Performing Lieder: Expert Perspectives and Comparison of Vibrato and Singer's Formant With Opera Singers

***Lynette Johnson-Read, *Anthony Chmiel, *Emery Schubert, and †Joe Wolfe,** **†Sydney, New South Wales, Australia*

Summary: This article reports three studies about performance of lieder, and in particular in comparison with opera performance. In study 1, 21 participants with experience in music performance and teaching completed a survey concerning various characteristics of lieder performance. The results showed that there was consensus between the literature and the assessment of an expert panel—that a "natural" and "unoperatic" vibrato was favored, and that diction, text, and variation of tone are all important aspects of lieder performance. Two acoustic analyses were conducted to investigate genre-specific differences of the singer's formant and vibrato parameters. The first analysis (study 2) used 18 single quasi-unaccompanied notes from commercial recordings of two lieder, and, for comparison, 20 single unaccompanied notes from an opera. Vibrato rate was statistically identical between the two genres at ∼6.4 Hz; however, lieder featured a longer delay in vibrato onset. Vibrato extent was smaller for lieder (∼112 cents) compared with opera (∼138 cents). The singer's formant, which is generally associated with opera, was at times observed in the lieder recordings; however, this was at an overall significantly weaker intensity than in the opera recordings. The results were replicated in study 3, where recordings using only singers who performed in both lied and opera were analyzed. This direct comparison used 45 lieder notes and 55 opera notes and also investigated three different methods of analyzing the singer's formant. A number of consistencies and inconsistencies were identified between acoustic parameters reported in studies 2 and 3, and the beliefs of singing teachers and scholars in the literature and study 1.

Key Words: Vibrato rate–Vibrato extent–Vibrato onset–Singer's formant–Singing–Lieder–Opera–Expert panel.

INTRODUCTION

Teaching and learning to sing in different musical styles involve a capacity to execute characteristics that nuance the singer's output so that the requirements of that style can be met. For example, Sataloff[1] states that vibrato can be produced in a different fashion for each genre of the singing voice, and that its presence is one factor which determines the character of each genre. Our broad interest was to investigate whether a set of principles for all genres and styles of singing could be available to reflect what performers do and to work toward well-defined guidelines to aid teachers and students. We were also interested in gathering perspectives from singing teachers and performance and comparing those, where possible, with acoustic measurement.

This article specifically looks at lieder, the 19th century German form of song that is usually accompanied by piano. Through a variety of mediums, certain vocal techniques are identified as being important for the performance of lieder. Vibrato and the singer's formant (henceforth, SF–a prominent peak in the spectral envelope at around 3 kHz)[2] are parameters that are specifically studied in voice acoustics. We also explored the use of a qualitatively "pure tone" in lieder.

Accepted for publication October 28, 2014.

From the *Empirical Musicology Group, School of the Arts and Media, The University of New South Wales, Sydney, New South Wales, Australia; and the †Acoustics Group, School of Physics, The University of New South Wales, Sydney, New South Wales, Australia.

Address correspondence and reprint requests to Emery Schubert, Empirical Musicology Group, School of the Arts and Media, University of New South Wales, Sydney, New South Wales 2052, Australia. E-mail: e.schubert@unsw.edu.au

Journal of Voice, Vol. 29, No. 5, pp. 645.e15-645.e32
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2014.10.020

Other genres of vocal performance have received more attention in recent studies. Research in the areas of opera and music theater are more plentiful, examining text intelligibility,[3] vibrato,[4,5] formant control,[6] and other vocal characteristics.[7,8] Similar topics are also examined under the generic label of classical singing.[9–20] Prame's studies[21,22] use F. Schubert's *Ave Maria* (an Art song, though not strictly a lied) to examine vibrato because it has many sustained notes throughout the piece. Sustained notes provide greater scope for vibrato effects, as will be discussed below.

In *Singing in style: a guide to vocal performance practices*, [23] Elliott discusses elements concerned with the singing of lieder. On vibrato and portamento, she writes "Vibrato can be used as an expressive device in lieder and is sometimes indicated with a wavy line or a variety of accent markings. A wide, continuous vibrato, however, is still inappropriate. Portamento, on the other hand, can be used liberally in many situations for both small and large gestures. It should be employed for expressive and dramatic purposes" (p. 192). The performance of lied is also discussed in *German lieder in the nineteenth century*[24] in which Spillman refers to diction, vowel formation, consonants, elisions, vocal "line," and the pianist. He notes that "[t]he singer must have a good ear for phonetic sounds and the technique and willingness to reproduce them faithfully … [as they are] frequently asked to do things that go against basic *bel canto* training" (p. 318). Miller's publication *Singing Schumann: an interpretive guide for performers*[25] lists examples of how lied is sung poorly. Some inappropriate techniques are identified: "scooping into important words … starting the vocal tone straight 'and then letting it wiggle' with vibrato … removing vibrancy on notes of

short duration, changing the dynamic intensity of each note in a phrase, using exaggerated 'vocal colouration' and 'word painting' to the detriment of vocal timbre" (p. 16). Although these texts state how lieder should or should not be sung, the ideas are mentioned briefly and no specific explanations or demonstrations are provided explicating how singers can use the techniques in practical situations. Nor do they apply specific principles of vocal acoustics, which can be valuable in documenting how lieder has, and can be, performed. No sources were found that made explicit statements about the use of SF in lieder, and reference to vibrato is mostly qualitative even though vibrato is quantitatively well defined. However, such an approach to understanding lieder was considered important in communicating the essence of lieder performance.

We therefore examined some of these issues by conducting three studies: one survey and two acoustic analyses of sound recordings. The general aims were to investigate views on lieder singing through a sample of experienced and expert musicians (study 1) and then to follow this up by examining selected acoustic parameters of lieder sound recordings in comparison with opera (study 2). The acoustic analysis was then refined to an examination of singers who have made recordings of lieder, but also in the comparison genre of opera (study 3).

STUDY 1: SURVEY OF MUSICIAN'S BELIEFS ABOUT LIEDER PERFORMANCE

The aim of study 1 was to investigate the perspectives among experienced teachers and performers of lieder, with regard to performance of the genre.

Methods

Participants. The recruitment of participants involved sending e-mails to a range of music organizations within Australia, including eisteddfods, examination boards, universities, opera and other vocal companies, lieder and art song societies, and various contacts. Around 80 people were contacted, of whom 21 completed the survey. Ethics approval was granted for the study.

From the recruited group, only three participants had prior contact with the authors. All but two of the participants had vocal training. One of the participants had a postgraduate education in music, and another was a professional accompanist specializing in lieder and vocal performance. Ten participants had completed a tertiary degree in music, and nine participants described themselves as singing teachers. Several participants had continued their vocal and instrumental training or performance careers overseas.

Stimuli and procedures. An online survey was prepared using *SurveyMonkey.*[26] The online survey consisted of 15 open-ended questions. These were as follows:

1. What is your understanding of "lieder"?
2. How would you describe a good lieder technique?
3. What should a singer do to produce good lieder? [Please provide as much detail as possible. Answer the question in every way that you can.]
4. What should a lieder singer NOT do? [Again, please provide as much detail as you can.]
5. [If not already mentioned] How would you say vibrato should be used?
6. Do you think there is the need for the voice to have a pure tone quality?
7. With respect to the previous question, what is your understanding of a pure tone quality?
8. Do you have a recording of a good lieder performance? If so, please provide the details of the recordings and CDs (up to the four best). Please provide as much information as possible (including catalogue number) so that I can locate the recording.
9. Select one of the recordings from the above question. Why is the performance good?
10. Is there anything you would change about the performance mentioned above?
11. Do you have a recording of lieder that you view as containing a poor performance? If so, please provide the details of the recording and CD. Please provide as much information as possible (including catalogue number) so that I can locate the recording.
12. If you have mentioned a specific recording in the above question, what do you think makes this poorly sung?
13. What would you recommend that this singer change? How would you explain this to the singer?
14. Please provide information as to your musical background. Have you undertaken singing or music training, what experience do you have in singing and in the field of lieder?

Results

The comments from the survey were content analyzed to pinpoint the techniques the participants of the survey identified as necessary for lieder performance, and whether there was a general consensus throughout. Each comment was put into a category. These categories included some that examined the form of the answer, such as metaphorical, adjectival, or a technical explanation. Other categories described what the comment was in reference to, including items from Wapnick and Ekholm's[27] criteria. Specifically, these categories were vibrato, diction (being the pronunciation of vowels and consonants), other translation issues (such as knowing the language and exposing subtleties, as well as literal translations of the text), text interpretation, references made to "words," comparison with other genres, and/or piano accompaniment.

The following summarizes the content analysis results question by question. Unless otherwise stated, all of the survey participants answered each question.

Question 1: "What is your understanding of 'lieder'?" The first question was solely present to separate any respondents without proper musical knowledge of lieder from other participants. All but four of the participants answered the question by

referring to the word "German," and only two did not mention the words "song," "Art Song," or "melody." These participants described lieder in a more conceptual style. The first wrote that lieder is "A partnership between poem, accompaniment, and vocal line," and the second participant wrote that lieder is "Generally Poetry in German set to music. However sometimes but rarely, poetry from other countries in other languages, set to music, is accepted in this genre." All of the responses provided sufficient information to confirm an understanding lieder.

Question 2: "How would you describe a good lieder technique?" Knowledge and importance of text were emphasized by eight participants. Good, clear diction, a centered sound, with variations of tone color, good breath control and support, and vocal flexibility were other typical comments. Operatic style was reported by three participants as a comparison with lieder, and this was often associated with an overdramatic vocal quality if applied to lieder. Technical responses occurred in all but four responses, and the possession of a good general vocal technique was mentioned throughout, such as "good breath support," "a beautiful tone," "[vocal] flexibility and control." In comparison, four participants used metaphors within their responses, such as "the sound has to be essentially sitting very high in the mask." Some examples fell into multiple coding categories and were coded once for each occurrence, eg, "their voice maintained an 'audible line' [metaphor], well balanced in tone throughout the range [technical], displaying sensitivity and color [adjectival], without an overriding vibrato [technical]." Comments concerning the importance of accurately conveying the meaning of the text through the voice were mentioned by 16 participants and consideration of the piano appeared in five responses.

Question 3: "What should a singer do to produce good lieder?" aimed to generate responses describing how the ideal techniques could be achieved by singers. Comments similar to the previous question were found here, but the responses were explained in further detail and with a more personal focus. There was a greater use of metaphorical (7) and adjectival (8) responses, and specific teaching techniques were mentioned. The following explanation is an example of one with a technical function: "vocal technique to produce a clear, pure, controlled sound." Another participant wrote that "lieder requires the singer to have great breath control, limited vibrato, even tonal quality throughout the range (no 'breaks' or 'holes'), linguistic proficiency, poetic interpretation, and a power over their audience—the ability to captivate the listener without the grandeur of an operatic scene."

Question 4: "What should a lieder singer not do?" The participants wrote that one should not oversing (seven respondents) or sing in the styles of music theater, opera, or other genres (9). Participants also noted that the language should be studied well and performed convincingly (4), that the pianist and singer should work as a team, being careful not to overshadow the piano (4), and that the emotion and story line that is being sung about should be carefully interpreted (8). Again, the text is emphasized: "A lieder singer should never perform a piece without intimate knowledge of their text - the poetry is the reason they are there."

