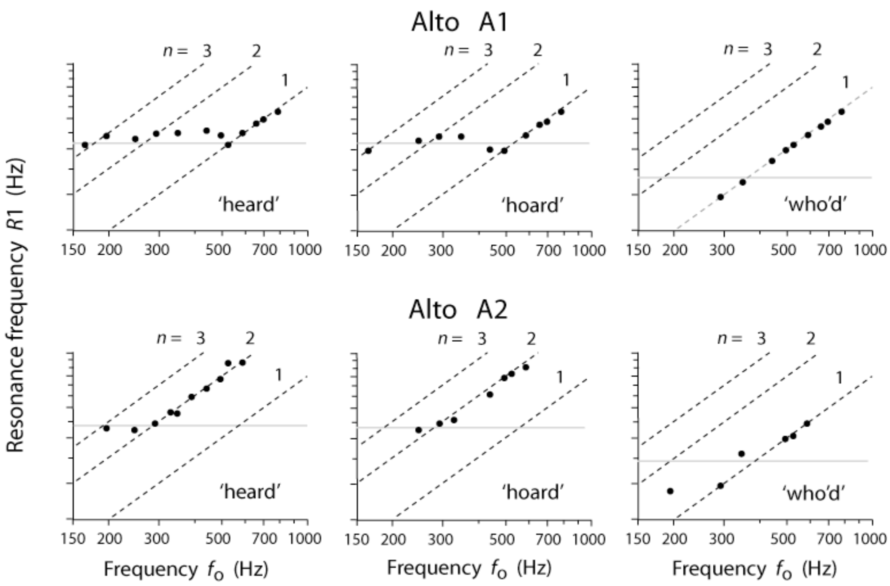


Figure 8. Examples of $R1{:}f_0$ and $R1{:}2f_0$ tuning by altos. Horizontal grey lines indicates the values of $R1$ measured for speech for those singers. From Henrich *et al.* 2011.

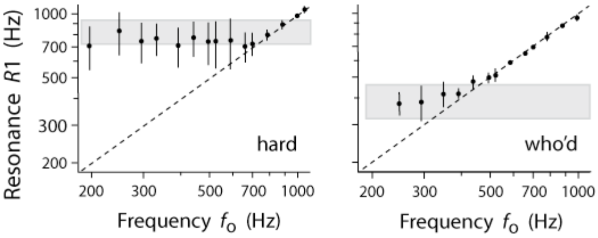


Figure 10. The variation in resonance frequency $R1$ among 10 female singers (altos + sopranos) as a function of frequency on a log-log scale. The dashed diagonal line indicates the relationship $R1 = f_0$. The error bars indicate the standard deviations; they are too small to be shown for several points close to or on the diagonal line indicating $R1{:}f_0$ tuning. The grey shaded areas indicate the range of ± standard deviation of $R1$ measured for these vowels and singers during speech. From Henrich *et. al.,* 2011.

Briefly, in the low pitch range, $R1$ has a value typical of its value for speech, as the shaded bars show. As soon as f_0 reaches this value, the singers tune $R1$ to f_0. The apparent precision suggested by Figure 10 can be examined more closely in Figure 11 which displays the histogram of $R1–f_0$, showing a narrow distribution around zero.

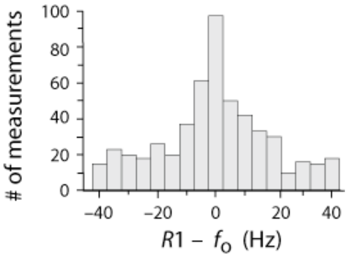


Figure 11. Histograms showing the distribution of the measured values of $R1$ about f_0 for 511 measurements on 27 soprano singing four different vowels. From Henrich *et. al.,* 2011.

TUNING *R*2

Figures 3, 5, 7 and 9 show that $R2$ can, in principle, be tuned to a higher harmonic of the voice. So far, we have only observed this in high voices. An interesting example is shown in Fig. 12, where a soprano simultaneously tunes $R1{:}f_0$ and $R2{:}2f_0$ over a range of most of an octave. While $R1$ is primarily determined by the size of the aperture between the lips, $R2$ is mainly determined by the position of the tongue constriction. However, $R2$ also increases with increasing size of the mouth aperture, which may render this double tuning less difficult that it might initially appear. Indeed this simultaneous tuning has been evident in approx. 40% of our measurements on female singers in the range C5 to F5.

