An Investigation of the Acoustic Vowel Space of Singing

Evan Bradley
Department of Linguistics and Cognitive Science, University of Delaware

Abstract

speaking and singing are two modes of the same system. These modes are subject to similar constraints, but have different goals. This study examined the acoustic vowel spaces, as defined by formant frequencies, used by singers in each mode. Differences between the modes is partialy explained by known articulatory processes used during singing.

Introduction

Speaking and singing use the same vocal apparatus, but to very different effect. While the articulatory and acoustic properties on speech cause the speaker to balance perceptibility and articulatory effort concerns, these pressures are moderated in singing by additional concerns for consistent resonance, expression, and style, and there is some evidence that vowels undergo articulatory modification in singing (Howard & Collingsworth, 1992; Wray et al., 2003). In addition, the acoustic properties of singing vary by gender, singer, and singing style (Bloothooft & Plomp, 1986); vocal training (Sundberg et al., 2005); and whether the singing is infant-directed (Trainor, 1997). This study aims to directly compare the acoustic properties of spoken and sung vowels, and to compare the vowel spaces of spoken and sung registers in order to better determine the factors which influence vowel modification in singing.

Sundberg (1987) provides a review of research on the articulation and acoustic properties of singing, especially as they relate to characteristics of the singer.

Gender

Males and females differ in average phonation frequency and vocal tract length. This difference in vocal tract length, however, is not to scale. The average mouth length of a female is 85% that of the average male, while the average female pharynx length is only 77% that of the average male, meaning that the pharynx to mouth ratio differs between men and women. This partially, but not fully, explains differences in formant frequencies between men and women, which may partitially be accounted for by 'sexolects', or gender-dependent articulation. Although fundamental frequency is the largest determininant of the

This project would not have been possible without the assistance of the University of Delaware Vocal Music Department and the Xi Mu Chapter of Phi Mu Alpha Sinfonia

gender of a speaker, differences in formant frequencies also contribute to the voice quality of men and women, and consequently to vocal timbre in singing. This is due in large part to differences in the fourth formant resulting from women's narrower larynx tubes; alto singers, for instance, have a higher fourth formant than do tenors, even though the lower formant frequencies are similar. Because the third and fourth formants are more independent of vowels than are the first and second, these formants are more similar across a range of articulations, and their proximity in males contributes to a harsher vocal quality, by increasing the amplitude of the partials between them as they near.

Within males, voice category (*e.g.*, tenor or baritone) also contributes to differences in formant frequencies (in addition to phonation frequency), and individual vowels appear to be articulated differently by tenors, baritones, and basses. These differences seem to be similar to the differences between males and females, which suggests that tenors and basses have different pharynx lengths.

Larynx Height

Larynx height varies in normal speech, and is associated with vowel identity. The larynx tends to be raised during the pronunciation of vowels produced with spread lips (such as /i/) and lowered during the pronunciation of rounded vowels. Sundberg explains that this is because the acoustic effects of changes in lip rounding and larynx height are similar, so adjusting larynx height reduces the amount of lip rounding needed and makes articulation easier. Larynx height also increases with phonation frequency in speech, as well as in untrained singing. Trained singers, however, aim to maintain a generally low larynx position, and in fact, larynx height in trained singers decreases slightly as pitch rises. A lowered larynx lengthens the vocal tract and consequently lowers its resonances.

The effect of larynx lowering causes the first and second formant frequencies of most vowels to be lowered, toward a set of values similar to /œ/. The effect on first formant frequency is greatest for /a/ and /æ/ (the lowerst vowels, and hence those with the highest f_1), and the effect on second formant frequency is greatest for /i/ and /e/ (front vowels with high f_2). Variance in larynx height specifically changes the length of the pharynx tube, and the fact that the formant changes seen here are similar to the differences between men and women suggest that pharynx length contributes to this sex difference. The fourth resonance of the vocal tract is determined by its length, but the exact fourth formant frequency is primarily determined by the shape of the larynx tube. The fourth formant is a major factor in determining individual voice timbre. Larynx height and the fourth formant also play a role in producing the *singer's formant*.

Singer's Formant

One result of this manipulation of larynx height is the production of the so-called *singer's formant*. In addition to the effects on the first and second formants, a lowered larynx has differential effects on the higher formants. In male opera singers, the fourth formant becomes much closer to the third than in speech, and the fifth formant lowers to a level at or below that of the fourth formant in speech. This clustering of formants greatly increases the energy in this frequency range (around 3,000Hz)

One probable reason for the production of the singer's formant is loudness. Although trained singers do produce a higher sound pressure level than novices, they can increase the perceived loudness of their voice by manipulating formants to increase energy within certain parts of the frequency spectrum. The singer's formant, which may be around 3,000Hz, is much higher than the spectral peak of an orchestra, which partly explained why an unamplified opera singer can still be heard above the accompaniment of the orchestra. This seems to be a more efficient means of increasing loudness than only increasing total sound pressure level. The singer's formant is affected by vowels, as well; this formant in vowels with a high second formant (*e.g.*, /e/, /i/) has a greater amplitude than that in vowels with a lower second formant (*e.g.*, /u/).

The singer's formant is more prominent in males than in females, and more prominent in altos than in sopranos. This may be due to the high phonation frequencies sung by female singers, which may exceed the formant frequencies of their vocal tract. Instead of 'wasting' these resonances, sopranos appear to use an alternative strategy to increase loudness; instead of using larynx height to produce the singer's formant, they increase the jaw opening to raise the first formant to the level of phonation frequency.

Questions

1. Manipulation of larynx height and other song-specific articulations appear to have differential effects on each vowel. How does this change the overall acoustic vowel space and the relationship between vowel classes as compared to speech?
2. Do sex, vocal training, and other factors influence or determine the acoustic vowel space of a singer?