Questions 5: "How would you say that vibrato should be used?" The general consensus (11 participants) was that a natural use of vibrato was preferred. A moderate use of vibrato was mentioned by seven participants although it was not clear what the difference was, if any, between a moderate and natural vibrato. Eight participants specifically discouraged the maximal use of vibrato in lieder, stating that this was an undesired operatic sound. Respondents advocated the varied use of vibrato to enhance emotion (4), and one mentioned that the reason lieder did not need an exaggerated use of vibrato was because the singer did not have to make their voice heard over an orchestra or in a large space; that the intimacy of lieder prescribed a more subtle use of vibrato. An additional possible acoustic implication of this belief is that lieder singers need not use SF to the same extent, or not at all, compared with the opera singer.[2,27] One participant summarized the responses with the comment that "[a]ll mature voices have vibrato which should be controlled and used to effect when appropriate."

Questions 6 and 7 asked the participants what their understanding of a pure tone quality was. A typical response was "a focused sound, little perceptible vibrato, ie, a narrow and even vibrato that gives vibrancy rather than intensity, a sound with a ringing quality to get a 'choir boy/early music' sound without airyness." One participant used more adjectival descriptions, writing that "[t]he *range* of vibrato and tone qualities is wider in lieder, but still a pure or 'white' quality should only be one of a vast range of colors and should be used very sparingly as it is very close to simply being poor technique." The responses to these questions suggest that purity of tone is a quality that is largely subsumed by the issue of vibrato. In combination with Question 5 and current literature, it also suggests consistent views on whether lieder should use wider or narrower vibrato extent than other vocal forms, such as opera. As a comparison, consider the suggestion by Elliott discussed in the introduction. A precise use of the term vibrato extent was not reported in the survey responses.

Question 8: "Do you have a recording of a good lieder performance?" The performers who were highlighted in good recordings of lieder varied a great deal. Included in these choices were non-classical singers, such as Barbara Streisand. The performers who were mentioned more than once are shown in Table 1. Of the pieces mentioned, the most frequently

TABLE 1.
Performers Mentioned More Than Once in Question 8, Pertaining to Good Performances of Lieder

Performer	Times Mentioned
Dietrich Fischer-Dieskau	10
Jessye Norman	6
Elisabeth Schwarzkopf	4
Bryn Terfel	4
Elly Ameling	3
Peter Schreier	2
Felicity Lott	2
Janet Baker	2

occurring composers were Schubert (12 participants), Schumann and Strauss (five participants each), and Brahms (three participants).

Question 9: "Select one of the recordings from the above question. Why is the performance good?" Eleven participants answered this question. A beautiful, flexible, and varied tone from the singer was reported as an important aspect by seven participants. Good diction and language (4) and an interesting and understood interpretation of the text (4) were again highlighted. An awareness of style and the composer's intentions was mentioned specifically by three participants.

Question 10: "Is there anything you would change about the performance mentioned above?" No participants provided useful information in this section, and in fact 12 participants stated that they could not improve the performance, and six participants did not respond at all. Participants seemed to be reluctant to critique the performers, with one participant noting that they "wouldn't dare" give criticism. Other respondents concluded that imperfection is part of what makes performance "real" and that maturity will "add depth to tone and artistry" of the performance.

Questions 11 to 12: Many of the participants did not answer the next three questions (Questions 11, 12, and 13). When asked to name recordings of lieder, they viewed as being "poor," and to identify what makes it poorly sung, only eight participants responded. Participants tended to mention performers rather than specific recordings for this question and did not go further into detail (3). Four participants responded to Question 12. The disliked features of the poor recordings included a "tremulous" or "wobbly" sound, not being able to "let their egos take a second place," being "indulgent in the beauty of their own voice," creating a "boring interpretation," use of inappropriate singing styles, and a poor use of the German language.

Question 13: "What would you recommend that this singer change? How would you explain this to the singer?" This question focused on pedagogic implications that could be made to enhance lieder singing. Only six participants answered this question, which again reflected the participant's reluctance to comment negatively or constructively on the performances of well-known singers. Alternatively, the poor response rate may also be interpreted as a sign of a lack on knowledge on the participant's part. Furthermore, placing critique on a professional singer may require that the survey respondent hold an equal or more advanced knowledge of singing to that singer. A more thorough and imaginative text interpretation (3) and better diction (2) were the only recurring comments for this question. Some of the respondents felt they had answered this question exhaustively in Question 12, and so referred to that question rather than restating their opinions.

Throughout the survey, some aspects were mentioned repeatedly. Commonly stated were words such as "text" (67 mentions across the entire data set) and "poet" or "poetry" (25), "diction," and other related words such as "language," "German," and "pronunciation" (50), "interpretation" (20). Phrases about the partnership between voice and piano (36) and the comparison to opera (24) were mentioned regularly.

Discussion

The responses from the survey highlighted some interesting opinions about the performance of lieder. The common words and phrases used indicate that many of the participants placed importance on similar aspects. These trends were also reflected in three texts on singing by Elliot, an edited volume by Hallmark and a guide to performance by Miller.[23–25]

The analysis indicated that a "natural" use of vibrato was preferable. Although a specifically "natural" use of vibrato for lieder is not explicitly stated in the texts, the results of the survey indicate that a supposedly wide or "operatic" use was considered inappropriate. However, as with several comments made in the self-reported comments and in the literature, a quantifiable definition of "natural" vibrato (eg, in terms of vibrato rate, onset, or extent, discussed in the following) was not forthcoming.

Why are text and diction mentioned so frequently? Can it not be assumed that all vocal performances need to have a high level of textual comprehension and diction? Possible interpretations could be that in lieder, the voice is more exposed because of the lower amount of accompaniment in terms of both instrumentation and potential volume of sound compared with genres such as music theater and opera, therefore creating a need for perfect diction and a more interesting interpretation. At its inception, lieder aimed to converge German poetry and music. The dual focus on text as well as the use of the German language was innovative at the time. Perhaps, this remains true in the performances of today. Furthermore, given that lieder performers do not have the advantage of props and staging found in much opera and music theater, clarity of text may receive greater weighting.

Although agreement was identified in the survey, some of which was also consistent with the literature, it also became apparent that a number of terms were not clearly defined. These include "good tone," "subtle vibrato," and "nonoperatic projection." These findings raised additional questions that were further investigated from an acoustic perspective. Given the prevalence of comparisons reported between lieder and opera, such as the former using less vibrato, and our interest in comparing qualitatively reported characteristics with acoustically measured variables, we conducted two additional studies. These studies took a more objective approach, measuring acoustic properties of singers and comparing lieder and opera singing. We limited our investigation to two acoustic parameters of the voice: vibrato and SF.

STUDY 2: A COMPARISON OF VIBRATO AND SF BETWEEN LIEDER AND OPERA

The aim of study 2 was to determine whether an acoustic analysis supported the more consistent views raised in the literature and the survey. This was done through a comparison of acoustic characteristics of lieder with opera. Specifically, comparisons were made between SF and three aspects of vibrato: vibrato onset time (delay from onset of a stable note/phoneme to onset of vibrato), pitch vibrato rate (number of regular pulsations in pitch per second) and vibrato extent (half the maximum to

minimum fundamental frequency fluctuation in vibrato; for more information see Sundberg[19]). Based on the literature review and qualitative data obtained in study 1, it was hypothesized that, when compared with opera.

H1: SF is lower in lieder excerpts (eg, because of the lack of competition from the piano accompaniment, compared with an orchestra).

H2: vibrato onset is the same between genres (an interpretation based on Miller's assertion that the tone should not be started "straight").

H3: vibrato extent is smaller (not as wide) in lieder excerpts.

In addition, we investigated the vibrato rate between opera and lied. However, we were not able to find any literature that explicitly predicted if there should be a difference between the two genres, and no survey responses specifically mentioned this characteristic.

Method

Design. Two designs could be used to investigate the hypotheses. One is to recruit singers and ask them to perform in certain ways, and an other is to use existing recordings of respected commercially recorded performances. We decided to limit our investigation to existing commercial sound recordings of well-known performers to avoid the problems of the former, being that either students would have to be employed, possibly reducing the control we had over different performance conditions (lied vs opera), or the prohibitive cost and time of employing highly respected performers. Furthermore, the difficulties of instructing any singer to perform a lied and an opera extract may reveal the design problem of demand characteristics, where the researcher cannot be sure that the singing participant is not adjusting their performance as an artifact of the experimental condition, rather than as an effect of the different genres. Although recruiting singers may be a tempting and interesting avenue for further study, we decided to use commercial recordings instead. However, even conducting acoustic analysis using commercial recordings presented several design challenges. Among the most important was the acoustic separation of vocal information from piano/orchestral accompaniment, which is generally hard to do with commercial sound recordings. With this in mind, we opted to choose sound recordings from lieder and opera where at least one note could be identified that had minimal interference from the audio signal of the accompaniment and we performed acoustic analyses on these.

Stimuli. Altogether, 38 single note excerpts were examined. Specific choices, such as the voice type and composer, were guided by the results of the survey (Table 1). Baritones were chosen because a high number of performers with that voice type were identified throughout the survey, and the literature associates SF with lower voice types.[28] A single unaccompanied note from each of Franz Schubert's songs *Wasserflut* and *Der Erlkönig* were used as the two lieder excerpts (Figure 1A and B). Two single notes from the extended recitative before the Count's aria "Vedrò mentr'io sospiro" in W. A. Mozart's *The Marriage of Figaro* provided two excerpts for comparison from the operatic repertoire (Figure 1C and D). A single, sustained note was chosen from each excerpt that was sung on a similar vowel sound, occurred in the same place in the overall phrase, and was performed in a comparable register.[4] These notes were chosen because they were unaccompanied, thus allowing a clear analysis without interference from the accompaniment.[4] Using a portion of a recitative from a Mozart opera, instead of Wagner for example, provided a conservative analysis, in that the Mozart style possesses relatively more subtle differences in vocal styles compared with lieder than some later, grander operatic forms. This way, we could be confident of the veracity of any measurable differences, should any be identified. A list of the performers and dates of recordings used are shown in Table 2.

Procedures. For acoustic measurements, the sound recordings were analyzed at each note of interest using sound editing and analysis software *Audacity 1.3.3 Beta* (GNU General Public License software, http://audacity.sourceforge.net/).[29] Vibrato measurements were made with the spectrogram in *Audacity*, set to a sampling rate of 44100 Hz and a Fast Fourier Transform (FFT) window size of 2048 samples, using Hann windows. For analysis of SF, measurements were made using the frequency spectrum plotter in *Audacity*, set to a sampling rate of 44100 Hz and a fairly course window size of 128 samples (344 Hz) for the easy identification of formants, rather than the details of the partials.

Vibrato parameters were calculated by inspection of the spectrogram. Vibrato rate was measured by identifying each full vibrato cycle, which consisted of a consecutive peak and trough, and the total number of these cycles was counted and then divided by the duration of those cycles in seconds. This produced a vibrato rate in hertz. Vibrato extent was estimated by reading the peak and trough frequencies during a region of sustained vibrato. The peak frequency was divided by the trough frequency, halved, and then converted into cents (a unit related to pitch that divides an equal temperament semitone into a further 100 units), as is the convention for vibrato extent[22] (see Figure 2 for an example of the readings taken to make these calculations).