Figure 9 shows that, above about C6 (the sopranos 'high C'), there is only one practical tuning strategy: $R2{:}f_0$. This suggests that coloratura sopranos and the jazz and rock singers who practise the very high range above C6 might use this tuning strategy. Recent studies have confirmed this (Garnier *et al.,* 2010).

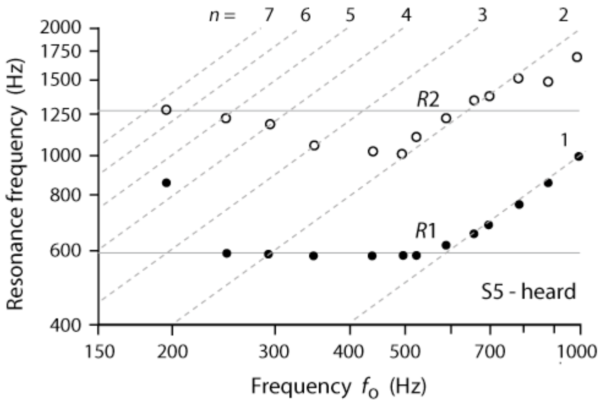


Figure 12. An example of simultaneous $R1{:}f_0$ and $R2{:}2f_0$ tunings by a soprano. The horizontal lines indicate the values of $R1$ and $R2$ measured for this singer and vowel in speech. The dashed diagonal lines indicate the relationships $Ri = nf_0$. From Henrich *et. al.,* 2011.

Figure 13 shows the tuning strategies of one singer studied. At C6, $R1$ has reached its practical maximum. For the next few notes she has no resonance tuning, but then begins $R2{:}f_0$ and continues this for several notes.

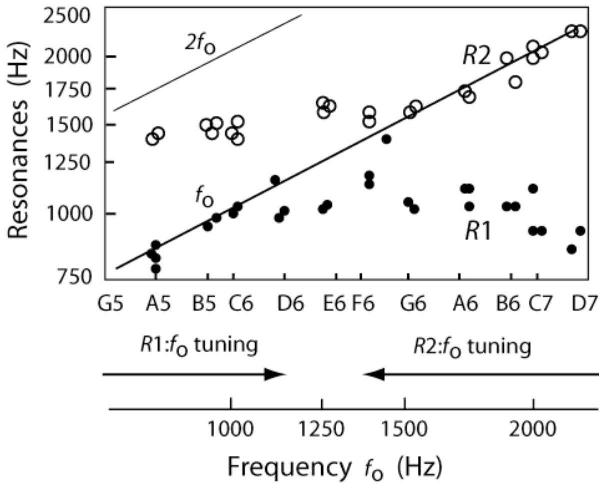


Figure 13. A transition from $R1{:}f_0$ to $R2{:}f_0$ tuning by a soprano as the pitch frequency f_0 increases. Closed and open circles indicate the measured $R1$ and $R2$ respectively. Diagonal lines indicate the relationships $Ri = f_0$ and $Ri = 2f_0$. From Garnier *et al.,* 2010.

What laryngeal mechanism are singers using in this very high range? A number of observations suggest a mechanism different from M2, and tentatively named M3. Unlike the usually discontinuous change (or 'break') from M1 to M2, the M2-M3 transition appears to be continuous and spread over several notes.

Most singers can vary the pitch of the M1-M2 transition by several notes, and thus are able to sing in either M1 or M2 for this range, i.e. an overlap between falsetto and normal voice, or head and chest register. Some sopranos can similarly vary the pitch of the M2-M3 transition, and thus have an overlap range over which they can produce two distinctly different voice qualities (Garnier *et al.,* 2010).

RESONANCE TUNING AND PHONEMES

In speech, as we have mentioned above, the ($R1$,$R2$) combination usually determines the vowel sound. Resonance tuning requires adjusting $R1$ and sometimes $R2$, often over substantial ranges. Obviously, this distorts vowel sounds and can produce a decrease in intelligibility as one vowel sound is replaced with another (Morozov, 1965; Scotto di Carlo and Germain, 1985). To take the most extreme case: as a soprano approaches high C, the mouth aperture must be increased and every vowel must move towards 'ah'. As well as the measurements of this vowel migration, we have also provided sound files illustrating the phenomenon (Music Acoustics, 2011).

