Vowel-pitch matching in Wagner's operas: Implications for intelligibility and ease of singing

John Smith and Joe Wolfe
School of Physics, University of New South Wales, Sydney, New South Wales 2052, Australia
john.smith@unsw.edu.au

Abstract: European vowels are mainly distinguished by the two lowest resonance frequencies ($R1$ and $R2$) of the vocal tract. Once the pitch frequency f_0 exceeds the value of $R1$ in normal speech, sopranos can deliberately "tune" $R1$ to match f_0. This increases loudness, uniformity of tone, and ease of singing, at some cost to intelligibility. Resonance tuning would be assisted if the pitch of the note written for a vowel corresponded with its usual range of $R1$. Analysis of several soprano roles indicates that Wagner aided the acoustics of the soprano voice at high pitch when setting text to music.

PACS numbers: 43.75.Rs [DD]
Date Received: December 24, 2008 **Date Accepted:** February 26, 2009

1. Introduction

In normal speech, the vibrating vocal folds generate a harmonically rich signal with pitch frequency f_0, which interacts with resonances of frequency Ri in the vocal tract to produce sound with a spectral envelope that exhibits broad peaks, called formants,[1] with frequency Fi (Fant, 1970). These resonances have bandwidths of 100 Hz or so, and can be controlled independently of f_0 by varying the position and shape of the tongue, jaw, lips, and larynx (Lindblom and Sundberg, 1971). Vowels in European languages are largely identified by the frequencies of the first two formants ($F1$,$F2$) and thus of the resonances ($R1$,$R2$) that produce them. In adult conversational speech, f_0 is typically in the range 100–300 Hz, whereas the resonances $R1$ and $R2$ lie in the approximate ranges 300–800 and 800–2000 Hz, respectively. Consequently, f_0 and/or some of its harmonics (i.e., $2f_0$, $3f_0$, etc.) usually fall close enough to each resonance to receive a useful power boost and to produce identifiable formants. The situation is similar for most singing ranges when f_0 is less than about 500 Hz (near the "high C" of tenors). However, f_0 for sopranos can range from 250 to 1000 Hz or even higher.

Four problems can arise when f_0 exceeds significantly the normal range of $R1$ for a vowel. First, the acoustic load of the tract on the vocal folds and glottis changes from inertive ("mass-like") when $f_0 < R1$ to compliant ("spring-like") when $f_0 > R1$. Vocal fold vibration can then become less efficient and less stable (Titze 1988, 2008; Titze *et al.*, 2008).

Second, the sound level is usually reduced, as little acoustic energy will be radiated at frequency f_0. Third, a strong variation in the amplitudes of the fundamental and/or second harmonic may occur as the pitch changes, producing possibly undesirable discontinuities in timbre. Finally, the effective absence of $F1$ means that vowels with a similar value of $F2$ become indistinguishable. Indeed, as f_0 increases, the spacing of harmonics can become so large that even $F2$ may also effectively disappear.

For a soprano, a solution to all but the last of these problems is to tune $R1$ close to, but slightly above, f_0 (Sundberg, 1975; Joliveau *et al.*, 2004a, 2004b; Titze *et al.*, 2008). This should increase power, increase the ease of singing, and help maintain timbral homogeneity. Accordingly, sopranos often "modify" their vowels when singing at high pitch (e.g., Coffin, 1974, 1976).

This solution, called resonance tuning, has no disadvantages when singing *vocalise* because there is no textual information. However, when singing text, $F1$ will consequently be similar to $R1$ and no longer have an appropriate value for many phoneme-pitch combinations. This increases the probability that vowels are confused as the difference between f_0 and the

value of $R1$ in normal speech increases (Morozov, 1965; Scotto di Carlo and Germain, 1985; Hollien *et al.*, 2000). In an earlier study (Dowd *et al.*, 1997), we found that correct recognition in speech decreases exponentially with displacement on the $(R1, R2)$ plane. (For French, the characteristic length is 0.36 times the average distance between vowels.) Thus, with resonance tuning, a vowel would be likely to be confused with another of higher $R1$ once f_0 exceeded its normal $R1$ by about 100–200 Hz. This confusion could be minimized if a composer took advantage of vowel-pitch matching, i.e., if each vowel were sung with a fundamental frequency f_0 that was consistent with its usual range of $R1$. Further, the singer would no longer need to "tune" the resonance significantly (Carlsson-Berndtsson and Sundberg, 1991): the libretto would partially do it for her, making it easier to sing with a high ratio of output power to input effort.