The relative level of the SF (henceforth LSF) was quantified by subtracting the sound level of the first formant (henceforth, formants are denoted as F_n, where n = formant number, and their levels are measured in units of dB) from the sound level of the SF (also in dB). Figure 3 illustrates how the calculation was read from the spectrum. This provided information about the relative strength of the SF, based on methods used by Rossing et al[30] and Omori et al[31] and removed the problem of using SF readings that were misleading because of overall sound recording level. We delve into more detail in the measurement of LSF in Study 3.

Results

Results of the acoustic measurements are summarized in Table 3. The overall strongest LSF was in the Tahu Rhodes A3 operatic note (3 dB higher than F_1; Figure 3), and the weakest LSF was found in Terfel's *Der Erlkönig* (29 dB below F_1). On average, the opera excerpts had a stronger LSF (6 dB below

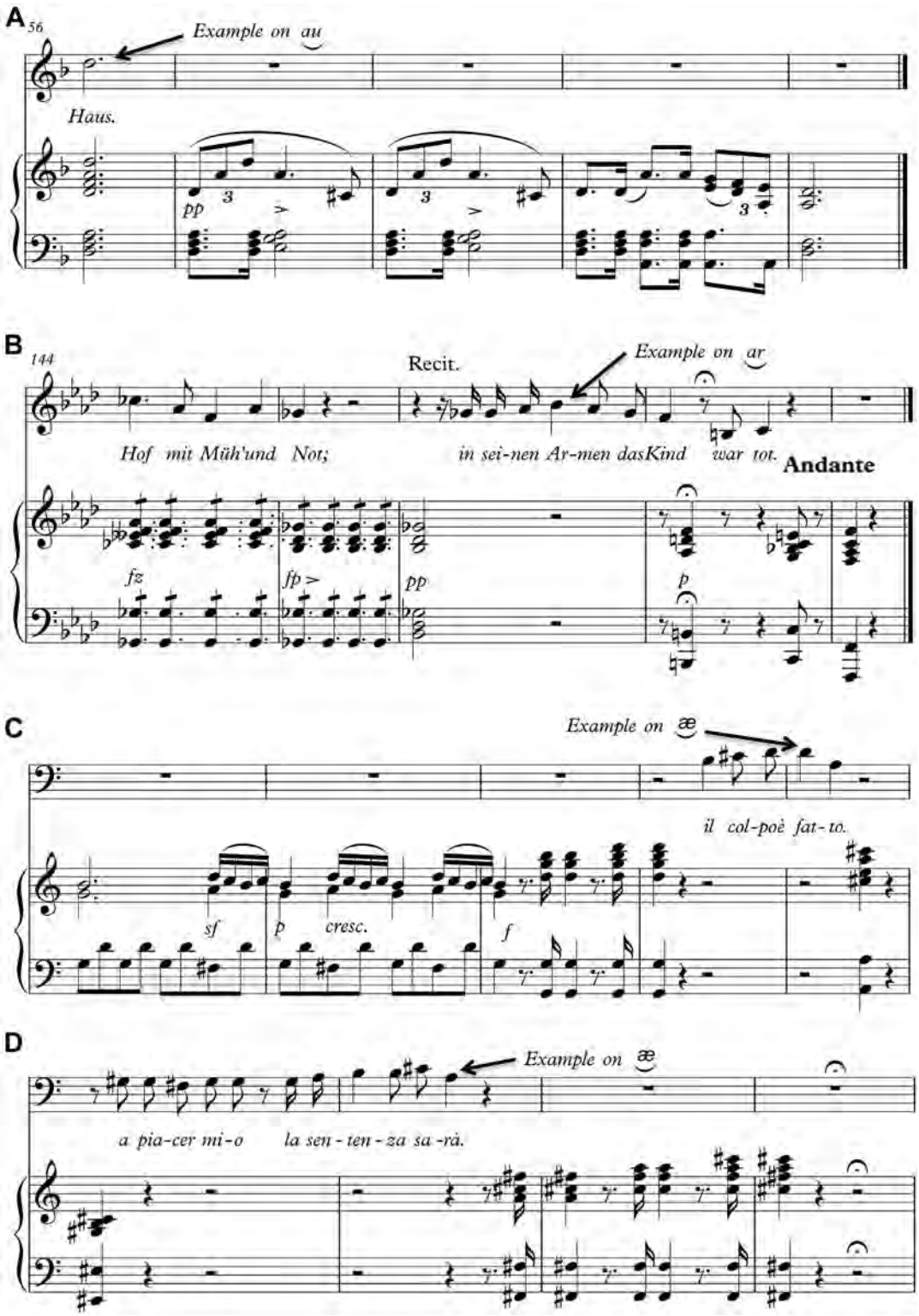

FIGURE 1. **A**. Extract from Schubert's *Wasserflut*, depicting one of the lieder examples used in study 2. **B**. Extract from Schubert's *Der Erlkönig*, depicting one of the lieder examples used in study 2. **C**. Extract from Mozart's *Le Nozze di Figaro*, depicting one of the opera examples used in study 2. **D**. Extract from Mozart's *Le Nozze di Figaro*, depicting one of the opera examples used in study 2.

F_1) than the lieder excerpts (12 dB below F_1). On average, vibrato extent was wider for opera excerpts (148 cents, about one and a half semitones) than for lieder excerpts (106 cents, about one semitone).

Mean vibrato onset time was much shorter for operatic excerpts compared with lieder excerpts (63 milliseconds and 276 milliseconds, respectively), but it should be noted that note durations for the *Wasserflut* excerpt were considerably

TABLE 2.
Performers and Year of Recording for the Excerpts Used for the Acoustic Analysis in Study 2

Wasserflut	Der Erlkönig	Le Nozze di Figaro (M1 and M2)
Dietrich Fischer-Dieskau (1988)	Dietrich Fischer-Dieskau (1988)	Dietrich Fischer-Dieskau (1968)
Dietrich Henschel (2006)	Bryn Terfel (2007)	Bryn Terfel (2006)
Hermann Prey (2008)	Hermann Prey (2008)	Teddy Tahu Rhodes (2003)
Gerhard Hüsch (1999)	Gerhard Hüsch (2001)	Thomas Allen (1982)
Matthias Goerne (2005)	Matthias Goerne (1997)	Thomas Hampson (1991)
Patrick Mason (1995)	Ernst Buscagne (2002)	Jonathan Lemalu (2005)
Jon Vickers (1992)	Marko Rothmüller (2008)	José Van Dam (2005)
Thomas Hampson (2005)	Wolfgang Holzmair (1993)	Bo Skovhus (2007)
Carl-Heinz Müller (1988)	—	Alfred Poell (1955)
Olaf Bär (2007)	—	William Stone (2009)

longer. *Der Erlkönig* had a more comparable duration with the Mozart excerpts (average duration 694 milliseconds), and the corresponding vibrato onset time was 135 milliseconds. On examining the proportional time of entry of the onset with respect to the duration, this gives a ratio of 135/694 = 0.195. For the opera excerpts, this ratio is 63/533 = 0.118 (Table 3), meaning that on average, lieder notes started vibrato proportionally later (with respect to note duration) than opera notes. Vibrato rate was similar across the two genres (6.16 and 6.20 Hz for lieder and opera notes respectively). LSF and vibrato extent, onset time, and rate were submitted to Mann-Whitney *U* tests. LSF, vibrato extent, and vibrato onset time were all significantly different between the two genres at $P = 0.01$, whereas vibrato rate was not (Table 4).

Discussion

Using acoustic analysis techniques, three hypotheses were tested in study 2. Two were supported, one rejected: lieder excerpts were performed with a weaker LSF (H1) and smaller vibrato extent (H3) than the opera excerpts. However, there was a difference in vibrato onset time, contrary to H2, with a longer delay in lieder than in opera. Additionally, no difference was observed in vibrato rate between the genres.

The results of these acoustic vibrato onset measurements indicate that, in comparison with opera, vibrato in lieder begins later within a note. In this study, Henschel's performance of the note of interest in *Wasserflut* had the most delayed vibrato onset of 1.85 seconds for a note of duration 2.85 seconds. The average delay of vibrato in lieder technique suggests that vibrato may be used as a coloration rather than as an immediate feature of vocal tone. This finding therefore questions the negative connotation in the use of delayed vibrato onset. As discussed in the introduction, a delay in vibrato onset has been argued to be indicative of a faulty technique, for example where Miller[25] states that a lieder singer should not let the vocal tone start straight and then let it "wiggle with vibrato" (p. 16). It must be noted, however, that Miller may have been referring to the avoidance of extreme use of vibrato *per se*, with which our data concur. The contribution of the present analysis allows for a more specific acoustic description of lieder performance. Furthermore, delay in vibrato onset may be an acoustic representation that describes the idea of pure tone reported in study 1.

The estimated vibrato extent showed that a narrower vibrato of around 100 cents was used in the lieder excerpts, compared with an average of 150 cents in opera. These values are consistent with those of Prame.[22] This is also consistent with reports made in both the survey and literature. Examination of the two sampled performers who had recordings in both genres,

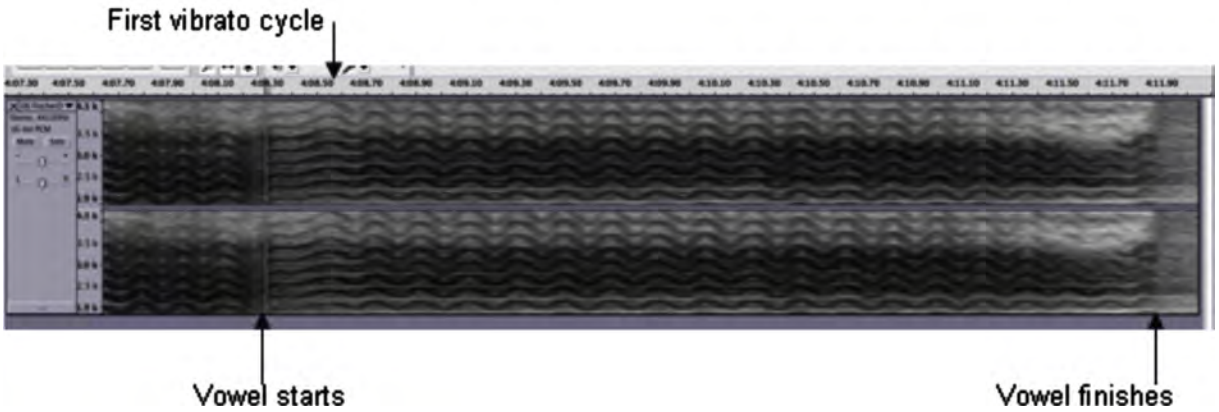

FIGURE 2. Spectrogram analysis from study 2 using *Audacity* to identify vibrato parameters.