In order to address these questions, the formant frequencies of spoken and sung vowels were measured from the same speaker/singers, reading or singing passages in a naturalistic way, with particular attention to the first and second formant, which define the traditional acoustic vowel space. The previous findings discussed above regarding the consequences of larynx lowering lead to several predictions about the vowel space of singing

1. Overall, formants will be lower in sung vowels, as compared to spoken vowels.
2. As vowels migrate, the variance in formant values will be lower in singing than in speech.
3. low vowels will show the greatest f_1 changes, while front vowels will show the greatest f_2 changes.
4. if larynx control requires training, then amateur singers should have vowel spaces which are more similar to their speaking voices than are those of trained singers.

Procedure

Singers

Five singers (3 female) were recruited from the Unviersity of Delaware community. All were between 18 and 23 years of age. Singers were classified as 'professional' or 'amateur' based on their previous vocal training. All singers had choral or solo singing experience, but only singers who had received individual instruction in solo singing were classified as 'professional'. All professional singers were studying voice at the college level. Table **??** summarizes the characteristics of the singers studied.

subject	sex	range	experience
1	female	soprano	professional
2	male	tenor	professional
3	female	soprano	professional
4	male	baritone	amateur
5	female	alto	professional

Table 1: Singer characteristics. 'professional' singers had at least five years of solo vocal training. 'Amateur' singers had various levels of singing experience, but no formal, individual vocal training

Music

The piece selected was an American folksong, *Shenandoah* (**?**) The melody is presented in figure 1. The piece was chosen for its simple melody, comfortable range for many voice types, and ballad style, which allows for easy vowel measurement. The lyrics were adapted by the author to create additional tokens of the vowels of interest (figure 2).

Recording and analysis

Singers were recorded in a sound-attenuated booth at the Experimental Psycholinguistics Laboratory. One singer (singer number 2) was recorded in an insulated practice room of the UDel Center for Fine Arts. All singers were recorded with a Labtec AM-22 microphone connected to a laptop computer. Singers were instructed first to read the lyrics

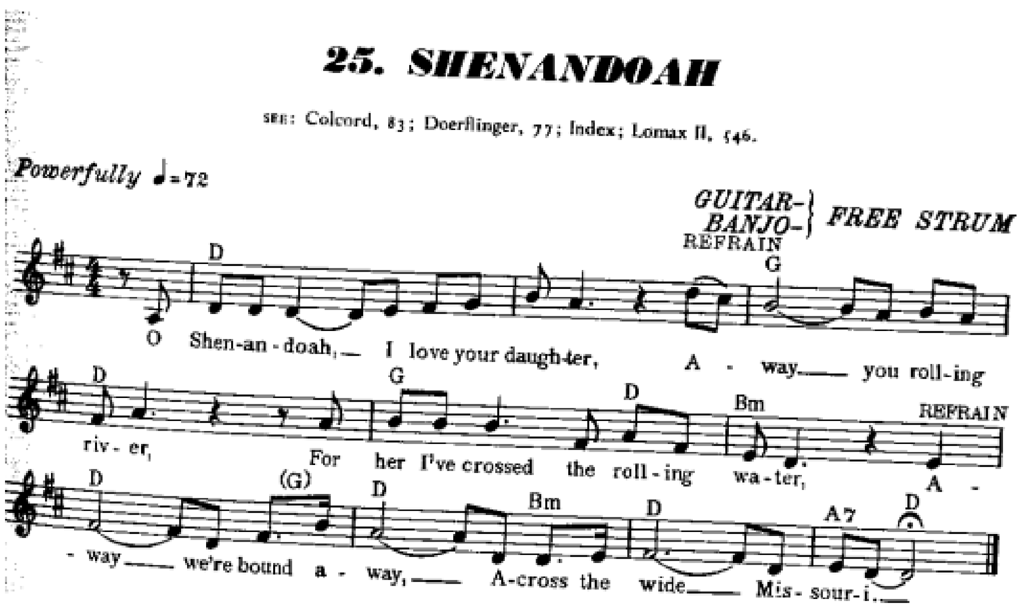

Figure 1. melody of *Shenandoah*, from Lomax (1960, p. 53)

O Shenandoah,
o ʃenændo
I long to hear you.
aɪ lɔŋ tu hir ju
Away, you rolling river.
əweɪ ju ɹolɪŋ ɹɪvɚ
O Shenandoah,
o ʃenændo
I long to hear you.
aɪ lɔŋ tu hir ju
Away, we're bound away,
əweɪ wɪɹ baʊnd əweɪ]
Across the wide Missouri.
ʌkɹɔs ðə waɪd mɪzʊɹi

Figure 2. adapted lyrics and IPA transcription of *Shenandoah*

of the piece, as if to an audience in the manner of a poem. Then, they sang the piece, as if for an *a cappella*, unamplified performance.

Recordings were transcribed and segmented. Pitch and formant frequency values of each vowel were measured using Praat (Boersma & Weenink, 2009). Monophthongs were measured near the midpoint of the vowel. Each segment of a diphthong was measured independently (though the second portion was not included in the analyses discussed here). Frequencies for pitch, and the first four formants were recorded. Recording quality was not sufficient to consistently measure the fifth formant.

Results

Because the first and second formants are the most important for determining the identity of a vowel, this anaysis will focus on these. I will also focus on a set of cardinal vowels (a, e, i, o, u), and ignore reduced vowels, non-point vowels, and diphthongs.

female singers

Female singers overall showed a decrease in f_1 and f_2 for nearly every vowel (figure 3). As predicted, /a/ showed the greatest reduction in f_1 during singing. /e/ and /u/ showed the greatest reduction in f_1, and /i/ showed little change at all.