Composers have long been aware of the problem of reduced intelligibility (Berlioz, 1844). But have they used vowel-pitch matching (whether consciously or unconsciously) to reduce it and/or to make the music easier to sing? If so, the vowel distribution would vary systematically with pitch: vowels associated with a low $R1$ would be sung less often at high pitch and vice versa. If not, we should expect the distribution of vowels to be independent of pitch, simply reflecting that of the libretto. A test of this hypothesis would require libretti with a large number of vowels in different words to ensure that any observed distribution was not a statistical accident. Compositions wherein the intelligibility of the text is important should also be more likely to show a non-uniform distribution of vowels with pitch. Not all operas are thus suitable: in some, the high pitch writing for soprano is aimed at technical display, rather than conveying text. In others, a relatively simple plot is adequately conveyed by repetition or by actions on stage. A composer who was also the librettist might be most likely to match vowels with pitch. For these reasons, this study focused initially on the two great Wagnerian soprano roles; Brünnhilde and Isolde. Each singer has to communicate lengthy, subtle aspects of plot via text alone. Further, the libretti were written by Wagner, published prior to composing the scores, and considered by him and others to be significant literary works in their own right. Four operas by other composers were then studied for comparison.

2. Methods

Using published scores, the phoneme and fundamental (pitch) frequency f_0 were recorded for each note sung by the chosen soprano(s) in each operas studied—see Table 1. Any obvious ornamentation, grace notes, trills, and mordants were not included as they carry no textual information. (Because resonances and formants are broad, high precision in fundamental frequency is not required, so $A4 = 440$ Hz was assumed throughout).

Wagner provided little ornamentation and generally writes only one note per syllable, (sometimes two for Isolde). To simplify presentation, the 12 vowels of German were grouped into the four standard categories according to their jaw height in the vocoid or Cardinal Vowel space. The range of $R1$ we associated with each category were as follows: *closed* 250–400 Hz, *close-mid* (or *half-close*) 400–550 Hz, *open-mid* (or *half-open*) 550–750 Hz, and *open* 750–1000 Hz. Of course the values of $R1$ will vary with language, dialect, accent, etc. It is also possible that listeners may learn to use a different formant "map" for sopranos (i.e., a different categorization of the vowel plane), in much the same way that we use different maps for men, women, and children. Ultimately, we can only make an informed guess at the vowel sound imagined by the composer-librettist. Although there might be uncertainty about the range of resonance frequencies in each category, the important feature for this study is that their order with increasing frequency is known. The soprano roles of Fiordiligi in *Così fan tutte* and Sophie in *Der Rosenkavalier* were analyzed in a similar fashion.

A slightly abbreviated analysis was made of the soprano roles in *Don Giovanni* by Mozart and *The Barber of Seville* by Rossini, which included only notes longer than a crochet (quarter note) or a quaver (eighth note) in slow sections, in the range from G4 to C6, a range below which resonance tuning by sopranos is not observed (Joliveau *et al.* 2004a, 2004b). There are thus no data on vowel-pitch matching for the closed vowels in these two operas.

Table 1. Details of the operas.

Opera	Composer	Librettist	Year	Language	Soprano role
Der Rosenkavalier[a]	Strauss	von Hofmannsthal	1909–1910	German	Sophie
Der Ring des Nibelungen[b,c]	Wagner	Wagner	1854–1874	German	Brünnhilde
Götterdämmerung[b]	Wagner	Wagner	1869–1874	German	Brünnhilde
Siegfried[c]	Wagner	Wagner	1856–1869	German	Brünnhilde
Tristan und Isolde[d]	Wagner	Wagner	1857–1859	German	Isolde
Die Walküre[c]	Wagner	Wagner	1854–1856	German	Brünnhilde
The Barber of Seville[e]	Rossini	Sterbini	1816	Italian	All
Così fan tutte[f]	Mozart	da Ponte	1790	Italian	Fiordiligi
Don Giovanni[g]	Mozart	da Ponte	1787	Italian	All

[a]The edition of the score used was from Boosey & Hawkes, London, 1943.
[b]The edition of the score used was from Klavierauszug-VEB Breitkopf und Härtel Musikverlag, Leipzig, 1980.
[c]The edition of the score used was from Klavierauszug-VEB Breitkopf und Härtel Musikverlag, Leipzig, 1984.
[d]The edition of the score used was from Ernst Eulenberg, London, 1900.
[e]The edition of the score used was from Kalmus 368, New York.
[f]The edition of the score used was from Neue Ausgabe sämtlicher Werke, Bärenreiter Kassel, Basel, 1991.
[g]The edition of the score used was from Kassel Bärenreiter, 1975.

