The vocal audition process: Do you hear what I hear?

Helen F Mitchell
Australian Centre for Applied Research in Music Performance (ACARMP),
Sydney Conservatorium of Music, University of Sydney, Australia
h.mitchell@usyd.edu.au

Dianna T Kenny
Australian Centre for Applied Research in Music Performance (ACARMP),
Sydney Conservatorium of Music, University of Sydney, Australia
d.kenny@usyd.edu.au - http://www2.fhs.usyd.edu.au/bach/staff/kenny/music.htm

In: K. Maimets-Volk, R. Parncutt, M. Marin & J. Ross (Eds.)
Proceedings of the third Conference on Interdisciplinary Musicology (CIM07)
Tallinn, Estonia, 15-19 August 2007, http://www-gewi.uni-graz.at/cim07/

Background in music performance and pedagogy. Experts' judgments in auditions and competitions may have significant impacts on musical careers. Few systematic research studies have identified objective (measurable) factors associated with musical potential or 'talent'. There are few high quality studies investigating the basis on which adjudicators make their determinations, or whether these are consistent across judges within an authentic performance assessment situation.

Background in music acoustics. There is a vast literature on vocal acoustic properties and the visual representation of voice. Characteristic long term average spectra (LTAS) shapes, particularly of classical or operatic voices have been accepted as an accurate visual representation of good sound in the genre and operatic singing is positively associated with an increase in energy between 2-4 kHz. However, matching measurements of comparative energy (between 0-2 and 2-4kHz) in LTAS curves to perceptual judgments of vocal quality have been inconclusive.

Aims. In this observational study, we assessed the relationship between audition outcomes, descriptions of vocal quality by audition adjudicators and acoustic parameters associated with successful and unsuccessful candidates in the classical singing stream who were applying for entry to the Bachelor of Music degree at a the Sydney Conservatorium of Music, Australia.

Method. Candidates (n=60) were assessed by a panel of expert pedagogues and performances were recorded using a head and torso simulator (HATS) placed adjacent to the panel. The top 24 singers were offered places (12 majors and 12 minors) and their recordings were matched with the next 23 singers who were not offered places in the program. Majors are considered to have more performance potential.
Candidates' songs were analysed for SPL and LTAS were performed on each singing performance. Measures of peak energy height [singing power ratio (SPR)], peak area [energy ratio (ER) and alpha (α)] were calculated on each LTAS.

Results. There were some significant differences between the acoustic output of females who were accepted versus female singers not accepted to the degree (mean SPL, α), but no differences between majors and minors or between any of the male candidates.

Conclusions. While standard acoustic measures successfully discriminated between Accepted/Not accepted, they were not sensitive enough to identify acoustic differences between genders or between Major/Minor. Future research must identify more sensitive acoustic measures to support perceptual judgments of differences between these two groups.

Implications. As findings of experimental research are being incorporated into texts on singing, researchers must interpret acoustic measures in relation to current singing pedagogy and practice so they can be useful tools for singing teachers and students. Acoustic science, psychoacoustics and perceptual psychology all have as yet unexplored potential for vocal pedagogy. Future research in singing will elaborate our findings for western classical voice into other voice genres.
While there is an established history of auditions for entry into a music performance degree; there is virtually no research that evaluates the audition process. This is the first study to match expert ratings and descriptors with acoustic measures of singers at tertiary entry level.

Tertiary entry auditions often mark the beginning or end to the musical careers of young singers. In a singing audition, judges make rapid decisions regarding the quality and potential of singing voices (Legge, 2001). The concepts of musical potential or 'talent'

remain elusive (Davidson & Da Costa Coimbra, 2001; Hollien, 1993; Watts, Barnes-Burroughs, Andrianopoulos, & Carr, 2003; Watts, Barnes-Burroughs, Estis, & Blanton, 2006) as does the process of adjudication (Davidson & Da Costa Coimbra, 2001; Stanley, Brooker, & Gilbert, 2002). Most available research studies on auditions use secondary level music students (Wapnick, Flowers, Alegant, & Jasinskas, 1993; Wapnick, Mazza, & Darrow, 1998).

Success at audition is the main criterion for music conservatorium admission (Subotnik, 2003). In an audition, assessors respond to an entire performance including the aural and visual cues (Legge, 2001) although their central concern remains on the vocal quality of singer and the degree to which the voice (and singer) is considered trainable. However, there is limited understanding of the selection criteria used by professional judges.