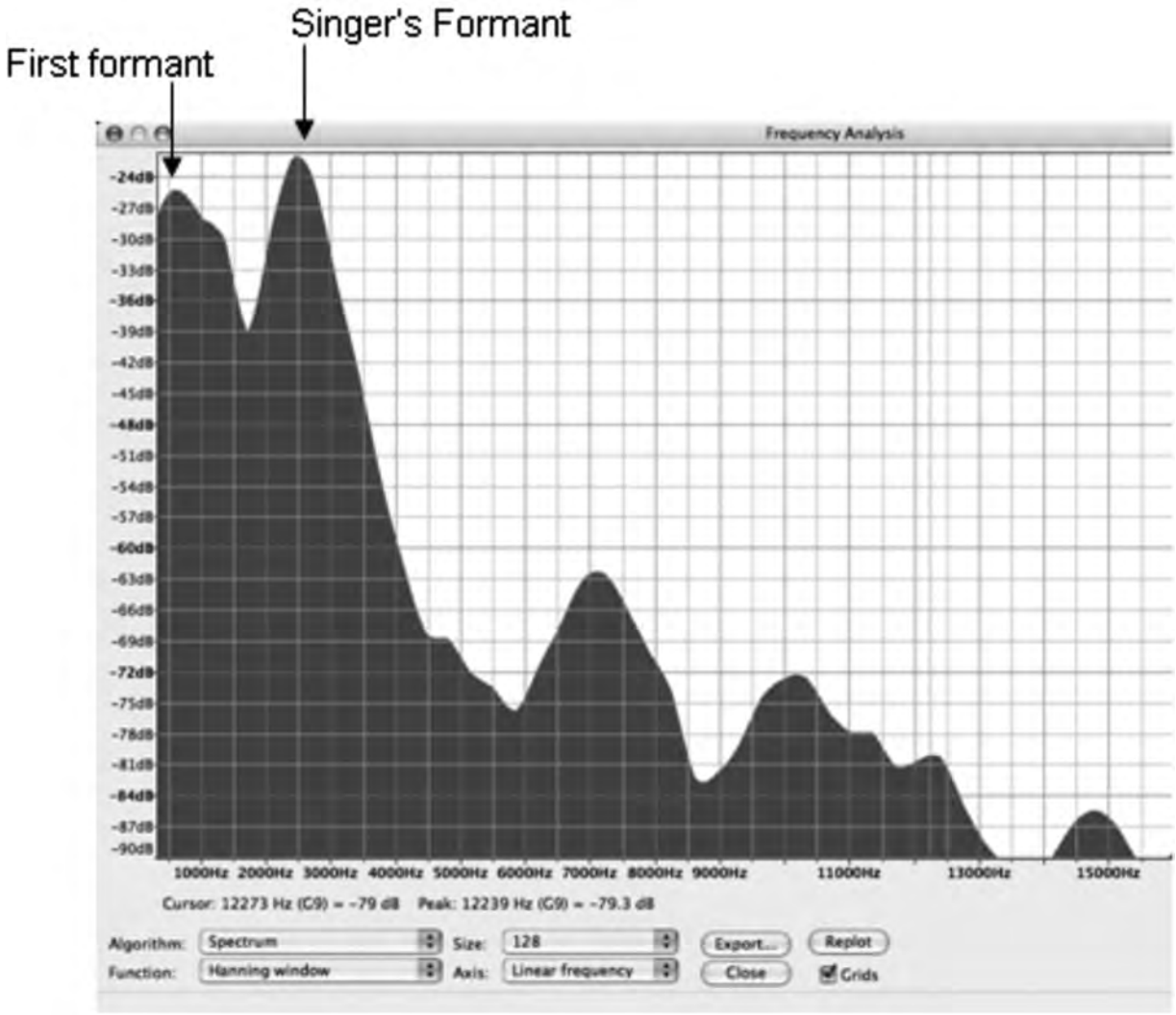

FIGURE 3. Spectral analysis in study 2 using *Audacity* to measure LSF.

Fischer-Dieskau and Terfel, demonstrated a distinction in vibrato onset delay across the two genres. For lieder excerpts, the delays were 100, 60, and 400 milliseconds, versus 50, 30, and 20 milliseconds for opera, whereas other parameters measured varied less systematically (Table 3).

SF was identified in both the opera and lieder recordings. Its use on a held note at the end of a phrase in *Wasserflut* is apparent in a majority of the recordings. However, the statistical analyses suggest that there is an acoustically significant difference between the two genres, with lieder singers being less reliant on the strength of that formant. Again comparing the performers who were sampled in both genres, Fischer-Dieskau and Terfel, these excerpts exhibited a stronger LSF in the opera excerpts (−7.8, −3.8, −7.4, and −8.3 dB) than in the lieder excerpts (−12.8, −5.9, and −29.4 dB). It is apparent from these data that professional, commercially recording singers have considerable flexibility in adapting to the different genres. The results are also consistent with Sundberg and Romedahl's[3] observation that a lack of SF increased text intelligibility. Furthermore, beyond not needing SF to project over the lighter accompaniment, a primary focus on text and diction in lieder compared with opera could be another explanation for why SF is lower.

STUDY 3: COMPARISON OF SINGERS WHO PERFORM IN BOTH GENRES

One of the criticisms of the design of study 2 is that not all singers performed in each genre. It can be argued that the results were confounded by variations across performer rather than across genre. Furthermore, no explicit attempt was made to control for the duration or pitch of notes because of the obvious difficulties in doing so, particularly with a small sample. To address these limitations, a third study examined only singers who had recordings available for both opera and lied. The number of excerpts used for analysis was also increased. Comparisons were made to further ascertain whether characteristics in SF and vibrato differed because of the genre performed.

The three hypotheses proposed in study 2 were also tested in this study. Additionally, this study included a fourth hypothesis that there would be no difference in mean vibrato rate between

TABLE 3.
One-Note Analyses of Singer's Formant Strength, and Vibrato Extent, Onset and Rate, for Lieder Versus Opera

Source	Performer	Year	SF Strength (dB)	Vibrato Extent (cents)	Note Duration (ms)	Vibrato Onset Delay (ms)	Vibrato Rate (Hz)
Der Erlkönig	Buscagne	2002	−3.1	59	600	130	6.38
	Fischer-Dieskau	1988	−12.8	165	700	60	6.25
	Goerne	1997	−8.9	113	800	150	5.88
	Holzmair	1993	−15	130	350	50	6.67
	Husch	2001	−18.8	142	870	130	6.76
	Prey	2008	−8.6	75	330	20	6.45
	Rothmuller	1944	−8.4	41	750	140	6.56
	Terfel	2007	−29.4	132	1150	400	6.67
Wasserflut	Baer	2007	−9.4	86	3150	100	6.39
	Fischer-Dieskau	1988	−5.9	95	3600	100	5.85
	Goerne	2005	−11.8	104	3000	350	6.04
	Hampson	2005	−6.5	116	2300	50	5.71
	Henschel	2006	−9.2	131	2850	1850	5.50
	Husch	1999	−13	77	2800	450	6.38
	Mason	1995	−11	93	3500	150	6.00
	Muller	1988	−6.5	92	3080	80	5.70
	Prey	2008	−8.3	158	3800	150	5.62
	Vickers	1992	−22.2	106	4300	600	6.10
Mean (SD) of all Lieder			−11.6 (6.42)	106 (33)	2107 (1377)	276 (424)	6.16 (0.394)
M1 (D4)	Allen	1982	−9.1	119	650	50	6.67
	Fischer-Dieskau	1968	−7.8	151	550	50	6.22
	Hampson	1991	−8.3	98	610	40	6.45
	Lemalu	2005	−5.2	185	700	100	6.67
	Poell	1955	−4.9	196	350	100	6.50
	Skovhus	2007	−5.1	146	260	50	5.98
	Stone	2009	−12	129	900	150	5.93
	Tahu Rhodes	2003	−2	106	470	40	5.97
	Terfel	2006	−7.4	177	220	30	6.71
	Van Dam	2005	−4.1	178	250	80	5.88
M2 (A3)	Allen	1982	−8.4	146	660	30	5.56
	Fischer-Dieskau	1968	−3.8	175	750	50	6.43
	Hampson	1991	−3.7	124	400	10	6.30
	Lemalu	2005	−3	193	750	50	6.11
	Poell	1955	−0.3	172	400	50	5.71
	Skovhus	2007	−7.7	64	220	20	6.28
	Stone	2009	−16.1	163	940	30	6.35
	Tahu Rhodes	2003	3.3	170	650	150	6.30
	Terfel	2006	−8.3	100	380	20	5.56
	Van Dam	2005	−3.8	172	550	160	6.41
Means (SD) of all Opera			−5.885 (4.212)	148 (36.5)	533 (220)	63 (45.5)	6.20 (0.351)

the two genres (H4). H4 aims to reproduce the results of study 2, in regards to vibrato rate. Furthermore, because the published literature has not reached a consensus on how to best measure LSF in relation to the overall sound level of a sung pitch, we decided to explore three methods, which are presented in the Procedures of the Method section.

TABLE 4.
Mann-Whitney *U* Tests for SF Strength, Vibrato Extent, and Vibrato Onset Time With Genre (Lieder or Opera) as the Independent Variable

	SF Strength	Vibrato Extent	Time Until Vibrato Onset	Vibrato Rate
Mann-Whitney U	68.5	69	70.5	170
Z	−3.261	−3.245	−3.223	−0.293
Asymp. Sig. (2-tailed)	0.001	0.001	0.001	0.77

Abbreviation: Asymp. Sig., asymptotic significance.

Method

Stimuli. One hundred excerpts, all performed by baritones, were selected for analysis of SF (Table 5). Inclusion criteria were as follows:

1. The performers had to be highly acclaimed, international professionals in both lied and opera.
2. The recordings were available.
3. Sung pitches could be identified that had minimal or no simultaneous accompaniment (eg, orchestral or piano).

All selected excerpts consisted of a single note with a minimum duration of 500 milliseconds. The sung pitches in the SF analysis ranged from Ab2 F_0(~103 Hz) to G4 (F_0~392 Hz). A second set of 90 excerpts from the same six performers were selected for vibrato analysis (Table 6). Of these 90 excerpts, 86 were taken from the initial batch of 100 SF excerpts, with 14 SF excerpts deemed unsuitable for vibrato analysis due to interference from accompaniment. The additional four excerpts added to the vibrato set were taken from the same contingent of performers and pieces. The inclusion criteria for vibrato excerpts were identical to that of the SF stimulus set. The sung pitches in the vibrato analysis ranged from F# (F_0~174 Hz) to G#4 (F_0~415 Hz).

Procedures. Excerpts meeting the selection criteria were extracted and saved as separate audio files. Each excerpt consisted of a single vowel, plus consonants, performed on a single pitch, as per study 2. Excerpts from the SF stimulus set were analyzed at each note of interest using the frequency spectrum analyzer routine in *MATLAB R2013b*.[32] Spectral analysis was performed with a Welch frequency spectrum using a Hann window. An FFT window size of 1024 samples was used for F_0, and 512 samples for F_1–F_5. Stimuli were sampled at a rate of 22050 Hz, and with a sample duration of 200 milliseconds. The spectra used a 20% overlap between segments. Frequency and intensity levels for F_0 and F_1–F_5 were determined using the *findpeaks* function. The F_0, formant frequency, and time readings were transferred onto a spreadsheet before statistical analyses. All statistical analyses were performed using *SPSS statistics package version 22*.[33]

Intensity levels for F_0, and F_1–F_5 were recorded in decibels; however, F_3, F_4, and F_5 were not always clearly identifiable. To aid with identification of ambiguous excerpts, F_3, F_4, and F_5 were assumed to lie in the region of the frequency spectrum where the SF is expected, from 2200 Hz to 4500 Hz.[2] F_1 was assumed to lie in the region of vowel definition, up to 900 Hz in the frequency spectrum.[34] Given the variety of options for quantifying SF, three methods were applied here, which we refer to as *A*, *B*, and *C*.