Most vowels also showed a small reduction in the standard deviations of their formant values, and overall variance in f_1 decreased from 192Hz to 135Hz, and f_2 from 498Hz to 417Hz.

Tables 2 and 3 summarize the first and second formant values for all female singers.

Singer 1 illustrates each of these patterns in her individual vowel measurements. Figure 4 illustrates the shift in f_2 (but not in f_1) for high vowels; /u/, which is speech is produced centrally, and so shows overlap with /i/, lowers its f_2, while /i/ hardly changes.

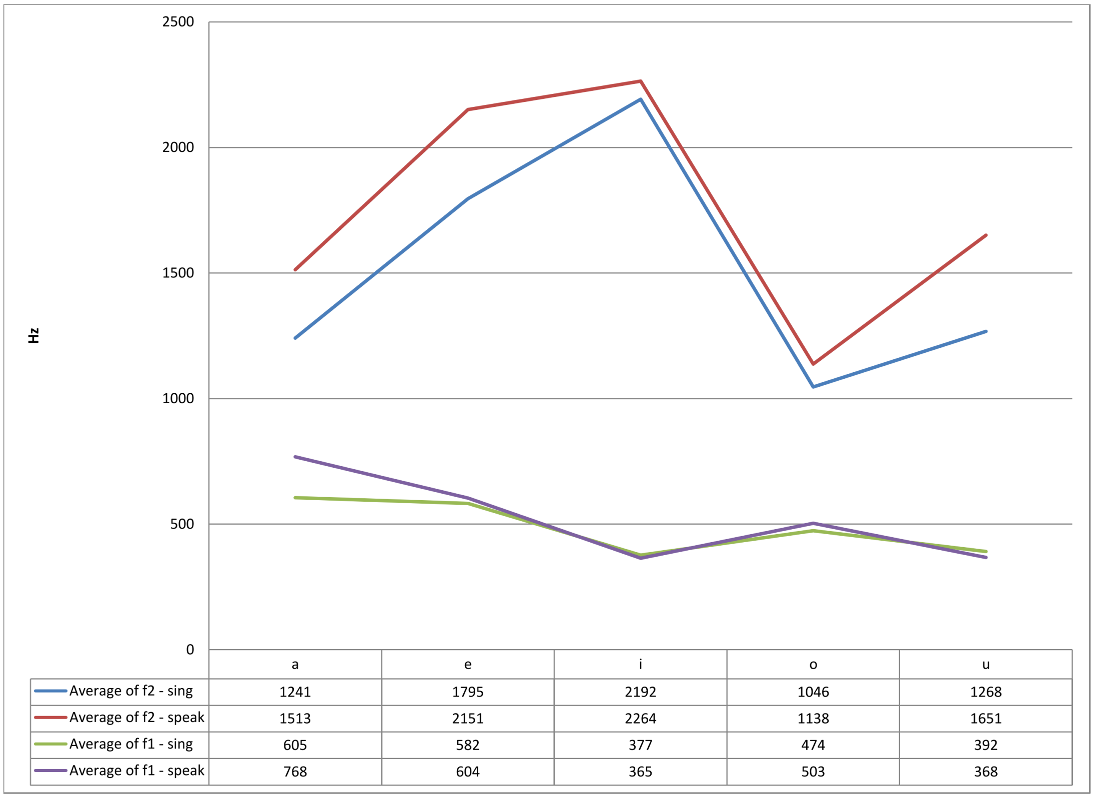

Figure 3. formant values during speaking and singing, females

vowel	f_1	f_2
æ	808	1783
ʊ	538	1138
a	768	1513
e	604	2151
i	365	2264
o	503	1138
u	368	1651
	559 (192)	1626 (498)

Table 2: first and second formant values during speaking for females

vowel	f_1	f_2
æ	611	1370
ʊ	609	1094
a	605	1241
e	582	1795
i	377	2192
o	474	1046
u	392	1268
	503 (135)	1392 (417)

Table 3: first and second formant values during singing for females

vowel	f_1	f_2
æ	636	1555
ʊ	544	1417
a	653	1305
e	533	1781
i	358	1978
o	535	2032
u	352	1724
	514 (176)	1717 (556)

Table 4: first and second formant values during speaking for males

Figure 5 illustrates a shift in f_2 for /e/ as well as /o/. Figure 6 illustrates a shift in both formants for /a/ and /o/; sung /a/ moves into the normal speaking range of /o/.

the total f_1/f_2 space used by the singer is decreased overall, and the space shifts toward the low f_1, low f_2 corner of the space (figure 7).

male singers

The two male singers analyzed did not show exactly the same pattern as the females (8. First formant values differed little between speaking and singing, while the second formant showed the greates difference for /o/. The male singers also showed greater variance in formants during singing.

Tables 4 and 5 summarize the first and second formant values for all male singers.

Singer 4, a musician who was not primarily a singer, showed a pattern of vowel migration similar to that reported above for female singers (figure 9); /u/ undergoes a large f_2 lowering during singing. His /a/ seems to have an already low f_1 during speech, and so does not shift much during singing.