3. Results and discussion

The upper part of Fig. 1(A) shows, for each pitch, the total number of notes sung for the combined soprano roles of Brünnhilde and Isolde. Each note has a particular pitch and an associated vowel. There are sufficient of these vowel-pitch combinations (more than 10 000) to provide useful statistics and a reasonably smooth distribution with pitch, although some preference for keys harmonically close to C major is apparent. Fewer notes are written at the extremes of the soprano range, presumably because very high notes are physically demanding and dramatically effective if used sparingly, and because sopranos' low notes often have reduced dynamic range.

The lower part of Fig. 1(A) shows the distribution of vowels grouped into the four categories for jaw height and thus $R1$. The vowel distribution does vary with frequency, and in a fashion that helps match $R1$ to f_0. Thus the closed and close-mid vowels with low $R1$ become less common as f_0 rises above 500–600 Hz. Conversely, the fraction of open-mid and open vowels increases significantly above 600 Hz. The open vowels are, of course, preferred at high pitch, whereas other vowels would be seriously distorted by resonance tuning. However, open vowels are also used across the whole pitch range. The unmodified $R1$ for these vowels is sufficiently high that, for almost the whole of the soprano range, the vocal tract load at the glottis is inertive. Further, for notes in the low soprano range, $R1$ may be excited by a second or even third harmonic. So the open vowels are least likely to be distorted by resonance tuning.

Figure 1(B) shows the data for the vowel-pitch combinations for the role of Fiordilgi in *Così fan tutte*, an opera buffa by Mozart. There is no significant variation of vowels with pitch, except surprisingly at low notes where the closed vowels (u and i) are less likely than at the highest notes; this could reduce intelligibility.

Figure 2 shows the extent of pitch-resonance matching for the eight operas studied in terms of γ, a parameter we define as follows. For all notes lying in a frequency band corresponding to a particular jaw height, let g be the average fraction of notes whose vowel corresponds to that jaw height. Let h be the average fraction of notes at all other frequencies having those vowels. The parameter $\gamma = g/h - 1$ then indicates the preference for the appropriate vowel-pitch combinations. Positive and negative values indicate favorable and unfavorable pitch-resonance matching, respectively. The vowel distribution is seen to change systematically with frequency