Perceptual research indicates that listeners show some degree of reliability and consistency in their overall judgments of good and poor vocal and instrumental performance (Ekholm, Papagiannis, & Chagnon, 1998; Geringer & Madsen, 1998; Saunders & Holahan, 1997; Smith, 2004; Wapnick et al., 1993) and more specifically in their assessment of overall voice quality (Kenny & Mitchell, 2006; Stanley et al., 2002; Wapnick & Ekholm, 1997). However, factors such as day, time of day, performer order and listener fatigue can all affect judge consistency (Bergee & Platt, 2003; Elliott, Schneider, & Zembower, 2000; Flores & Ginsburgh, 1996). Evaluation of music performance in tertiary institutions revealed that the assessment of singing and instrumental playing was based initially on the whole performance, before appraisal of specific technical and performance components (Davidson & Da Costa Coimbra, 2001; Stanley et al., 2002). In tertiary singing auditions, the panel expects an acceptable level of technical and vocal mastery. Unanimous consensus in a panel indicates that the voice is 'destined for greatness' (Subotnik, 2003, 2004). One study that evaluated the process of singing assessment (Davidson & Da Costa Coimbra, 2001) during the second year of a music degree found that the four assessors commented on technical facility, visual appearance and stage presence in their ratings, but did not define or discuss the basic quality of the voice in detail. Davidson and Coimbra (Davidson & Da Costa Coimbra, 2001) suggested that judges may have accepted the vocal quality as a 'stable element' without which candidates would not have been admitted to the course. Clarifying the characteristics of vocal quality is complex (Ekholm et al., 1998; Wapnick, Darrow, Kovacs, & Dalrymple, 1997; Wapnick & Ekholm, 1997). Since no research to date has linked expert judgments or rankings for admission to tertiary level music performance degrees with acoustic and perceptual factors of vocal quality, the study reported here attempted to address this central issue in vocal assessment.

There is a vast literature on vocal acoustic properties and the visual representation of the singing voice, which is regularly incorporated into singing texts (Miller, 1996; Nair, 1999; Sundberg, 1988). Acoustic measures such as long term average spectra (LTAS) have been used to differentiate between singing styles (pop versus classical) and different levels of professional achievement (eg performance at international, national or local venues). For classical and operatic voices, LTAS demonstrates an spectral reinforcement between 2-4 kHz (Borch & Sundberg, 2002; Sundberg, 1974; Thorpe, Cala, Chapman, & Davis, 2001), which is associated with singers in this genre (Barnes, Davis, Oates, & Chapman, 2004; Bartholomew, 1934; Sundberg, 1974).

A conventional way of reducing the information in the LTAS to a single meaningful number is to compute the ratio of energies above and below 2 kHz (Löfqvist, 1986; Löfqvist & Mandersson, 1987; Omori, Kacker, Carroll, Riley, & Blaugrund, 1996; Thorpe et al., 2001). Ratio measures (Omori et al., 1996; Thorpe et al., 2001) enable inter- and intra-singer comparisons. They have been used successfully to differentiate between singing and speaking, by singers of varying training (Omori et al., 1996) and to test optimal projection in professional singers (Barnes et al., 2004; Thorpe et al., 2001).

For singing students, ratio measures were not successful in differentiating between singing and speaking quality (Lundy, Roy, Casiano, Xue, & Evans, 2000) however, more recently, Watts et al (2006) showed that these measures could successfully differentiate between the vocal qualities of untrained talented and untrained non-talented singers.

Identifying talent or potential, as in a singing audition setting may be more problematic as singers are younger, and will have technical facility appropriate to age, training and experience, but are unlikely at that stage of their musical development to be able to demonstrate a fully mature or developed vocal instrument.

Although LTAS measures (ER and SPR) have successfully matched rankings of professional singing level (Bunch & Chapman, 2000) of six sopranos (Barnes et al., 2004), they were not successful in matching expert listeners' rankings of vocal quality in a group of advanced female singing students with the same voice type and level of singing training (Kenny & Mitchell, 2006; Mitchell & Kenny, 2004a). Visual inspection of LTAS in these studies showed a marked difference between individual singers (Barnes et al., 2004; Kenny & Mitchell, 2006; Thorpe et al., 2001). The sensitivity of measures of ER and SPR require further investigation with larger groups of singers to evaluate their use as an objective measure of vocal timbre.