Singer 2, a trained oepratic tenor, shows nearly an opposite effect, with /u/ showing

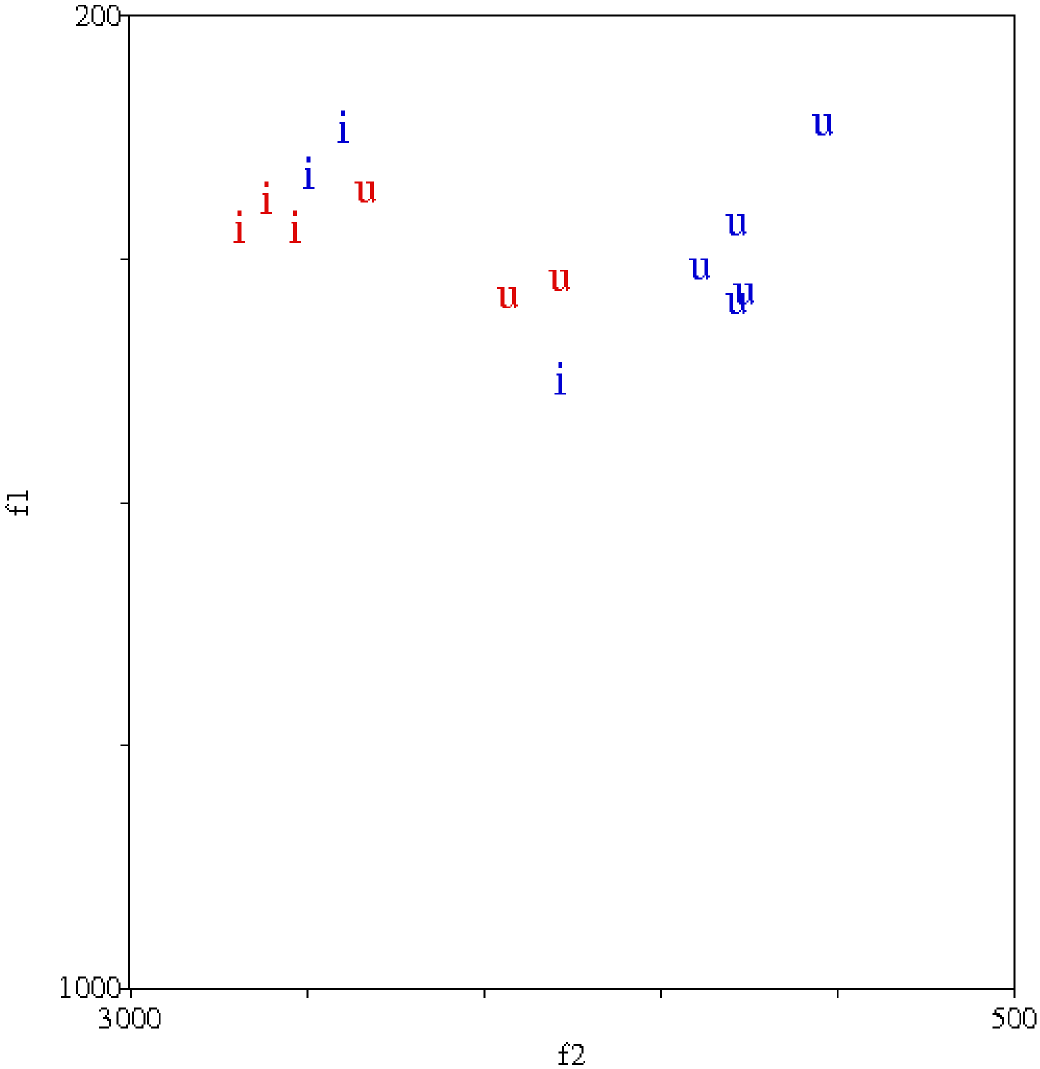

Figure 4. spoken (red) and sung (blue) high vowels of subject 1 (professional soprano)