- Method *A* compared the level of the putative (visually obvious) peak in the spectral envelope in the region of vowel definition (henceforth 'vowel putative') with the level of the putative peak in the region in which SF (henceforth 'SF putative') is visually expected. The difference in level between these two points was recorded as the LSF in units of decibels, as summarized in Figure 4.
- Method *B* is identical to the method used in study 2. This method compared F_1 instead of vowel putative with the SF putative as summarized in Figure 4.

TABLE 5.
Performers and Works From Which Excerpts Were Selected for SF Analysis in Study 3 (Lied Compared With Opera)

Performer	Piece	Genre	Number of Excerpts
Bryn Terfel	*Der Erlkönig*	Lied	3
Bryn Terfel	*Schwanengesang*	Lied	3
Bryn Terfel	*Der Fliegende Holländer*	Opera	9
Bryn Terfel	*Don Giovanni*	Opera	2
Bryn Terfel	*Le Nozze di Figaro*	Opera	5
Dietrich Fischer-Dieskau	*Schwanengesang*	Lied	2
Dietrich Fischer-Dieskau	*Winterreise*	Lied	7
Dietrich Fischer-Dieskau	*Der Fliegende Holländer*	Opera	8
Dietrich Fischer-Dieskau	*Le Nozze di Figaro*	Opera	4
Hermann Prey	*Morgen*	Opera	5
Hermann Prey	*Winterreise*	Lied	6
Hermann Prey	*Le Nozze di Figaro*	Opera	5
José Van Dam	*Schwanengesang*	Lied	2
José Van Dam	*Winterreise*	Lied	4
José Van Dam	*Der Fliegende Holländer*	Opera	10
Matthias Goerne	*Schwanengesang*	Lied	2
Matthias Goerne	*Winterreise*	Lied	4
Matthias Goerne	*Le Nozze di Figaro*	Opera	5
Matthias Goerne	*Tannhäuser*	Opera	2
Thomas Allen	*Winterreise*	Lied	5
Thomas Allen	*Don Giovanni*	Opera	2
Thomas Allen	*Le Nozze di Figaro*	Opera	5

TABLE 6.
Performers and Works From Which Excerpts Were Selected for Vibrato Analysis in Study 3 (Lied Compared With Opera)

Performer	Piece	Genre	Number of Excerpts
Bryn Terfel	*Der Erlkönig*	Lied	2
Bryn Terfel	*Schwanengesang*	Lied	5
Bryn Terfel	*Der Fliegende Holländer*	Opera	9
Bryn Terfel	*Le Nozze di Figaro*	Opera	1
Dietrich Fischer-Dieskau	*Schwanengesang*	Lied	3
Dietrich Fischer-Dieskau	*Winterreise*	Lied	7
Dietrich Fischer-Dieskau	*Der Fliegende Holländer*	Opera	7
Dietrich Fischer-Dieskau	*Le Nozze di Figaro*	Opera	3
Hermann Prey	*Morgen*	Lied	5
Hermann Prey	*Winterreise*	Lied	6
Hermann Prey	*Le Nozze di Figaro*	Opera	4
José Van Dam	*Winterreise*	Lied	4
José Van Dam	*Der Fliegende Holländer*	Opera	10
Matthias Goerne	*Schwanengesang*	Lied	1
Matthias Goerne	*Winterreise*	Lied	6
Matthias Goerne	*Le Nozze di Figaro*	Opera	5
Matthias Goerne	*Tannhäuser*	Opera	2
Thomas Allen	*Winterreise*	Lied	5
Thomas Allen	*Le Nozze di Figaro*	Opera	5

- Method *C* compared the level of F_1 with the level of F_3. The difference in level between these two points was recorded as the LSF in units of decibels as summarized in Figure 4.

Method *A* is similar to one used by Schutte and Miller.[35] Schutte and Miller used a vowel putative and SF putative; however, their method differs to method *A* when multiple prominent peaks are present in either of these spectral areas. In this circumstance, Schutte and Miller replaced the putative peak level with the level half way between the two highest peaks. Schutte and Miller do not outline any further criteria for identifying exactly how prominent a secondary peak must be or discuss steps for encountering a larger number of competing peaks. When compared with method *A*, we feel this is a limited method for measuring LSF, as both the spectral areas used for

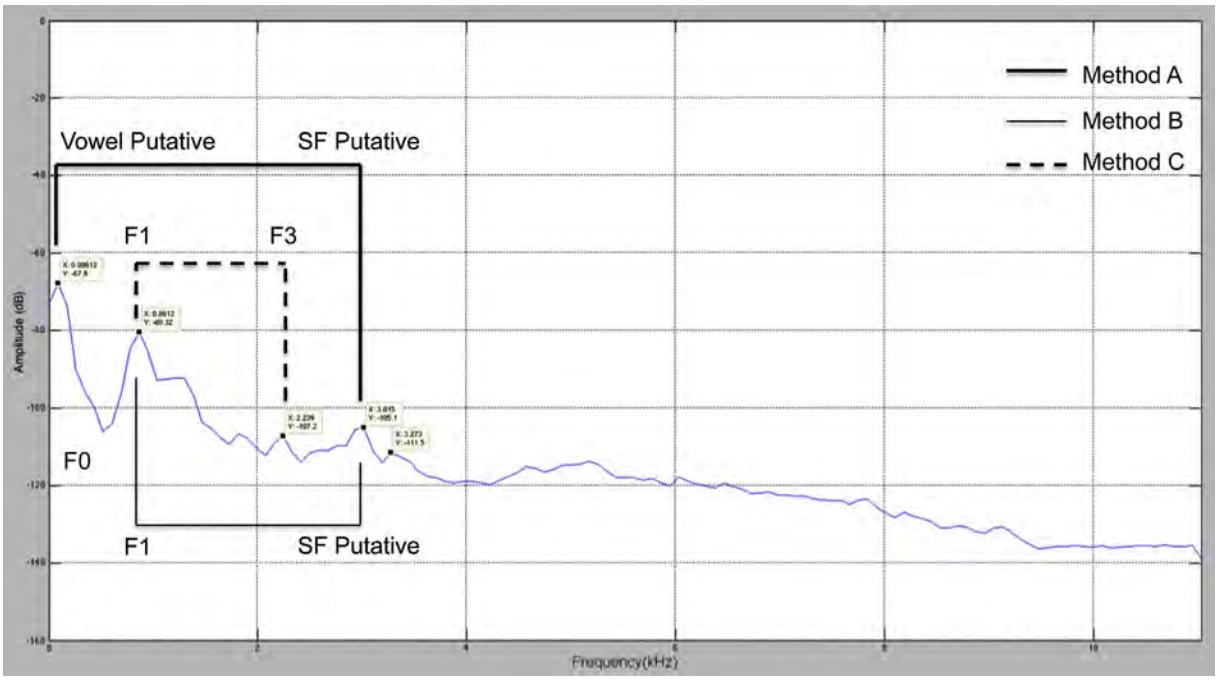

FIGURE 4. An example of methods A, B and C, as used to calculate LSF in study 3. This example depicts an excerpt from *Der Erlkönig*, as performed by Bryn Terfel.

the vowel putative and SF putative will often be comprised of several peaks. We argue that the use of a single measurement from the most prominent peak is a more reliable point for comparison among a large number of stimuli than an average of multiple peaks.

In a study such as this, in which varying vowel sounds are compared, we propose that methods without a vowel putative may suffer from inconsistencies. Methods *B* and *C* fall into this category. Specifically, any change in the positioning of the lips, jaw, tongue, velum, and larynx can affect the amplitude and frequency of F_1.[34] Furthermore, although F1 commonly exhibits the highest intensity level in the area of the spectrum near F_0, which is often the basis for its use as a reference point, this is not always the case. It is possible that a certain vowel sound with a lower F_1 level could be measured at a lower intensity than the reading F_0 or another prominent peak of the same vowel would give. In this situation, we argue a vowel putative has greater flexibility. We do, however, acknowledge that in a study comparing examples of matched vowel sounds, in which F_1 could be expected to be similar across examples, vowel measurement based solely on F_1 could prove more reliable.

Method *C* may also produce an inaccurate representation of the overall level of the SF due to the potential distance from the actual, though not graphically visible SF putative. This inaccuracy arises from the formation of the SF, which is sometimes proposed to be a clustering of F_3, F_4, and F_5.[2] As an example, the SF clustering in Figure 4 shows F_3 at a lower level than F_4 and could therefore be an inaccurate estimate of the level of SF intensity, which by simple visual extrapolation appears to be closer to the intensity of F_4.

Additionally, the frequency spectra used in study 2 featured a significantly lower window size (128 samples) than those used in this study. This lower sample size often yielded two broad peaks in the spectral envelope, labeled as F_1 and SF, as opposed to a collection of partials across the spectrum. The deliberately low frequency resolution of this method avoids the confusion of partials and formants and negated the need for putative peaks. However, the low frequency resolution can also be seen as presenting lower precision.

Excerpts on the vibrato list were analyzed at each note of interest using the *MIRtoolbox 1.3.3* in *MATLAB R2013b.*[32] Pitch vibrato measurements were determined with the *mirspectrum* spectrogram using an FFT window length of 1024 samples and a Hamming window. The spectrogram frame ranged from 0 to 3500 Hz. Peaks and troughs were determined with the *ginput* function. The pitch and time readings were transferred onto a spreadsheet prior to statistical analyses. All statistical analyses were performed using *SPSS statistics package version 22.*[33] Vibrato calculations were made following the same methods as study 2 with an example shown in Figure 5.

Results

Results pertaining to hypotheses tested. SF was observed in all 55 opera excerpts and in 40 of the 45 lieder excerpts. All three methods yielded a stronger mean LSF for operatic excerpts compared with lieder excerpts. The results of method *A* are summarized in Figure 6.