While there is an established history of auditions for entry into a music performance degree, there is no research of which we are aware that evaluates the singing audition process. This study assessed the relationship between acoustic and perceptual characteristics of voices awarded a place at major (highest quality singers) and minor levels and compared these with unsuccessful auditions applicants to identify acoustic characteristics that discriminated between the three groups. A secondary analysis explored the relationship between acoustic parameters and with the panel's audition rankings.

Method

Participants

Singers (n=60) who auditioned for the Bachelor of Music degree at the Conservatorium of Music in 2006 were the participants in this study. Auditions took place in the Music Workshop at the Conservatorium of Music, Sydney, Australia. Singers were asked to sing two of the four songs or arias they had been instructed to prepare for their audition.

The audition panel comprised members of the vocal unit at the Sydney Conservatorium of Music, all highly renowned for their singing studio and singing pedagogy. The panel engaged in a group decision making process and arrived at selections and rankings using discussion and consensus (Subotnik, 2003).

Recordings

Auditions took place on a 13.94x16.04m stage area of a performance hall (Music Workshop) at the Conservatorium of Music. The audition panel sat on the stage area approximately 9m from the singer and piano. All singers were accompanied by the same expert pianist from the Conservatorium of Music.

A Head and Torso Simulator (HATS; Brüel & Kjær 4100-D) plus a matched pair of stereo microphones (Neumann KU140, in the ORTF configuration) were placed directly behind the audition panel (height to pinnae 155cm) approximately 10m from the piano and singer. The HATS mannequin has built-in ear simulators that provide a realistic reproduction of the acoustic properties of an average adult human head and torso. HATS approximates the frequency response of the human ear.

The acoustic signals were digitised (Digidesign ProTools HD192 A-D convertor) and captured as 24 bit, 48kHz AIFF lossless files using Pro Tools. Recordings were calibrated (Brüel & Kjær DP 4231) so the singer's absolute sound pressure was known at these microphones.

We were able to achieve a high quality recording environment that produced audio

files that emulated the experience of the listener (Fletcher & Munsen, 1933).

Acoustic Analysis

Initial acoustic analyses of the audio files were performed in PRAAT (Boersma & Weenink, 2006). Measures of mean and maximum sound pressure level (SPL) were performed on each song file. Long-term average spectra analyses (performance LTAS) were performed on each song file to describe the spectra of the singers' entire performance. Data were copied into Excel for further calculation.

Conventional single number ratings of singing power ratio (SPR) (Omori et al., 1996), energy ratio (ER) (Thorpe et al., 2001) (which have been used to characterise the extent of singer's formant and hence vocal projection) and Alpha (α-1 and α-2) were performed on these performance LTAS. SPR measures the difference between the highest peaks in the 0-2 kHz and 2-4 kHz regions (Omori et al., 1996). ER (Thorpe et al., 2001), α-1 and α-2 (Sundberg & Nordenberg, 2006) measure the difference between the spectral energy below the curves between 0-2 and 2-4, 0-1 and 1-6 and 0-2 and 2-6kHz respectively.

For all LTAS measures, the direction of all calculations was high frequency (HF) minus low frequency (LF). The values typically increase from negative to less negative and if HF > LF the measure becomes positive.

Analysis

Order effects. The singing auditions occurred over 2.5 days in five sessions (morning and afternoon). Judge rankings of singers were assessed against the order in which the singers auditioned in order to identify any effects of day (day1, 2, 3) or time (morning/afternoon).

Perceptual and acoustic relationships. Relationships between acoustic parameters and perceptual ratings of voices awarded a place at major and minor level and voices not accepted into the course were examined to determine the degree of concordance between the acoustic parameters and their group assignment.

Statistical analysis. Day and time differences were tested using Fisher's Exact Test. Data were subjected to analysis using the general linear model (GLM) with planned contrasts with audition outcome and sex as the fixed factors. In the first analysis main effects for audition outcome (major/minor/not accepted) and sex and interaction effects (contrasts) were calculated for each dependent measure (α-1, α-2, SPR, ER, mean and maximum SPL).