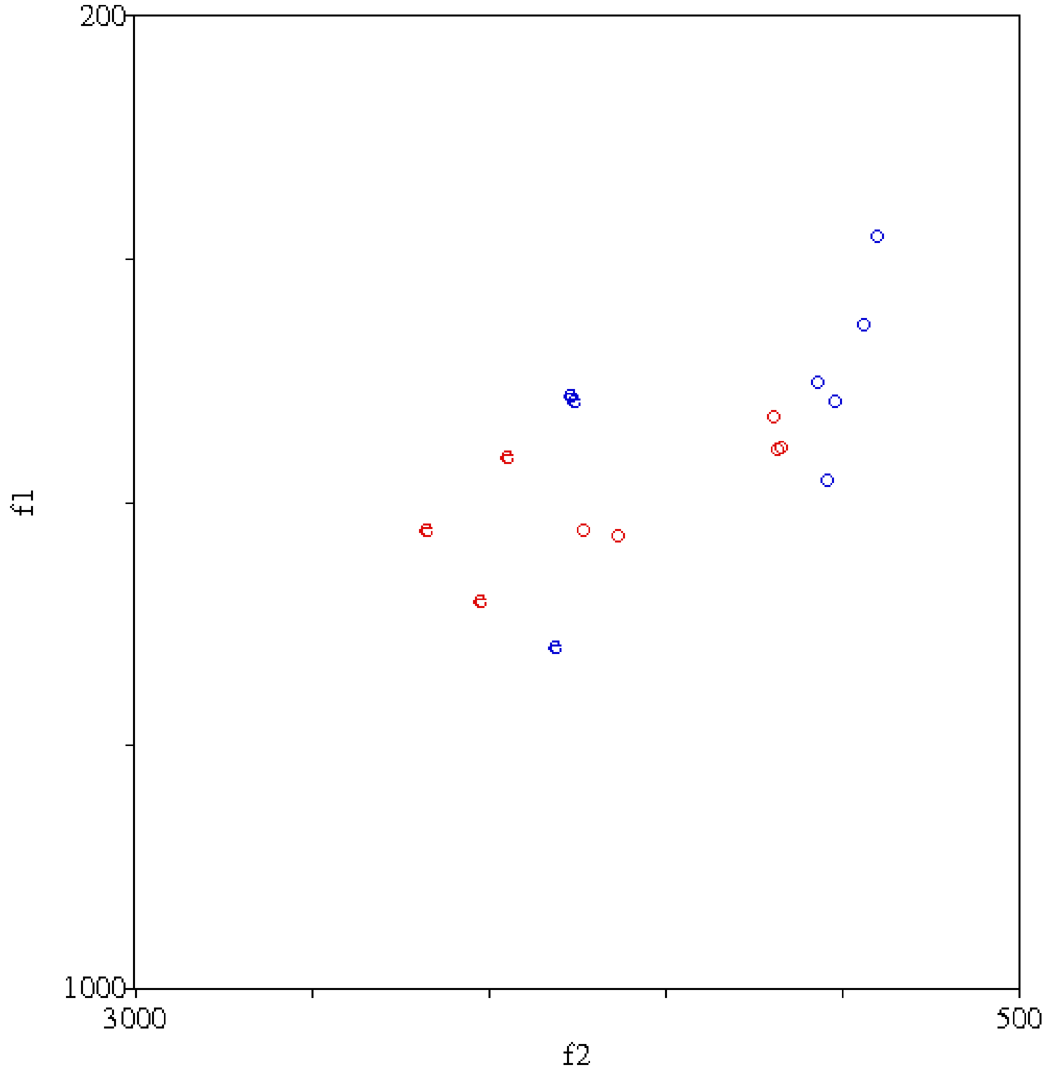

Figure 5. spoken (red) and sung (blue) mid vowels of subject 1 (professional soprano)

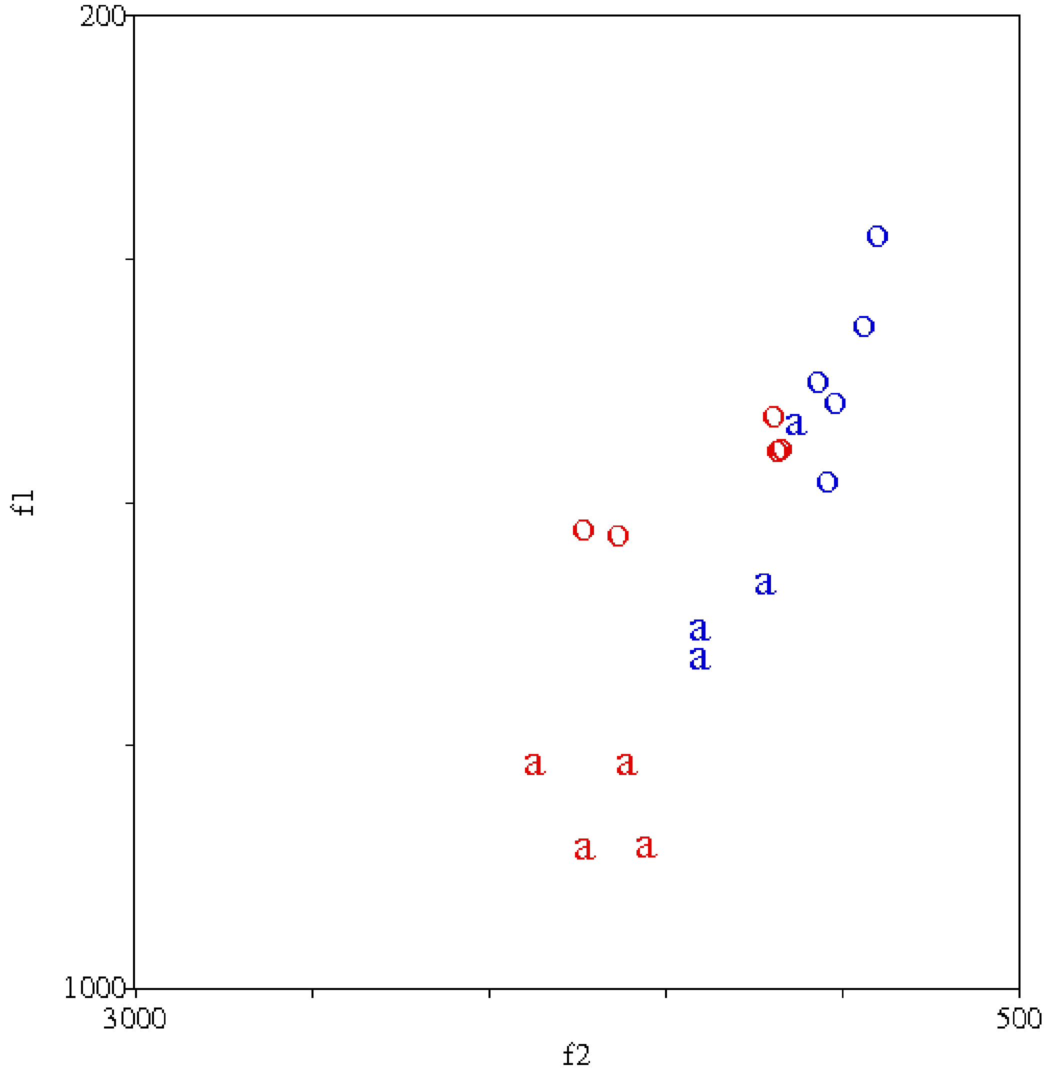

Figure 6. spoken (red) and sung (blue) low vowels of subject 1 (professional soprano)