ACKNOWLEDGMENTS

We thank the Australian Research Council for support. We are also indebted to the singers who volunteered for this study.

REFERENCES

Bourne, T & Garnier, M 2010, ' Physiological and Acoustic Characteristics of the Female Music Theatre Voice in ‘belt’ and ‘legit’ qualities,' *Int. Symp. Music Acoustics*, Sydney & Katoomba, 2010. J. Smith, editor.

Dowd, A, Smith, JR & Wolfe, J 1998, 'Learning to pronounce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time', *Language and Speech* vol . 41, pp. 1–20.

Epps, J, Smith, JR & Wolfe, J 1997, 'A novel instrument to measure acoustic resonances of the vocal tract during speech', *Meas. Sci. Technol*. vol. 8, pp. 1112–1121.

Fant, G 1970, *Acoustic Theory of Speech Production* (Mouton, The Hague), pp. 15–26.

Fletcher, NH 1993, 'Autonomous vibration of simple pressure-controlled valves in gas flows', *J. Acoust. Soc. Am.* vol. 93, pp. 2172–2180.

Garnier, M, Henrich, N, Smith, J & Wolfe, J 2010, 'Vocal tract adjustments in the high soprano range', *J. Acoust. Soc. Am*. vol. 127, pp. 3771–3780.

Henrich, N 2006, 'Mirroring the voice from Garcia to the present day: Some insights into singing voice registers', *Logoped. Phoniatr. Vocol*. vol. 31, pp. 3-14.

Henrich, N, Kiek, M, Smith, J & Wolfe, J 2007, 'Resonance strategies used in Bulgarian women's singing style; A pilot study,' *Logoped. Phoniatr. Vocol*. vol. 32, pp. 171–177.

Henrich, N, Smith, J & Wolfe, J 2011, 'Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and basses', *J. Acoust. Soc. Am*. vol. 129, pp. 1024-1035.

Joliveau, E, Smith, J & Wolfe, J 2004a, 'Tuning of vocal tract resonances by sopranos', *Nature* vol. 427, p. 116.

Joliveau, E, Smith, J, & Wolfe, J 2004b, 'Vocal tract resonances in singing: The soprano voice', *J. Acoust. Soc. Am.* vol. 116, pp. 2434–2439.

Lindblom, BEF & Sundberg, JEF 1971, 'Acoustical consequences of lip, tongue, jaw, *and larynx movement', J. Acoust. Soc. Am.* vol. 50, pp. 1166–1179.

Mithen, S 2007, *The singing Neanderthals: the origins of music, language, mind and body*, Harvard Univ.Press.

Monsen, RB, & Engebretson, AM 1983, 'The accuracy of formant frequency measurements: a comparison of spectrographic analysis and linear prediction", *J. Speech Hear. Res.* vol. 26, pp. 89–97.

Morozov, VP 1965, 'Intelligibility in singing as a function of fundamental voice pitch', *Soviet Physics– Acoustics* vol. 10, pp. 279–283.

Music Acoustics 2011, *Sopranos: resonance tuning and vowel changes,* viewed 27 September, 2011 www.phys.unsw.edu.au/jw/soprane.html

Scotto di Carlo N, & Germain, A 1985, 'A perceptual study of the influence of pitch on the intelligibility of sung vowels', *Phonetica* vol. 42, pp. 188–197.

Sundberg, J (1975, 'Formant Technique in a Professional Female Singer'. *Acustica* vol. 32, pp. 89–96.

Sundberg, J & Skoog, J 1997, 'Dependence of jaw opening on pitch and vowel in singers', *J. Voice* vol. 11, pp. 301–306.

Swerdlin, Y, Smith, J & Wolfe, J 2010, 'The effect of whisper and creak vocal mechanisms on vocal tract resonances', J. *Acoust. Soc. Am.* vol. 127, pp. 2590–2598.

Titze, IR 1988 'The physics of small–amplitude oscillations of the vocal folds' *J. Acoust. Soc. Am*. vol. 83, pp. 1536–1552.

Wolfe, J. (2002) “Speech and music, acoustics and coding, and what music might be ‘for;”, *Proc. 7th International Conference on Music Perception and Cognition*, Sydney. eds K Stevens, D. Burnham, G.McPherson, E. Schubert, J. Renwick, pp 10-13.