Because this was an exploratory study, a number of sub analyses (eg accepted/not accepted; sopranos only) were also conducted to clarify emergent relationships between factors and dependent variables.

Results

Twenty four singers were offered a place as voice majors (n=12) and voice minors (n=12). The performances of these 24 top ranked singers were compared with performances of the next 23 ranked singers who were unsuccessful at audition and not offered a place to study voice.

Effects of day and time of day

There was no statistically significant effect of day or time of day on the audition results (Fisher's Exact Test = 3.472, p=0.54).

Long term average spectra

Figures 1 and 2 present LTAS performed on selected samples for females and males. Each figure shows LTAS from two candidates accepted into the course (ranked highest and lowest within the group) and two candidates not accepted into the course. Below each LTAS exemplar is the corresponding assessment from the audition panel including the singer's overall perceptual rating (%), group designation (to major, minor or not accepted) and voice type. The panel's comments illustrate their response to vocal and performance features that shaped their judgments (e.g. vocal quality, technique and overall musical performance). From the adjudicators' comments, acoustic cues appeared to be a critical influence on outcome and it was hypothesized that LTAS may

provide further clarification of the acoustic signal.

Group differences

Song 1 performance parameters for the accepted and not accepted groups are shown in Table 1, first as a single group and then divided by gender. There were 14 females accepted to the course (5 majors, 9 minors) and 10 males (7 majors, 3 minors) and they were matched by the next 23 ranked singers (20 females and 3 males) who were not accepted into the course.

Audition Result: Accepted vs. not accepted

The average differences between the accepted v not accepted groups for the 6 dependent variables showed significant differences between mean SPL, and all LTAS measures (α-1, α-2, ER and SPR). That is, for SPL measures accepted candidates' produced greater mean SPL than those not accepted. For LTAS measures, accepted candidates produced lower results for α-1, α-2, ER and SPR in each case, as expected. When the energy is focused in the 0-2 kHz region, it results in higher SPR results and there is typically less energy reinforcement in the 2-4 kHz region. A low SPR indicates a stronger energy peak >2 kHz.

Audition result by gender

The analysis was repeated for females and males separately. For females, there were statistically significant differences between the groups for mean SPL and LTAS measure α-1. Accepted females demonstrated greater mean SPL and lower results for α-1 (Figure 3a-b).

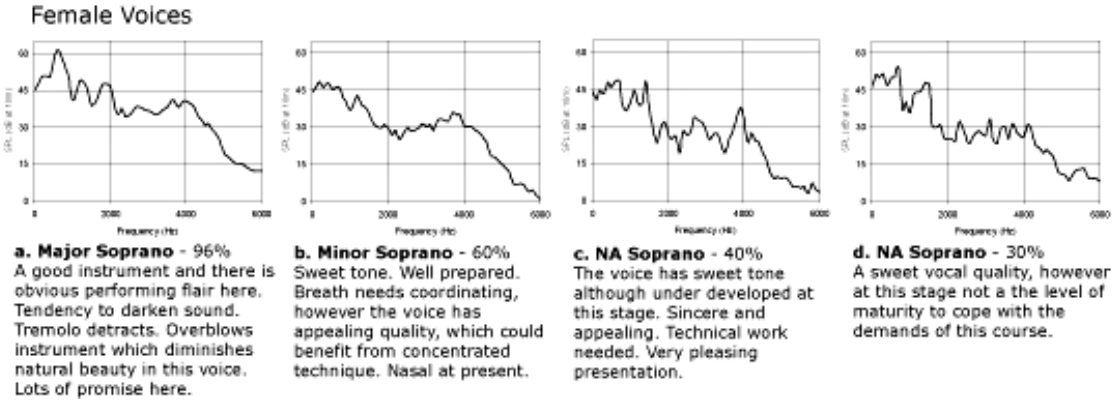

Figure 1a-d. Exemplars of LTAS for the highest and lowest rated female singers within each accepted/not accepted group with voice type, overall rating (%) and audition panel's comments.