In regards to method *A*, the results of five of the six singers show a trend in which operatic performances feature a stronger mean within-singer SF than the lieder performances. These observations support the view that the SF is present in lieder performance, albeit in a slightly weaker form. Goerne featured a higher LSF in lieder excerpts; however, these results may be

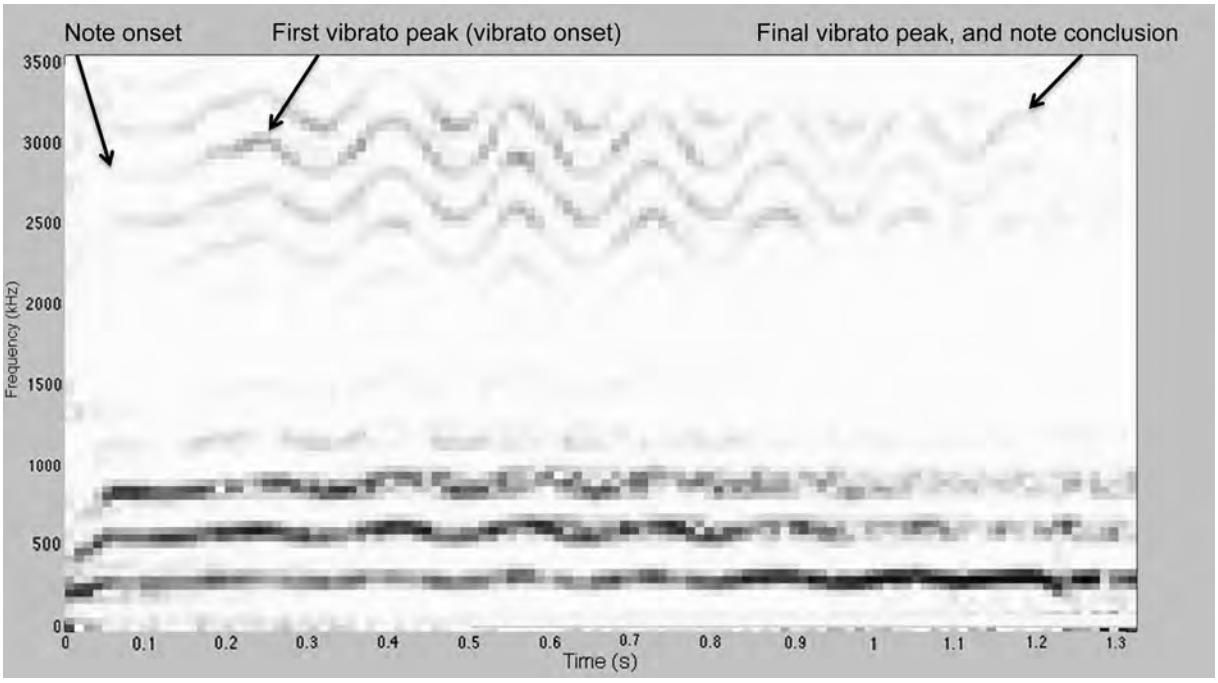

FIGURE 5. Spectrogram analysis used to identify vibrato parameters in study 3. This example depicts an excerpt from Schubert's Die Wetterfahne, as performed by Thomas Allen.

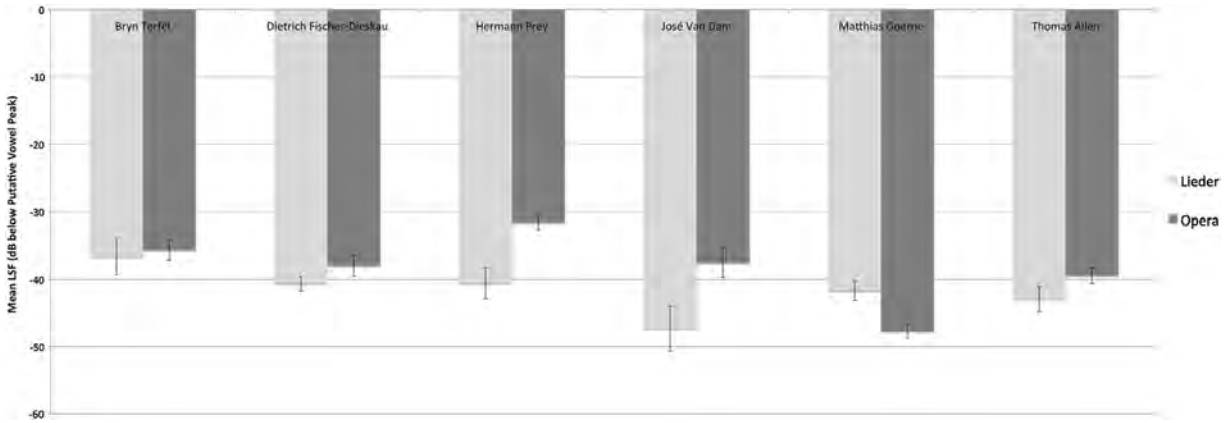

FIGURE 6. Summary of LSF results in study 3, according to method A. Results are displayed by performer and genre. Error bars = ±1 standard error.

because of the low number of excerpts used in each genre for this singer.

To examine whether any of the SF and vibrato parameters were different between lieder and opera performance, a series of *t* tests were conducted. Because of the low number of within performer excerpts, a meaningful within-singer paired comparison was considered inappropriate. Instead, these observations were further investigated through inferential statistical methods on a pooled genre-by-genre basis. These results are summarized in Table 7.

Using LSF method *A*, $M = -38.5$ dB, standard deviation [SD] = 6.32, n = 55 for opera excerpts which is higher than $M = -41.8$ dB, SD = 7.0, n = 45 for lieder excerpts. The difference in LSF between the two genres was statistically significant according to an independent samples *t* test (t(98) = 3.209, p = 0.002). These results are summarized in Figure 7.

Method *B* yielded $M = -32.7$ dB, SD = 6.7, n = 55 for opera excerpts and $M = -36.5$ dB, SD = 7.2, n = 45 for lieder excerpts. The difference in LSF between the two genres was statistically significant according to an independent samples *t* test (t(98) = 3.26, p = 0.002). Method *C* yielded $M = -37.5$ dB, SD = 6.1, n = 55 for all opera excerpts and $M = -40.3$ dB, $SD = 8.3$, n = 45 for all lieder excerpts. The difference in LSF between the two genres was statistically significant according to an independent samples *t* test (t(98) = 2.302, p = 0.023).

The results of method *A* showed the lieder excerpts to have a mean LSF 3.3 dB weaker than the operatic excerpts. The results of method *B* showed the lieder excerpts to have a mean LSF 3.8 dB weaker than the operatic excerpts, and method *C* showed the lieder excerpts to have a mean LSF 2.8 dB weaker than the operatic excerpts. In comparison, the results of study 2 showed the lieder excerpts to have a mean LSF 5.7 dB weaker than the operatic excerpts. Overall, the mean relative SF intensity is significantly lower in lieder compared with opera by approximately 3 to 5 dB regardless of the method used to perform the calculation.

Overall, vibrato onset time was shorter for the operatic excerpts ($M = 80$ ms, SD = 110, n = 46) when compared with the lieder excerpts ($M = 1605$ ms, SD = 180, n = 44). This difference was statistically significant according to a *t* test (t(88) = −2.153, p = 0.007, but note that the assumptions of normality for these temporal data are quite likely violated and should be treated with caution). On examining the proportional time of entry of vibrato onset with respect to note duration (vibrato onset time/note duration time), the ratio was 0.11 for lieder and 0.07 for opera. This means that, on average, lieder excerpts began vibrato proportionally later (with respect to note duration) than opera excerpts. That is, typically 11% of the note has elapsed before vibrato commences in a lied note, whereas only 7% of the note has elapsed in an opera note. These results correspond with the findings of study 2, in which onset time was also statistically different between the genres.

Operatic excerpts featured a slightly greater vibrato rate ($M = 6.8$ Hz, SD = 1.9, n = 46) than lieder excerpts ($M = 6.4$ Hz, SD = 1.1, n = 44). In similar fashion to the results of study 2, this study was not able to identify a difference in vibrato rate between the two genres with statistical significance according to an independent samples *t* test (t(88) = 1.205, p = 0.231). Additionally, in concert with study 2, the average vibrato extent was wider in operatic examples ($M = 127$ cents, SD = 56, n = 46) than in lieder excerpts ($M = 118$ cents, SD = 49, n = 44). No significant difference was found in vibrato extent between the two genres according to an independent samples *t* test (t(88) = 0.867, p = 0.256). The difference in vibrato extent between the two genres was slightly less than in study 2, which featured an average extent of 148 cents in opera and 106 cents in lieder. The lower values could be an artifact of the slight differences in signal processing and analysis software used across the two studies.

Comparison of three methods for calculating level of the SF.

The three methods used to calculate LSF produced fairly consistent results with each other, and they were consistent with the calculations in study 2 although small discrepancies could be explained by the different software (*Audacity* in study 2, *MATLAB* in study 3) and software settings used across the two studies. Methods *A* and *B* produced comparable results, in which the mean LSF of each singer was consistently weaker in lieder by a similar difference. Method *C* produced LSF readings

TABLE 7.
Summary of Acoustic Analysis for Study 3

	Vocalist	Mean LSF (Method A, dB)	Mean LSF (Method B, dB)	Mean LSF (Method C, dB)	Mean Vibrato Onset Time (ms)	Mean Vibrato Onset Time Ratio	Mean Note Duration (ms)	Vibrato Extent (cents)	Vibrato Rate (Hz)
Lieder excerpts	Bryn Terfel	−37	−33	−35	224.7	0.116	2240	165.44	5.43
	Dietrich Fischer-Dieskau	−41	−37	−40	133.76	0.101	1540	124.07	6.36
	Hermann Prey	−41	−31	−34	111.35	0.132	820	105.91	7.09
	José Van Dam	−47	−42	−44	199.28	0.102	1880	84.59	5.8
	Matthias Goerne	−42	−37	−42	217.04	0.135	1430	103.66	6.75
	Thomas Allen	−43	−39	−47	84.62	0.065	1450	110.79	6.12
	Mean (SD) of all Lieder excerpts	−41.8 (6.93)	−36.5 (7.24)	−40.3 (8.34)	156.25 (176.92)	0.113 (0.106)	1450 (991.31)	117.77 (48.42)	6.38 (1.05)
Opera excerpts	Bryn Terfel	−36	−30	−35	21.89	0.019	1207	164.86	6.97
	Dietrich Fischer-Dieskau	−38	−33	−35	119.69	0.08	1783	114.46	6.33
	Hermann Prey	−32	−26	−37	183.65	0.151	1271	84.43	6.21
	José Van Dam	−38	−32	−34	74.75	0.073	1056	138.47	6.57
	Matthias Goerne	−48	−40	−46	107.71	0.105	822	130.02	6.96
	Thomas Allen	−39	−35	−38	41.54	0.029	1152	86.8	7.8
	Mean (SD) of all Opera excerpts	−38.5 (6.32)	−32.7 (6.68)	−37.5 (6.14)	83.9 (107.44)	0.07 (0.08)	1241 (754)	127.39 (55.23)	6.77 (1.86)

that were slightly closer between the two genres. Additionally, when examining LSF for individual singers, methods *A* and *B* produced relatively higher values of LSF in opera for five of the six singers, whereas method *C* produced results in which only three singers exhibited a stronger LSF in opera and the remaining three a stronger LSF in lieder. The results of Goerne's performances, which had a stronger LSF in lieder regardless of the method used, can be attributed to a personal technique choice by this performer. However, we decided to perform a visual spectral investigation on the results of Terfel and Prey, who showed a stronger LSF in opera for methods *A* and *B* and the opposite for method *C*. That is, method *C* produced the most conservative readings (identifying fewer differences), but with some inconsistencies for the Terfel and Prey excerpts.

Although all three methods showed a statistically significant difference in LSF between the two genres for the pooled data, an investigation of the spectra revealed that method *C* may be less accurate than either of the other two methods that used putative peaks in the calculation. This stems from the use of only F_3 to measure the entire SF intensity, as discussed in the Method section. Of the 22 excerpts performed by Terfel, the level of F_3 appeared to be a relatively inaccurate representation of the overall SF in 10 examples (44% of the excerpts). Similarly, with the 16 excerpts performed by Prey, spectral investigation revealed F_3 to be a relatively inaccurate indicator of the overall SF in seven examples (45% of the excerpts). Two of these examples are displayed in Figure 8.