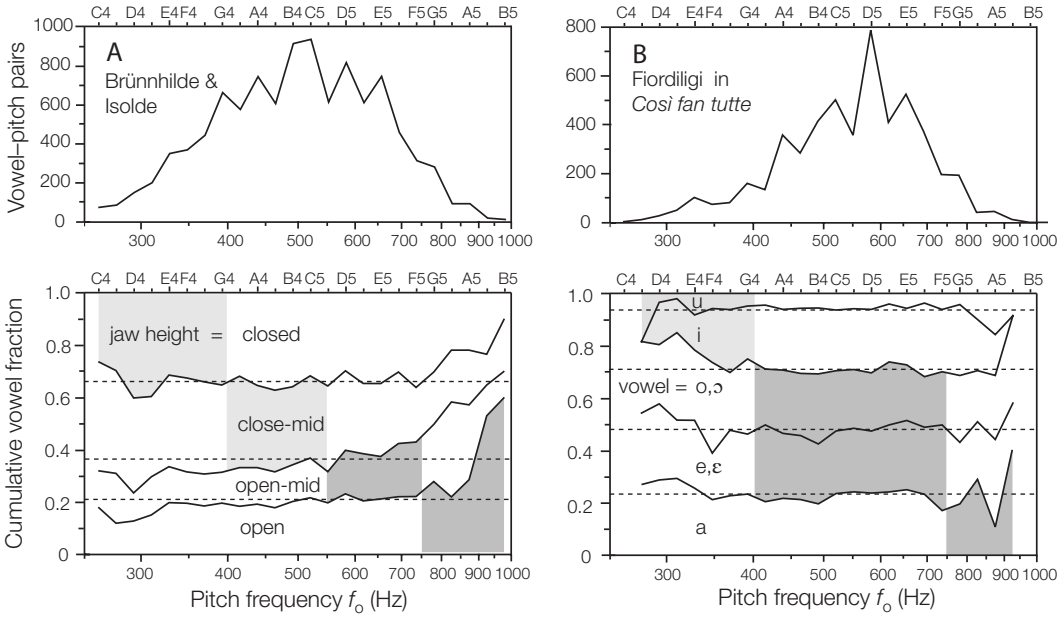

Fig. 1. Semi-logarithmic plots of the number of vowel-pitch combinations (upper) and the cumulative vowel fraction (lower) as function of written pitch frequency f_0 for soprano roles in operas by Wagner and Mozart. On the lower figures, the open vowels lie between the axis and the lowest continuous line, open-mid between the two lowest continuous lines, etc. The dashed lines show the null hypothesis: equal distribution of vowels across pitch. Shaded areas indicate regions where $R1$ and pitch might be favorably matched. Heavily shaded areas indicate approximate regions where resonance tuning would be particularly favorable.

in all four Wagner operas studied. The direction of the changes strongly supports the hypothesis of non-accidental matching of vowels with pitch, but it is worth considering non-acoustic causes. For example, high notes are often used for dramatic emphasis, so important declarative text might alter the distribution via preferential repetition at high pitch. Isolde often sings "Tristan" and "Liebe," usually with descending pitch. Brünnhilde is also heavily involved with "Sieglinde," "Siegmund," "Siegfried," and "Liebe." However, removal of these words from the analysis did not significantly alter the overall distribution of vowels with frequency. The data presented in the figures do not include the war-whoops of Brünnhilde in *Die Walküre* because, although her repeated cries of "Hojotoho Heiaha" occur at high pitch, they convey negligible textual information. Their inclusion does not significantly alter the vowel distribution.

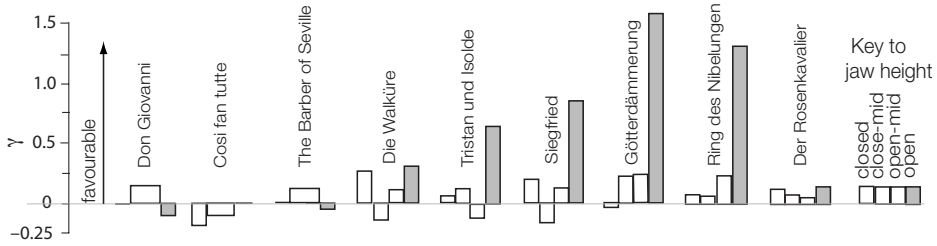

Fig. 2. The extent of pitch-resonance matching for soprano roles in the eight operas studied. The degree of matching is indicated by γ, a parameter that indicates the preference for the appropriate vowel-pitch combinations (see text). Positive and negative values of γ indicate favorable and unfavorable pitch-resonance matching, respectively. Values associated with open vowels (high $R1$) are indicated by shading. The operas are shown in historical order from left to right.

Do other composers or librettists match vowels with pitch? Figure 2 indicates that, for the limited sample of works studied, a significant increase in γ for open vowels occurs only in the operas by Wagner. Interestingly, γ for open vowels increased systematically as Wagner's experience as a composer increased. For comparison, the role of Sophie in the later opera *Der Rosenkavalier* by Richard Strauss shows no significant effect.

Wagner's idea of opera was a continuous music drama. Earlier operas often linked separate arias and choruses with explanatory recitative and thus had less need for intelligibility at high pitch. Furthermore Wagner wrote for much larger orchestras than those available to Mozart or Rossini, and wrote vocal parts that severely test the stamina and capabilities of singers. Thus the employment of vowel-pitch matching could have helped satisfy the concomitant requirements of intelligibility, vocal power, and easier singing of difficult parts.

A number of complications and differences are also relevant: in some operas, important phrases can be repeated several times at different pitches—rarely in Wagner and Strauss, but often in Mozart. One might also consider the time available to composers to "polish" the operas; Rossini wrote some 26 operas in 7 years whereas Wagner wrote 14 in over 50 years.