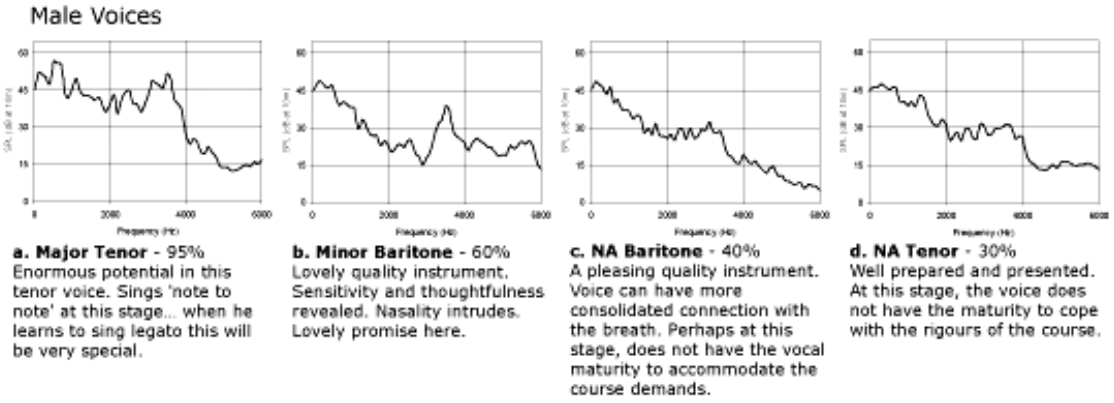

Figure 2a-d. Exemplars of LTAS for the highest and lowest rated female singers within each accepted/not accepted group with voice type, overall rating (%) and audition panel's comments.

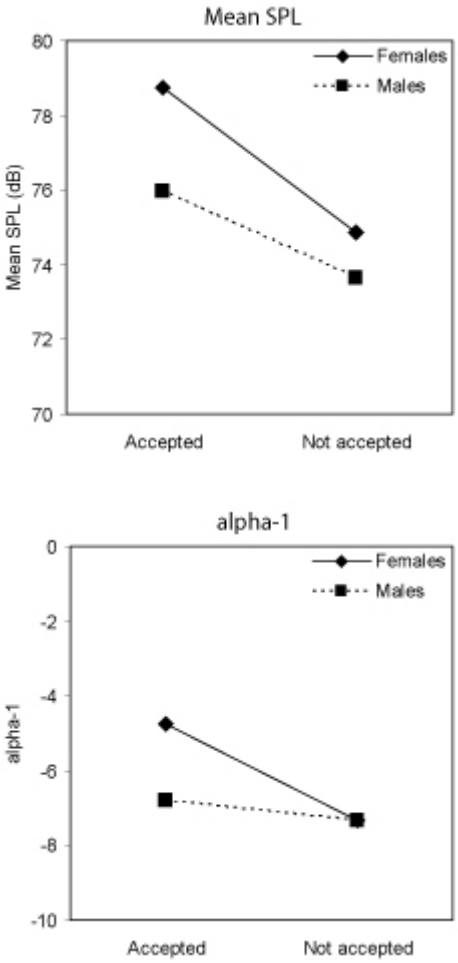

Figure 3a-b. Differences in estimated marginal means by gender for mean SPL and alpha-1 for the accepted and not accepted group.

	Accepted/Not accepted			
	All	Female	Male	Sop
LTAS measures				
alpha-1	0.027	0.006		0.019
ER	0.005			
alpha-2	0.005			
SPR	0.008			
SPL				
Max				
Mean	0.003	0.002		0.005

Table 1. Statistically significant findings for Accepted v Not accepted for sound pressure level (SPL) and LTAS measures.

For the males, there were no statistically significant differences for any measure between accepted and not accepted singers.

Sopranos

Sopranos represented the largest number in of a single voice type. The analysis was repeated for sopranos and there were statistically significant differences between groups for mean SPL and LTAS measure α-1, as for all female voices.

Major vs. Minor

A subgroup analysis of major and minor singers by sex showed that none of the variables were significantly different between female majors and minors or male majors and minors.

Discussion

Although the audition panel heard specific vocal qualities in the successful candidates at the three levels that differentiated them perceptually, conventional acoustic measures used in this study could only identify differences between the acoustic outputs of female singers who were accepted versus females not accepted to the degree. The measures used could not differentiate between majors and minors or between any of the three groups of male candidates.