A second investigation of the spectra was also conducted to ascertain the accuracy of method *B* compared with method *A*. Although F_1 is often the highest intensity formant[36] in the vowel region, it is not unusual for F_0 or another peak to dominate. In 40 of the 100 excerpts used in method *B*, F_1 was significantly lower level than F_0. These 40 excerpts were spread relatively evenly across the two genres. They accounted for 14 out of the 38 lieder excerpts (37%), and 26 out of 62 opera excerpts (42%). An example of this is displayed in Figure 9. With this in mind, we argue that method *A* corresponds most closely with the visually apparent location of the SF of the three methods. This largely stems from the flexibility the method is afforded from having two putative peaks and is therefore able to produce plausible results regardless of whether F_1 is the most prominent vowel peak or not. Method *C*, then, may also be viewed as a more conservative approach, but further research will be needed to determine the validity and reliability of the different methods in more detail.

Discussion

Study 3 investigated three hypotheses that were tested in study 2 and that were based on the literature review and qualitative findings of study 1. A fourth hypothesis was included in an attempt to reproduce the vibrato rate results of study 2. In all four hypotheses, the findings of study 2 were replicated. As with study 2, a stronger SF was observed in opera excerpts (H1), and a shorter mean onset time was observed in opera examples (rejecting H2). A slightly wider mean vibrato extent for opera examples was found, in accordance with H3. No difference was observed in mean vibrato rate between the two genres (H4).

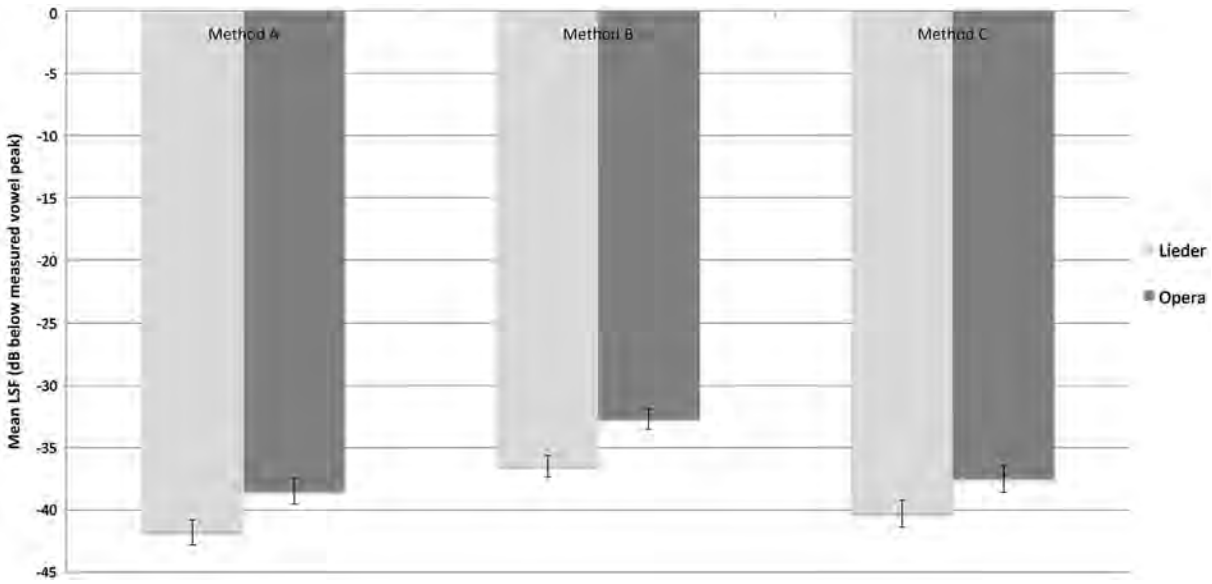

FIGURE 7. Summary of LSF results in study 3, displayed by genre and method. Error bar = ±1 standard error.

With the replication of the findings, and consistency across studies, we need to explain why lied singers do not use SF to the same extent as opera singers, even if the same performer is observed. Explanations from study 1 suggest that opera singers include SF in lieder to retain the specific tonal qualities of the technique; however, the smaller scale of accompaniment in lieder requires less voice projection and therefore a lesser degree of SF. Importantly, when a singer is competing with an orchestra, they require additional resources apart from sheer intensity as the required intensity could not be achieved against large orchestral forces accompanying at high volume. The singer is more likely to project over the orchestra by applying a stronger SF as well as a near continuous vibrato and wider vibrato extent. Analyses have revealed that SF occupies a spectral region in which orchestral intensity is relatively weak.[36,37] It is through this spectral "gap" that a combination of SF and vibrato presents an alternative and contrasting path for the singer to allow identification of their vocal line. The findings from study 1 also suggest that a lesser use of the SF found in lieder makes the voice more speech-like, allowing greater focus on the communication of the words in a more natural language way that is required in opera.

In both studies 2 and 3, mean vibrato onset time was found to be significantly shorter in opera excerpts. However, not all singers conformed to this finding. Prey produced a longer mean vibrato onset time in opera (180 milliseconds) compared with lieder (110 milliseconds). This may be explained by the considerably shorter mean duration of Prey's lieder excerpts (820 milliseconds) compared with opera excerpts (1708 milliseconds). However, examining the temporal ratio

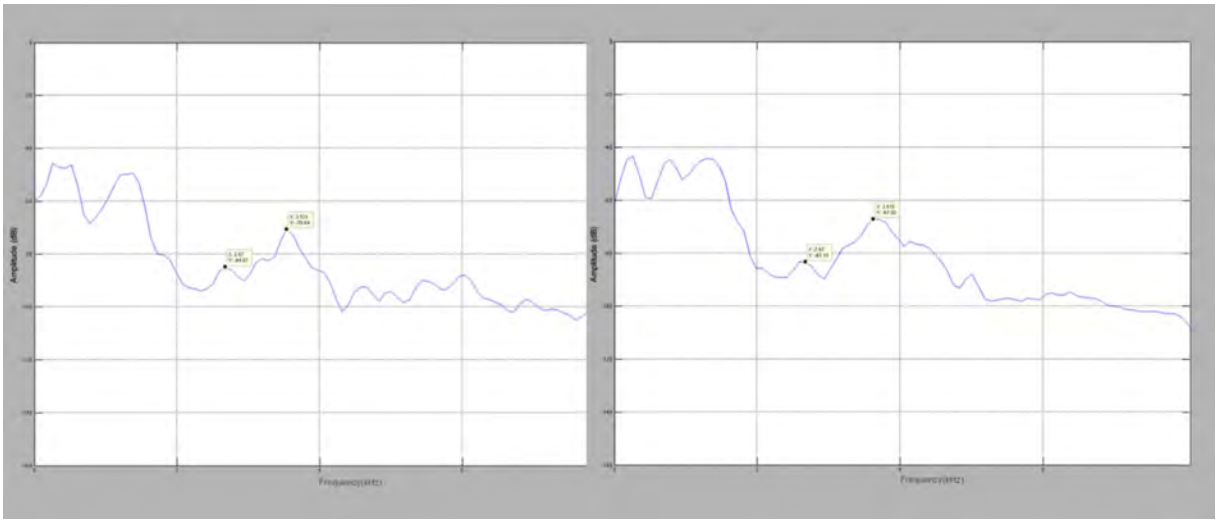

FIGURE 8. Two spectral examples in which the level of F_3 is at a significantly lower level than the rest of the SF (measured here with an SF putative). The example on the left is an excerpt from *Don Giovanni*, as performed by Bryn Terfel. The example on the right is an excerpt from *Le Nozze di Figaro*, as performed by Hermann Prey.

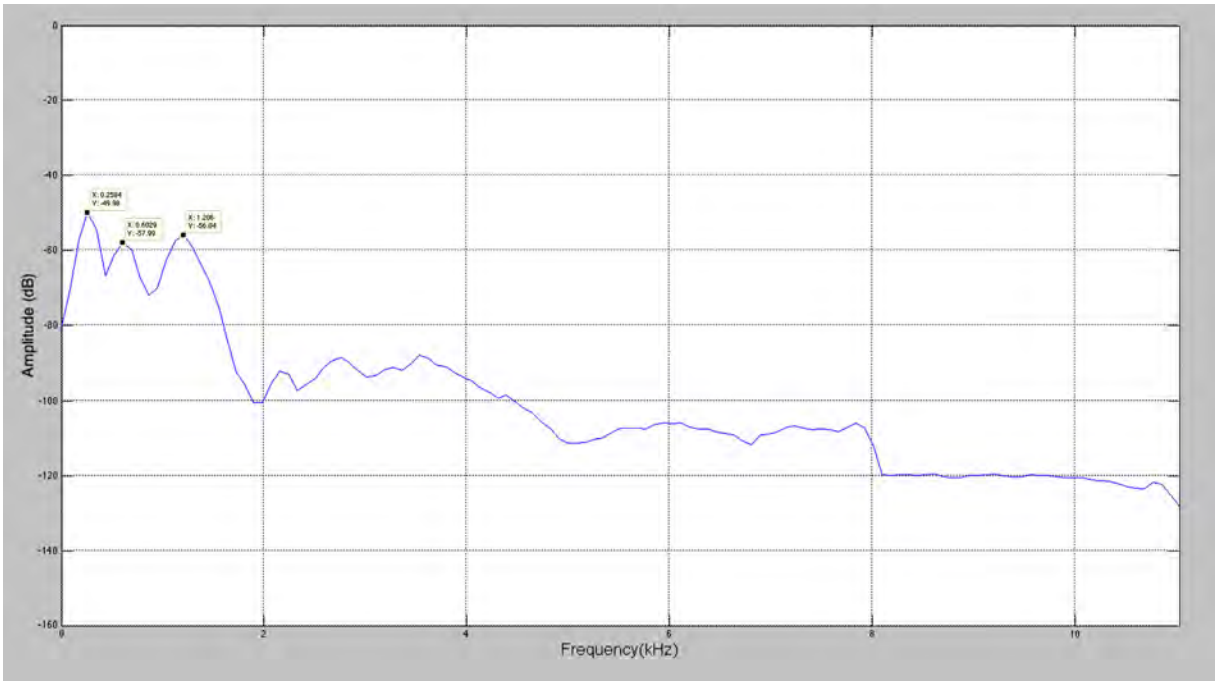

FIGURE 9. A spectral example in which F_1 is not the dominant peak in the vowel region. This example depicts an excerpt from *Der Fliegende Höllander*, performed by José Van Dam. Measurements are shown for F_0, F_1, and F_2 (left to right).

of vibrato onset time with respect to the overall duration of the note, hence adjusting for possible biases caused by comparing notes of different duration, we find a rather similar ratio of 15% vs 13% for opera versus lied respectively (Table 7). Similarly, Fischer-Dieskau also demonstrated a longer mean onset time in opera (130 milliseconds) compared with lieder (120 milliseconds) although the difference here is much less pronounced. The vibrato onset ratio is again very similar (8% for opera and 10% for lied). This too can be attributed to note durations, in which the lieder excerpts were on average slightly shorter (1540 milliseconds) than opera (1780 milliseconds). With this in mind, the data strongly suggest a connection between the performance of lied and a delayed vibrato onset, in relation to the length of the overall note duration.