It is also possible that a composer-librettist might alter the overall vowel distribution from that of normal speech to favor more open vowels, particularly when writing text for orchestral accompaniment. First, singers can produce higher sound pressure levels using open vowels at any pitch (Gramming and Sundberg, 1988). Furthermore, the time-averaged power spectrum produced by an orchestra peaks around 500–600 Hz and then decreases with increasing frequency (Sundberg, 1977). Consequently, the voice of a singer using more open vowels would be expected to have more power above 500 Hz through reinforced harmonics, even when singing at lower pitch. Thus a libretto to be sung with orchestra might advantageously include a higher proportion of open vowels than written text. This was not investigated here, in part because of the difficulty in choosing appropriate texts to use as the non-operatic controls.

Many factors, apart from vowels, are involved in intelligibility. For instance, if only one vowel in a string of phonemes produces a real word, then the vowel need not be recognized. Similarly, context in a sentence can mean that a weird can be understood, independently of the vowel. The results of this study also suggest that those translating a libretto for performance in another language might occasionally consider vowel-pitch matching—among all of the other constraints.

4. Conclusions

The authors are unaware of any written evidence about the composers' intentions nor of whether they were advised on this issue by sopranos, with whom they sometimes had quite close relations. However, it appears that Wagner, either consciously or unconsciously, did take the acoustics of the soprano voice at high pitch into account when setting text to music. This is consistent with the increased importance of textual information in his operas, the increasing size of his orchestras, and the more complex vocal parts.

Acknowledgment

The support of the Australian Research Council is gratefully acknowledged.

References and links

[1]We follow the original definitions of Fant (1970) and reserve the term "formant" for broad peaks in the spectral envelope and "resonance" for the acoustic resonances of the vocal tract that produce them—see Wolfe *et al.* (2009) for further discussion.

Berlioz, H. (**1844**). *Grand traité d'instrumentation et d'orchestration modernes (A treatise on modern instrumentation and orchestration)*, translated by M. C. Clarke (Novello, London).
Carlsson-Berndtsson, G., and Sundberg, J. (**1991**). "Formant frequency tuning in singing," Speech Transm. Lab. Q. Prog. Status Rep. **32**, 29–35.
Coffin, B. (**1974**). "On hearing, feeling and using the instrumental resonance of the singing voice," NATS Bulletin **31**, 26–30.

Coffin, B. (**1976**). *Coffin's Sounds of Singing* (The Scarecrow, Lanham).
Dowd, A., Smith, J., and Wolfe, J. (**1997**). "Learning to pronounce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time," Lang Speech **41**, 1–20.
Fant, G. (**1970**). *Acoustic Theory of Speech Production* (Mouton, The Hague).
Gramming, P., and Sundberg, J. (**1988**). "Spectrum factors relevant to phonetogram measurement," J. Acoust. Soc. Am. **83**, 2352–2360.
Hollien, H., Mendes-Schwartz, A. P., and Nielsen, K. (**2000**). "Perceptual confusions of high-pitched sung vowels," J. Voice **14**, 287–298.
Joliveau, E., Smith, J., and Wolfe, J. (**2004a**). "Tuning of vocal tract resonances by sopranos," Nature (London) **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.
Lindblom, B. E. F., and Sundberg, J. E. F. (**1971**). "Acoustical consequences of lip, tongue, jaw, and larynx movement," J. Acoust. Soc. Am. **50**, 1166–1179.
Morozov, V. P. (**1965**). "Intelligibility in singing as a function of fundamental voice pitch," Sov. Phys. Acoust. **10**, 279–283.
Scotto di Carlo, N., and Germain, A. (**1985**). "A perceptual study of the influence of pitch on the intelligibility of sung vowels," Phonetica **42**, 188–197.
Sundberg, J. (**1975**). "Formant technique in a professional female singer," Acustica **32**, 89–96.
Sundberg, J. (**1977**). "The acoustics of the singing voice," Sci. Am., 82–91.
Titze, I. R. (**1988**). "The physics of small-amplitude oscillations 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., Tobias, R., and Popolo, P. (**2008**). "Nonlinear source-filter coupling in phonation: Vocal exercises," J. Acoust. Soc. Am. **123**, 1902–1915.
Wolfe, J., Garnier, M., and Smith, J. (**2009**). "Vocal tract resonances in speech, singing and playing musical instruments," HFSP J., **3**(1), 6–23.