This audition panel was consistent in its decision making and was not influenced by the day of the audition or the time of day (morning/afternoon) that the audition occurred. This panel functioned more effectively than other panels reported in the literature that showed listener inconsistencies and fatigue over the course of long adjudications and listening tests (Bergee & Platt, 2003; Elliott, 1996; Flores & Ginsburgh, 1996). This audition panel operated under different circumstances to those reported on previously. In addition, this panel arrived at their decision collectively, after discussion and this way of reaching a decision may have countered possible problems with judge fatigue and prevented extreme assessments that may occur during individual adjudications. (Smith, 2004)

This panel, through their comments, showed that they were primarily focused on vocal quality although there was also interested in overall presentation, style and personality of the singers. Voice majors were deemed to have a superior 'instrument' or a voice of 'high' or 'lovely' quality. In addition, the highest ranked majors were credited with 'potential', 'promise' or 'talent' for a future in the singing profession. Minor voices were described as having a 'pleasing' vocal quality and musical performance but were more likely to attract technical comments or criticisms. These candidates had an aptitude for singing that would benefit from further training. Unsuccessful voices from this audition showed (in some cases) a degree of 'sweet' or 'pleasing' core vocal quality but the panel focused on a lack of 'vocal maturity', in terms of quality and also of technique. Those candidates not accepted into the course were considered not to be vocally or technically equipped for the rigours of the degree.

Female singers accepted into the degree program (major/minor) produced a statistically greater mean SPL through their performance than those not accepted. There is evidence to suggest that classical training or advanced technique enables singers to produce a louder voice than untrained singers or increase their average SPL over time (Akerlund & Gramming, 1994; Mitchell & Kenny, 2004b; Murbe, Sundberg, Iwarsson, Friedmann, & Hofman, 1999).

In α-1, the accepted females produced a significantly smaller ratio between 0-1 and 1-6 than those not accepted. Ultimately, it proved important to consider two delimiting bandwidths in α (1 and 2 kHz). The spectral information in the energy band 1-2kHz is usually incorporated into ER or indeed α-2. We were initially concerned that 1kHz may not be as representative of the singing range for F0 as 2kHz, particularly for female voices. For females, α-2, ER and SPR showed no significant effects between groups. However, the energy band 1-2kHz may hold clues to the timbral differences between accepted and not accepted females.

Previous studies were able to rank voices using these spectral measures (Omori et al., 1996) (Barnes et al., 2004; Thorpe et al., 2001). More recently, SPR discriminated between untrained talented and non-talented female singers (Watts et al., 2006). These studies indicated that classical singing requires a degree of this characteristic energy to produce the classical sound quality.

Visual inspection of the LTAS curves showed a similarity in shape, with prominent peaks above 2 kHz particularly for males and accepted females. While LTAS shape has not been confirmed as a measure of quality when assessed against perceptual evaluation (Kenny & Mitchell, 2006; Vurma & Ross, 2000), the LTAS characteristics of successful voices in this study are striking. We expect professional classical and operatic singers and singers in training demonstrate a characteristic spectral energy above 2 kHz and the energy observed in these voices showed clearer individual peaks of energy. There is inconclusive longitudinal evidence to support the development of high range energy of voices in training (Mendes, Rothman, Sapienza, & Brown, 2003). Indeed, while high frequency energy may increase over time for singing students, it is not necessarily linked to an improvement in overall vocal quality (Vurma & Ross, 2000). It may be possible to quantify this common spectra as a general classical or 'operatic timbre' (Davidson & Da Costa Coimbra, 2001).

Although these LTAS were calibrated to be representative of known dB levels and are relative to each other, they do not account for the individual's overall sound level. Energy levels increase in high frequencies as SPL rises at a greater extent than in low frequencies. Further spectral measurements are required to investigate the relationship between LTAS bandwidths and their relationship to SPL for each singer to identify any group differences (Nordenberg & Sundberg, 2003; Sundberg & Nordenberg, 2006; White, 1998), particularly as mean SPL was a key discriminator in identifying accepted from not accepted candidates even when LTAS measures could not.

Acoustic results in this study showed some differentiation between those who were accepted into the course as majors and those who were not but they did not identify the

more subtle features of the very best voices from the good voices (major to minor). Although acoustic measures were never intended to provide vocal quality descriptors (Bartholomew, 1934), their extended use in the singing studio requires that they be validated, scrutinised and developed in authentic singing settings. Future research must refine these acoustic measurements in accordance with pedagogical judgments and listeners' assessments.