Mean vibrato extent was shown to be slightly wider in the opera examples. This result is consistent with other studies, such as one[8] in which vibrato extent was found to be greater in opera (∼98 cents) than Broadway musical vocal styles (∼78 cents). The difference in vibrato extent also supports the findings of study 2 although the results in this study are not as pronounced. The overall wider extent and earlier onset of vibrato relative to lied singing found in this article suggests another strategy singers use to compete with the larger forces of their accompaniment, in addition to a strong SF. Orchestral instruments usually have vibratos with a rather smaller amplitude than that of the voice.[38] Furthermore, an tutti orchestra almost never coordinates vibrato. Consequently, a singer with a large amplitude vibrato, maintained for a relatively long fraction of the note, gives the audience a better chance to identify the singer. This is both through contrast with the relatively stable tones of the orchestra, and possibly also through the perceptual advantage of creating a varied channel of pitch delivery, allowing audience detection of that voice through auditory stream analysis.[39] Expressly, vibrato and SF address practical performance issues. Aesthetic issues may emerge as a result of the success or otherwise of solving the practical problems. That is, if "more" vibrato allows for clearer communication in opera, the use of vibrato may then become the aesthetic preference in opera as well.

As with study 2, mean vibrato rate was found to be similar between the two genres. The slight difference in rate between study 2 and study 3 is comparable with those found by Prame[21] although that study did not use examples of opera or lied.

GENERAL DISCUSSION AND CONCLUSION

This research demonstrated that several of the assumptions made by teachers and performers about the techniques that are appropriate for the performance of lieder can be investigated through acoustic measurement. An important contribution made in the present research is that it has enabled, for the first time, a systematic scientific approach to examining some of the features associated with good lieder technique and facilitates clearer definitions that may be applied by singers and teachers of the genre. Although previous qualitative reports of singing style provide generic explanations with poorly defined terms, such as "natural vibrato," we are able to propose here some of the ways that these descriptions may manifest themselves in the physical, acoustic properties of a voice aiming to produce an appropriate and perhaps even good performance of lieder, eg, in terms of vibrato "width" (in terms of pitch) and amount of delay before the onset of vibrato for a

particular pitch, vowel, or phoneme. We suggest that when a vocal pedagogue refers to purity of tone, from an acoustic perspective, this corresponds to a delayed vibrato onset time with a narrower vibrato width compared with vocal styles that use a larger, louder combined accompanying force. Relatively lower relative intensity of the SF may also contribute to the pedagogues concept of purity of tone.

By marrying the quantitative and qualitative data, we are able to conclude that, with respect to opera at least, good lieder performance will be sung with a more delayed vibrato onset, less vibrato extent, and a weaker level of singers formant. Although descriptions of what makes a good lied performance can encapsulate the essence of these acoustic characteristics, until the present research there was no evidence that could clarify and directly link such language to its acoustic counterparts. We believe that bringing together the scientific and conventional pedagogical approaches will help to provide a better understanding of the nature of singing and, here, has provided a better understanding of the performance of lieder.

The lack of consensus between the survey (study 1) and the acoustic analyses (studies 2 and 3) in regards to vibrato onset time may suggest that discrepancies are occurring from theoretical performance instruction to practical performance. There are also variations among singers in the features they exploit in the operatic context. For example, not all singers may want to or need to use a particularly strong SF, early vibrato onset, and wide vibrato extent all of the time. Furthermore, although some discrepancies were encountered, several descriptions of good lieder technique were complemented by the results of the analyses. Among these, Elliott's[23] description of a "wide, continuous vibrato" being undesirable in lieder is congruent with the findings of Studies 2 and 3. But care should be taken not to assume that acoustic measurement provides a ground truth that informs pedagogical approaches. Methods are limited in acoustic measurement contexts, too. For example, we had limitations on the kinds of notes we could analyze because of our decision to use commercial sound recordings and of those to find notes that were not or weakly accompanied to allow accurate measurement. This approach could well have biased the sample of notes investigated, such as restricting the notes to parts of an opera that were recitative or recitative-like and therefore perhaps more speech-like. We only examined the lower voice types used in males. But confirmation of the effects of vibrato is an area for further research for female vocalists.

Furthermore, we identified a lack of consensus on how to measure the SF, having proposed three different methods in this research, which each produced slightly different results. Of these, our analysis suggested that method *C* has good utility: It compares the level of the third formant against the level of the first formant and although it produces some inaccuracies, it was found to be overall conservative and relatively simple to calculate via visual inspection of a frequency spectrum. However, method A appeared to show good face validity because of the phonetically plausible locations for selecting the two key readings on the frequency spectrum for calculating LSF.

Even with such complications, we still found consistencies across the qualitative and quantitative studies. Our point is that one should not assume one method of investigation is sovereign. Acoustic measurement and the accounts of teachers and scholars are all important kinds of data that can be considered in helping to build a rounder picture of what goes on in the process of producing lieder, and for that matter, good lieder.

We recommend further research on this topic, preferably with more comparable matching in regards to note duration, pitch, and the nature of the phonemes sung. Future investigations should also include more excerpts for each performer so that further comparisons can be made on a performer-by-performer level, rather than as a genre-specific analysis with pooled degrees of freedom. However, even with these factors, Studies 2 and 3 have cast light on quantitative differences in lied and opera singing styles and have demonstrated that although singing teachers have a good understanding of the required techniques that differentiate the styles, they may also benefit from examination and application of findings from quantitative measurements in terms of SF and vibrato. What remains to be seen is how the multitude of variables that constitute human vocal production (we only examined a small selection, concerned with SF and vibrato) will provide a deeper understanding of this fascinating and complex generator of aesthetic pleasure. Additionally, the practical consideration of accompaniment and environmental factors, such as piano accompaniment in a small auditorium compared with a symphony orchestra in a large auditorium, rather than genre alone, should receive further attention as the ontological basis of apparent differences between the two styles.

REFERENCES

1. Sataloff RT. *Professional Voice: The Science and Art of Clinical Care*. New York: Raven Press; 1991.
2. Sundberg J. Level and center frequency of the singer's formant. *J Voice*. 2001;15:176–186.
3. Sundberg J, Romedahl C. Text intelligibility and the singer's formant - a relationship? *J Voice*. 2009;23:539–545.
4. Bretos J, Sundberg J. Measurements of vibrato parameters in long sustained crescendo notes as sung by ten sopranos. *J Voice*. 2003;17:343–352.
5. Howes P, Callaghan J, Davis P, Kenny D, Thorpe W. The relationship between measured vibrato characteristics and perception in Western operatic singing. *J Voice*. 2004;18:216–230.
6. Sundberg J. My research on the singing voice from a rear-view perspective. *THM-QPRS*. 2003;45:11–22.
7. Sundberg J, Gramming P, Lovetri J. Comparisons of pharynx, source, formant and pressure characteristics in operatic and musical theatre singing. *J Voice*. 1993;7:301–310.
8. Stone RE, Cleveland TF, Sundberg J, Prokop J. Aerodynamic and acoustical measures of speech, operatic, and Broadway vocal styles in a professional female singer. *J Voice*. 2003;17:283–297.
9. Bezerra A, Cukier-Blaj S, Duprat A, Camargo Z, Granato L. The characterisation of the vibrato in lyric and Sertanejo singing styles: acoustic and perceptual auditory aspects. *J Voice*. 2009;23:666–670.
10. Vurma A, Ross J. Where is a singer's voice if it is placed "forward"? *J Voice*. 2002;16:383–391.
11. Kenny D, Mitchell H. Acoustic and perceptual appraisal of vocal gestures in the female classical voice. *J Voice*. 2006;20:55–70.
12. Kenny D, and Mitchell H, 'Vocal quality in female classical singers: the role of acoustics, perception and pedagogy'. In Proceedings of the 3rd conference on Interdisciplinary Musicology, Tallinn, Estonia; 2007.
13. Ekholm E, Papagiannis G, Chagnin F. Relating objective measurements to expert evaluation of voice quality in Western classical singing. *J Voice*. 1998;12:182–196.

14. Bloothooft G, Plomp R. Spectral analysis of sung vowels: III characteristics of singers and modes of singing. *J Acoust Soc Am*. 1986;79:852–864.
15. Mitchell H, Kenny D. The impact of 'open throat' technique on vibrato rate, extent and onset in classical singing. *Logoped Phoniatr Vocol*. 2004;29:171–182.
16. Mitchell H, Kenny D. Open throat: acoustic and perceptual support for pedagogic practice. *J Singing*. 2008;64:429–441.
17. Diaz J, Rothman H. Acoustical comparison between samples of good and poor vibrato in singers. *J Voice*. 2003;17:179–184.
18. Hakes J, Shipp T, Doherty E. Acoustic properties of straight tone, vibrato, trill and trillo. *J Voice*. 1987;1:148–156.
19. Sundberg J. Acoustic and psychoacoustic aspects of vocal vibrato. *STL-QPSR*. 1994;35:45–67.
20. Weiss R, Brown W, Morris J. Singer's formant in sopranos: fact or fiction. *J Voice*. 2001;15:456–468.
21. Prame E. Measurements of the vibrato rate of ten singers. *J Acoust Soc Am*. 1994;96:1979–1984.
22. Prame E. Vibrato extent and intonation in professional Western lyric singing. *J Acoust Soc Am*. 1997;102:616–621.
23. Elliot M. *Singing in Style: A Guide to Vocal Performance Practices*. Yale University Press; 2006.
24. Spillman R. Performing lieder: the mysterious mix. In: Hallmark R, ed. *German lieder in the nineteenth century*. New York: Schirmer Books; 1996.
25. Miller R. *Singing Schumann: An Interpretive Guide for Performers*. New York: Oxford University Press; 1999.
26. Finley R, and Finley C, 'SurveyMonkey'. California, U.S.A.; 2009.
27. Wapnick J, Ekholm E. Expert consensus in solo voice performance evaluation. *J Voice*. 1997;11:429–436.
28. Joliveau E, Smith J, Wolfe J. Vocal tract resonances in singing: the soprano voice. *J Acoust Soc Am*. 2004;116:2343–2439.
29. Audacity 1.3.3 Beta, multi-track audio editor and recorder, http://audacity.sourceforge.net. GNU General Public License software.
30. Rossing TD, Sundberg J, Ternström S. Acoustic comparison of voice use in solo and choir singing. *J Acoust Soc Am*. 1986;79:1975–1981.
31. Omori K, Kacker A, Carroll L, Riley W, Blaugrund S. Singing power ratio: quantitative evaluation of singing voice quality. *J Voice*. 1996;10:228–235.
32. Lartillot O, Toiviainen P, Eerola T. "A matlab toolbox for music information retrieval," in *Data analysis, machine learning and applications*. In: Preisach C, Burkhardt H, Schmidt-Thieme L, Decker R, eds. Berlin: Springer; 2008, pp.261–268.
33. SPSS IBM, 'SPSS Statistics Package Version 22'. New York, U.S.A.: 2013.
34. Sundberg J. *The Science of the Singing Voice*. Dekalb, Illinois: Northern Illinois University Press; 1987.
35. Schutte HK, Miller R. Intraindividual parameters of the singer's formant. *Folia Phoniatr(Basel)*. 1985;37:31–35.
36. Barnes JJ, Davis P. The relationship between professional operatic soprano voice and high range spectral energy. *J Acoust Soc Am*. 2004;116:530–538.
37. Sundberg J. The acoustics of the singing voice. *Sci Am*. 1977;236:82–91.
38. Fletcher NH. "Vibrato in Music". *Acoustics Australia*. 2001;29:91–102.
39. Bregman AS. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Massachusetts: MIT Press; 1990.