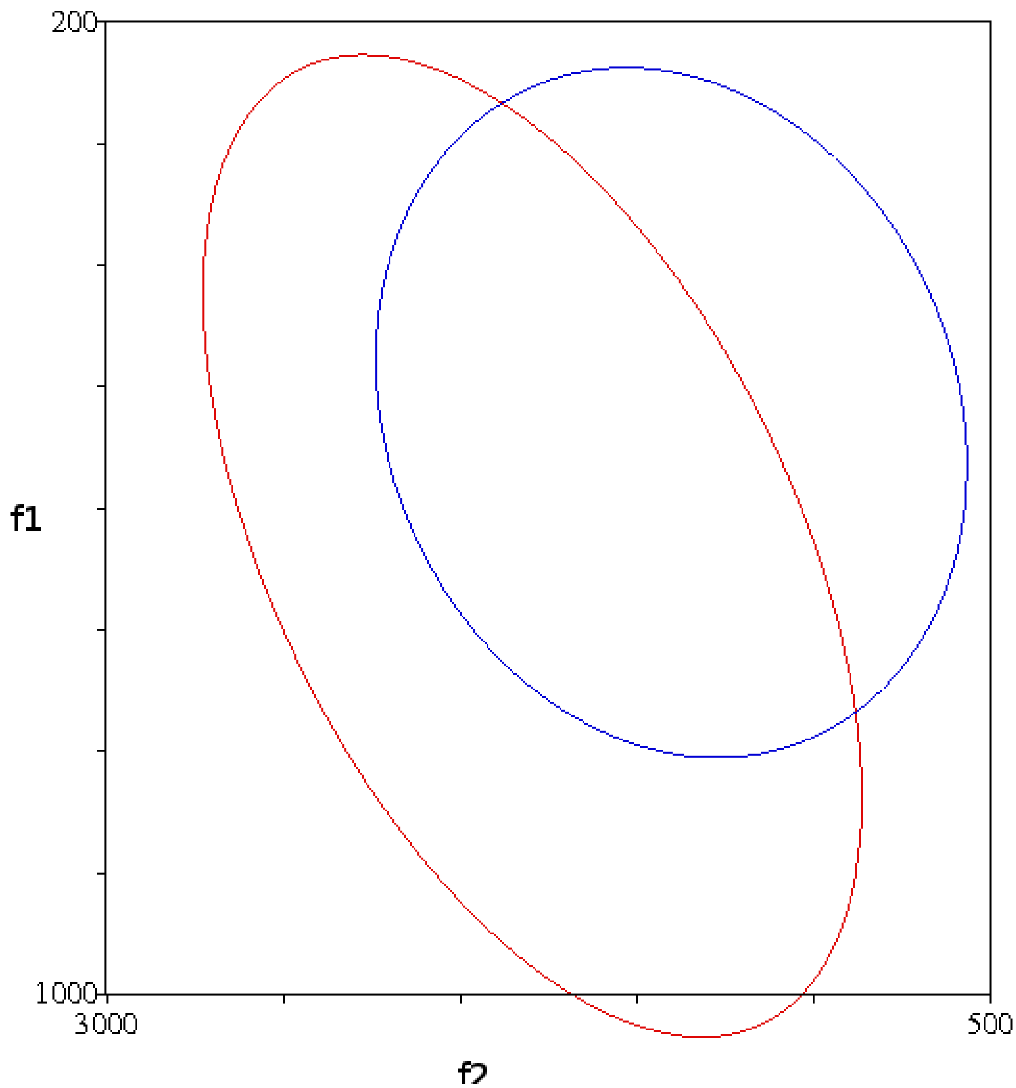

Figure 7. variance of all spoken (red) and sung (blue) vowels of subject 1 (professional soprano). Ellipses represent two standard deviations from the mean.

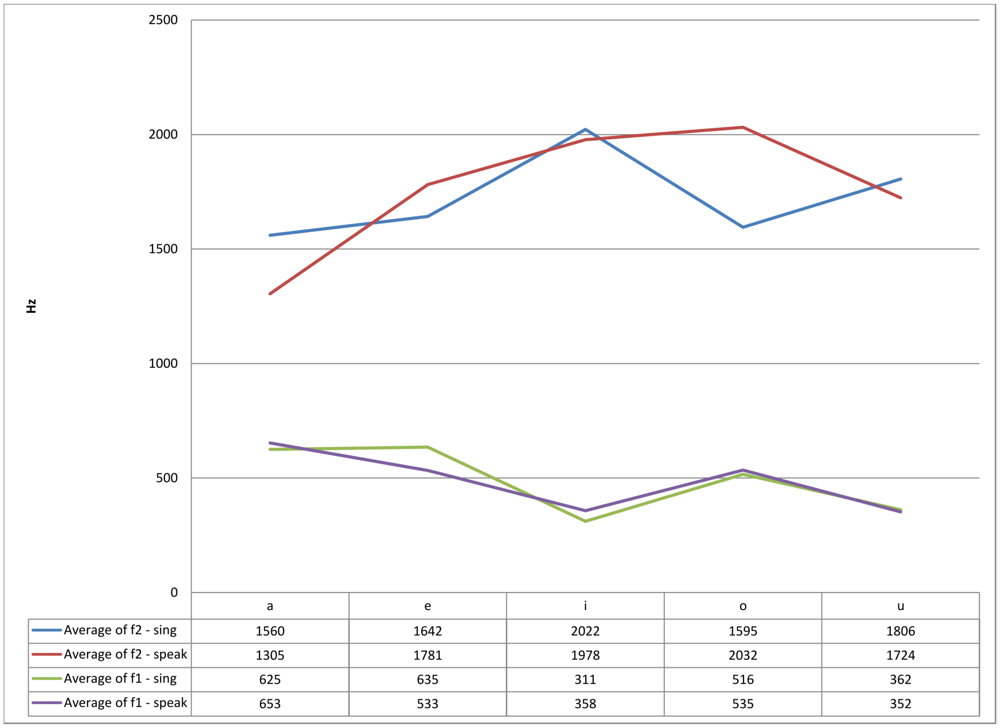

Figure 8. formant values during speaking and singing, females

vowel	f_1	f_2
æ	631	1502
ʊ	598	1806
a	625	1560
e	635	1642
i	311	2022
o	516	1595
u	362	1806
	508 (215)	1702 (756)