Acknowledgments. This project was funded by an ARC Discovery Grant [DP0558186] to Dianna Kenny, Helen Mitchell, Densil Cabrera and Michael Halliwell.

References

Akerlund, L., & Gramming, P. (1994). Average loudness level, mean fundamental frequency, and subglottal pressure: Comparison between female singers and nonsingers. *Journal of Voice, 8*(3), 263-270.

Barnes, J. J., Davis, P. J., Oates, J., & Chapman, J. (2004). The relationship between professional operatic soprano voice and high range spectral energy. *Journal of the Acoustical Society of America, 116*(1), 530–538.

Bartholomew, W. (1934). A physical definition of "good voice quality" in the male voice. *Journal of the Acoustical Society of America, 6*, 25-33.

Bergee, M. J., & Platt, M. C. (2003). Influence of selected variables on solo and small-ensemble festival ratings. *Journal of Research in Music Education, 51*(4), 342-353.

Boersma, P., & Weenink, D. (2006). Praat: Doing phonetics by computer (Version 4.5.04).

Borch, D. Z., & Sundberg, J. (2002). Spectral distribution of solo voice and accompaniment in pop music. *Logopedics Phoniatrics Vocology, 27*, 37-41.

Bunch, M., & Chapman, J. (2000). Taxonomy of singers used as subjects in scientific research. *Journal of Voice, 14*(3), 363-369.

Davidson, J. W., & Da Costa Coimbra, D. (2001). Investigating performance evaluation by assessors of singers in a music college setting. *Musicae Scientiae, 5*(1), 33-53.

Ekholm, E., Papagiannis, G. C., & Chagnon, F. P. (1998). Relating objective measurements to expert evaluation of voice quality in Western classical singing: Critical perceptual parameters. *Journal of Voice, 12*(2), 182-196.

Elliott, C. A. (1996). Race and Gender as Factors in Judgments of Musical Performance. *Bulletin of the Council for Research in Music Education*, 50-56.

Elliott, C. A., Schneider, M. C., & Zembower, C. M. (2000). Influence of the audition hour on selection to an all-state band. *Journal of Band Research, 35*(2), 20-31.

Fletcher, H., & Munsen, W. (1933). Loudness: Its definition, measurement, and calculation. *Journal of the Acoustical Society of America, 5*, 82-108.

Flores, R. G., & Ginsburgh, V. A. (1996). The Queen Elisabeth musical competition: How fair is the final ranking? *The Statistician, 45*(1), 97-104.

Geringer, J. M., & Madsen, C. K. (1998). Musicians' ratings of good versus bad vocal and string performances. *Journal of Research in Music Education, 46*(4), 522-534.

Hollien, H. (1993). That golden voice--talent or training? *Journal of Voice, 7*(3), 195-205.

Kenny, D. T., & Mitchell, H. F. (2006). Acoustic and perceptual appraisal of vocal gestures in the female classical voice. *Journal of Voice, 20*(1), 55-70.

Legge, A. (2001). *The Art of Auditioning: A handbook for singers, accompanists and coaches*. London: Peters.

Löfqvist, A. (1986). The long-time-average spectrum as a tool in voice research. *Journal of Phonetics, 14*, 471-475.

Löfqvist, A., & Mandersson, B. (1987). Long-time average spectrum of speech and voice analysis. *Folia Phoniatrica et Logopedica, 39*, 221-229.

Lundy, D. S., Roy, S., Casiano, R. R., Xue, J. W., & Evans, J. (2000). Acoustic analysis of the singing and speaking voice in singing students. *Journal of Voice, 14*(4), 490-493.

Mendes, A. P., Rothman, H. B., Sapienza, C., & Brown, W. S., Jr. (2003). Effects of vocal training on the acoustic parameters of the singing voice. *Journal of Voice, 17*(4), 529-543.

Miller, R. (1996). *The structure of singing: System and art in vocal technique*. New York: London: Schirmer Books.

Mitchell, H. F., & Kenny, D. T. (2004a). The effects of open throat technique on long term average spectra (LTAS) of female classical

voices. *Logopedics Phoniatrics Vocology*, *29*(3), 99-118.

Mitchell, H. F., & Kenny, D. T. (2004b). The impact of "open throat" technique on vibrato rate, extent and onset in classical singing. *Logopedics Phoniatrics Vocology*, *29*(4), 171-182.

Murbe, D., Sundberg, J., Iwarsson, J., Friedmann, P., & Hofman, G. (1999). Longitudinal study of solo singer education effects on maximum SPL and level in the singers' formant range. *Logopedics Phoniatrics Vocology*, *24*, 178-186.

Nair, G. (1999). *Voice tradition and technology: A state-of-the-art studio*. San Diego: Singular Publishing Group.

Nordenberg, M., & Sundberg, J. (2003). Effect on LTAS of vocal loudness variation. *Logopedics Phonatrics Vocology*, *29*(4), 183-191.

Omori, K., Kacker, A., Carroll, L. M., Riley, W. D., & Blaugrund, S. M. (1996). Singing power ratio: Quantitative evaluation of singing voice quality. *Journal of Voice*, *10*(3), 228-235.

Saunders, T. C., & Holahan, J. M. (1997). Criteria-specific rating scales in the evaluation of High School instrumental performance. *Journal of Research in Music Education*, *45*(2), 259-272.

Smith, B. P. (2004). Five judges' evaluation of audiotaped string performance in international competition. *Bulletin of the Council for Research in Music Education*(160), 61-69.

Stanley, M., Brooker, R., & Gilbert, R. (2002). Examiner perceptions of using criteria in music performance assessment. *Research Studies in Music Education*, *18*, 43-52.

Subotnik, R. (2003). Adolescent pathways to eminence in science: Lessons from the music conservatory. In P. Csermely & L. Lederman (Eds.), *Science Education. Talent Recruitment and Public Understanding* (pp. 295-301): IOS Press.

Subotnik, R. (2004). Transforming elite level musicans into professional artists: A view of the talent development process at the Juilliard School. In L. Shavinina & M. Ferrari (Eds.), *Beyond knowledge: Extracognitive aspects of developing high ability* (pp. 137-167). Mahwah, MJ, London: Erlbaum Associates.

Sundberg, J. (1974). Articulatory interpretation of the "singing formant". *Journal of the Acoustical Society of America*, *55*(4), 838-844.

Sundberg, J. (1988). *The science of the singing voice*. DesKalb, Illinois: Northern Illinois University Press.

Sundberg, J., & Nordenberg, M. (2006). Effects of vocal loudness variation on spectrum balance as reflected by the alpha measure of long-term-average-spectra of speech. *Journal of the Acoustical Society of America*, *120*(1), 453-457.

Thorpe, C., Cala, S., Chapman, J., & Davis, P. (2001). Patterns of breath support in projection of the singing voice. *Journal of Voice*, *15*(1), 86-104.

Vurma, A., & Ross, J. (2000). Priorities in voice training: Carrying power or tone quality. *Musicae scientiae*, *4*(1), 75-93.

Wapnick, J., Darrow, A. A., Kovacs, J., & Dalrymple, L. (1997). Effects of physical attractiveness on evaluation of vocal performance. *Journal of Research in Music Education*, *45*(3), 470-479.

Wapnick, J., & Ekholm, E. (1997). Expert consensus in solo voice performance evaluation. *Journal of Voice*, *11*(4), 429-436.

Wapnick, J., Flowers, P., Alegant, M., & Jasinskas, L. (1993). Consistency in piano performance evaluation. *Journal of Research in Music Education*, *41*(4), 282-292.

Wapnick, J., Mazza, J. K., & Darrow, A. A. (1998). Effects of performer attractiveness, stage behavior and dress on violin performance evaluation. *Journal of Research in Music Education*, *46*(4), 510-521.

Watts, C., Barnes-Burroughs, K., Andrianopoulos, M., & Carr, M. (2003). Potential factors related to untrained singing talent: A survey of singing pedagogues. *Journal of Voice*, *17*(3), 298-307.

Watts, C., Barnes-Burroughs, K., Estis, J., & Blanton, D. (2006). The singing power ratio as an objective measure of singing voice quality in untrained talented and nontalented singers. *Journal of Voice*, *20*(1), 82-88.

White, P. (1998). A study of the effects of vocal intensity variation on children's voices using long-term average spectrum (LTAS) analysis. *Logopedics Phoniatrics Vocology*, *23*, 111-120.