Table 5: first and second formant values during singing for males

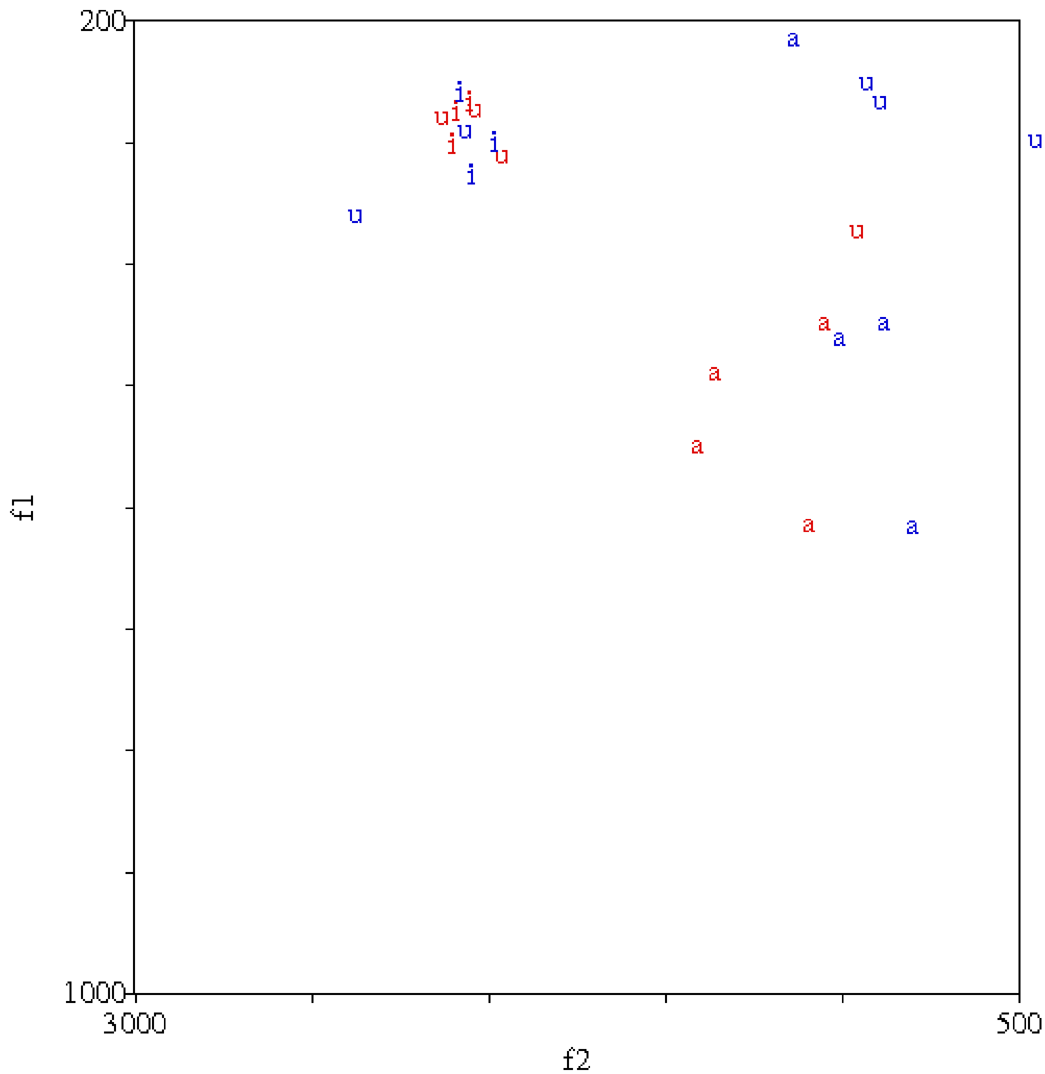

Figure 9. spoken (red) and sung (blue) point vowels of subject 4 (amateur baritone)

an increase in f_2 frequency during singing (figure 10)

Discussion

1. Overall, formants will be lower in sung vowels, as compared to spoken vowels.
2. As vowels migrate, the variance in formant values will be lower in singing than in speech.
3. high vowels will show the greatest f_1 changes, while front vowels will show the greatest f_2 changes.
4. if larynx control requires training, then amateur singers should have vowel spaces which are more similar to their speaking voices than are those of trained singers.

Clearly, vowels are articulated differently in singing than they are in speech. One of these differences is larynx height, and several of the observations made here are consistent with the hypothesis that singers maintain a lowered larynx during singing. The first formant frequency of non-high vowels was lowered, and the second formant frequency of nearly all vowels, especially front vowels, was also lowered.

Some observations were inconsistent with this hypothesis. Curiously, among both males and females, /i/ showed almost no change between speaking and singing. This is puzzling, given that in speaking, it is typically produced with a raised larynx, meaning that it should be most affected by an overall lowering of the larynx. This could be due to inconsistent application of larynx height by the singers

One singer (singer 2) showed vowel migrations which could be described as the opposite of that expected; some of his formant frequencies became higher during singing. This singer reported studying a method which advises the singer to change his articulation of vowels dependent on where the pitch of the note on which they are sung falls within the individual singer's range in order to increase their resonance and improve timbre (Coffin, 1987). This could have led the singer to produce unordinary formant frequencies for some vowels, because, in effect, he was changing the identity of the vowels somewhat

too few 'untrained' singers have yet been analyzed to fully speak to the effects of training, so a continuation of this work with more singers of various backgrounds may be informative. Even the 'amateur' singer recorded for this study had significant singing experience, so it may be interesting to examine whether some of these acoustic effects may be 'naturally occuring', or require formal training in order to acquire.

Conclusion

The acoustic vowel space used in singing shifts from that which occurs in speech, due to demands of musical aesthetics and articulatory constraints. This study directly compared the acoustic properties of vowels in the sung and spoken modes. Some of the acoustic effects are predictable given knowledge about singers, and what kinds of articulatory changes singers are likely to produce.

References

Bloothooft, G., & Plomp, R. (1986). Spectral analysis of sung vowels. iii. characteristics of singers and modes of singing. *The Journal of the Acoustical Society of America*, *79*(3),

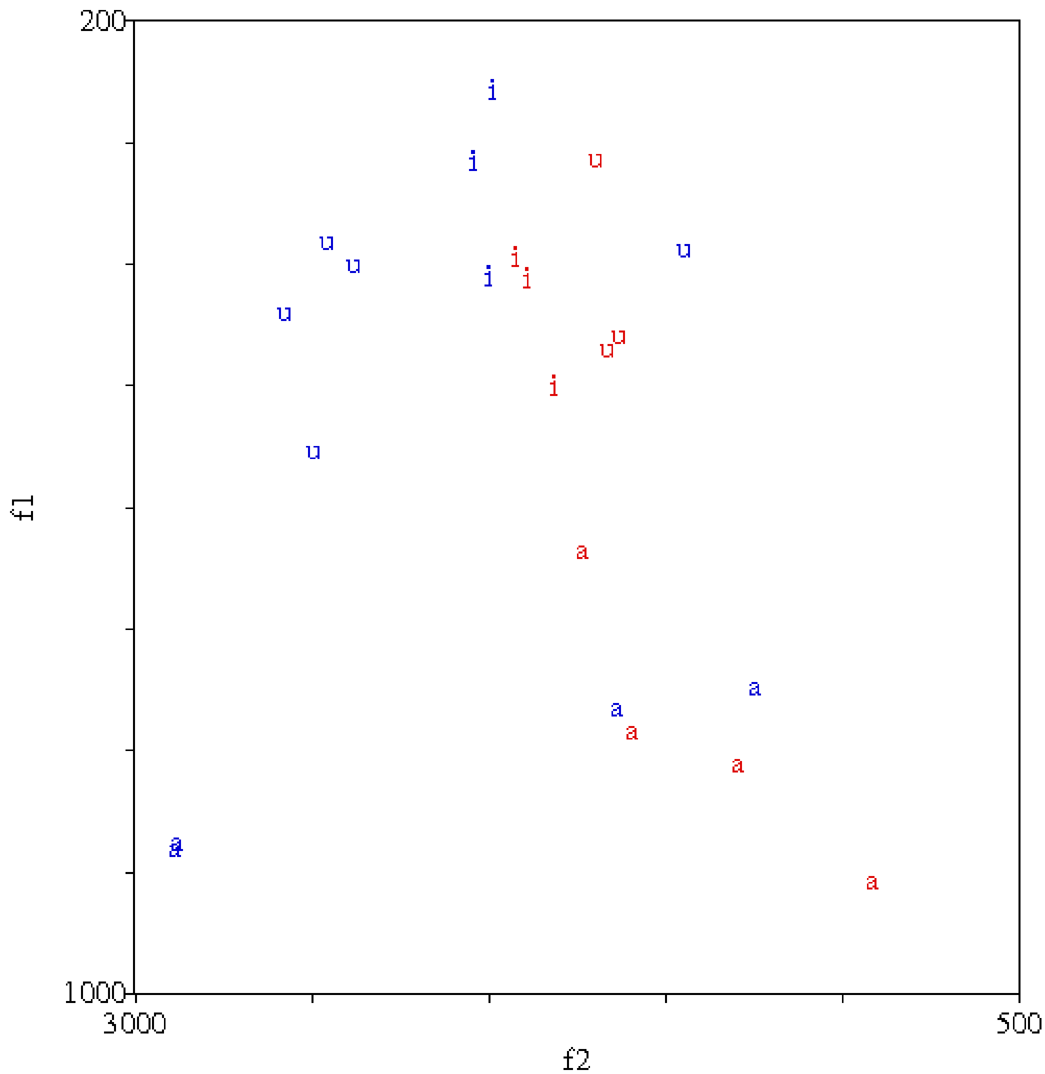

Figure 10. spoken (red) and sung (blue) point vowels of subject 2 (professional tenor)

852–864. Available from http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=JASMAN000079000003000852000001&idtype=cvips&gifs=yes

Boersma, P., & Weenink, D. (2009, April). *Praat: doing phonetics by computer (version 5.1.04).* Available from http://www.praat.org/

Coffin, B. (1987). *Sounds of singing* (second ed.). Metuchen, N.J.: The Scarecrow Press.

Howard, D., & Collingsworth, J. (1992). Voice source and acoustic measures in singing. *Acoustics Bulletin*, *17*(4), 5–12.

Sundberg, J. (1987). *The science of the singing voice*. DeKalb, IL: Northern Illinois University Press.

Sundberg, J., Fahlstedt, E., & Morell, A. (2005). Effects on the glottal voice source of vocal loudness variation in untrained female and male voices. *The Journal of the Acoustical Society of America*, *117*(2), 879–885. Available from http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=JASMAN000117000002000879000001&idtype=cvips&gifs=yes

Trainor, L. (1997, September). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development*, *20*(3), 383–396. Available from http://dx.doi.org/10.1016/S0163-6383(97)90009-6

Wray, A., Evans, B., Coupland, N., & Bishop, H. (2003). Singing in welsh, becoming welsh: turfing a grass roots identity. *Language Awareness*, *12*(1), 49–71.