A Comparative Analysis of the Singer's Formant Cluster

by
Byron-Mahieu van der Linde

Dissertation presented for the partial fulfilment of the degree of *Master of Music in Choral Conducting, Faculty of Arts, at Stellenbosch University*

Supervisors: Mr. T. Herbst and Prof. H. J. Vermeulen

December 2013

DECLARATION

By submitting this dissertation I declare that the entirety of the work contained therein is my own original work, that I am the owner of the copyright thereof and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature

Byron-Mahieu van der Linde

Name in Full

Date

ACKNOWLEDGEMENTS

I have one acknowledgment always: thank you to my Friend and Brother, Jesus Christ, for increasing my capacity to excel and to achieve that which is so far beyond my reach. This thesis bears testament to this fact.

I would first, and foremostly, thank my loving parents Mr. Nico & Mrs Lyzette van der Linde. You always believed I was capable of more than I thought. Thank you for your undying love and assistance on every level; but, most importantly for covering me in the most meaningful way possible. I am a most privileged son.

Thank you to my awesome grandfather Dr. Ben Nel who, alongside – and with – my late grandmother, Mrs. Constance Nel, has always been my biggest supporter and biggest fan. Thank you for sharing your unprecedented knowledge with me, ever prayerful for me and pointing me to my exceedingly great Reward.

Alanna Rebelo – my other supervisor: you have carried me through this project with understanding, encouragement, your serving heart and love. May all know that, were it not for you, this study would have imploded. I could not thank you enough…

Further gratitude is extended to the following people:

- Mr. Adel al Zogaiby, who spent many an hour with me contemplating the endless possibilities of acoustic research. Thank you for making time to elucidate the essence of the spectral domain behind the intimidating physics.
- Prof. André Nel, Mr. Dieter von Fintel, Mr. Adriaan de Villiers and Mr. Gerhardus du Preez for always encouraging, probing and channeling my thoughts productively.
- Mr. Gerhard Roux, for being available for consultation and making sound editing look like so much fun!
- Prof. Hendrik. J. Vermeulen, your highly positive outlook on this project; for investing your expertise and having so much understanding and patience for my situation.

- Mr. Theo Herbst, since our first meeting you have inspired me to not take things for granted – to knock at the door of the unknown. Thank you for believing that this study was worth it and for teaching me that there's so much more to music than meets the ear.
- Prof Martin Kidd, for not making me feel that I was completely ignorant of the significance of the world of statistics.
- Doctors Rudolf & Marinda de Beer, were it not for your motivation, I might never have reaped the enriching benefits of delving into the fascinating fields of choral music, as well acoustic science. Thank you for all the years of mentorship and facilitation that culminated in where I am today.
- Mrs. Chantel Swartz, for meeting me where I was at and trusting that my research was relevant.
- Zaan Bester for her editing assistance, valuable commentary and kind support.

I am very grateful to the trustees of the Ernst & Ethel Ericksen Trust, the Harry Crossley Foundation and the Molteno Bursary Fund, for allowing me to apply myself fully, throughout the course of this research project, through very generous financial support.

Lastly, and certainly not least, thank you to the eight singer-subjects that participated in this study, without your positive involvement, we might have never had our target group. Much thanks also to the two professionals that were willing to give of their time to act as auditors. Your input is indispensable.

ABSTRACT

It is widely accepted that the singer's formant cluster (F_s) – perceptual correlates being *twang* and *ring*, and pedagogically referred to as *head resonance* – is the defining trait of a classically trained voice. Research has shown that the spectral energy a singer harnesses in the F_s region can be measured quantitatively using spectral indicators Short-Term Energy Ratio (STER) and Singing Power Ratio (SPR). STER is a modified version of the standard measurement tool Energy Ratio (ER) that repudiates dependency on the Long-Term Average Spectrum (LTAS). Previous studies have shown that professional singers produce more F_s spectral energy when singing in ensemble mode than in solo mode; however for amateur singers, the opposite trend was noticed. Little empirical evidence in this regard is available concerning undergraduate vocal performance majors. This study was aimed at investigating the resonance tendencies of individuals from the latter target group, as evidenced when singing in two performance modes: ensemble and solo. Eight voice students (two per SATB voice part) were selected to participate. Subjects were recorded singing their parts individually, as well as in full ensemble. By mixing the solo recordings together, comparisons of the spectral content could be drawn between the solo and ensemble performance modes. Samples (n=4) were extracted from each piece for spectral analyses. STER and SPR means were highly proportional for both pieces. Results indicate that the singers produce significantly higher levels of spectral energy in the F_s region in ensemble mode than in solo mode for one piece ($p<0.05$), whereas findings for the other piece were insignificant. The findings of this study could inform the pedagogical approach to voice-training, and provides empirical bases for discussions about voice students' participation in ensemble ventures.

OPSOMMING

Dit word algemeen aanvaar dat die *singer's formant cluster* (F_s) – die perseptuele korrelate is die Engelse "*twang*" en "*ring*", en waarna daar in die pedagogie verwys word as *kopresonansie* – die bepalende eienskap is van 'n Klassiek-opgeleide stem. Navorsing dui daarop dat die spektrale energie wat 'n sanger in die F_s omgewing inspan kwantitatief gemeet kan word deur die gebruik van *Short-Term Energy Ratio* (STER) en *Singing Power Ratio* (SPR) as spektrale aanwysers. STER is 'n gewysigde weergawe van die standaard maatstaf vir energie in die F_s, naamlik *Energy Ratio* (ER), wat afhanklikheid van die *Long-Term Average Spectrum* (LTAS) verwerp. Vorige studies het getoon dat professionele sangers meer F_s energie produseer in ensemble konteks as in solo konteks, in teenstelling met amateur sangers waar die teenoorgestelde die norm is. Min empiriese data in hierdie verband is beskikbaar, m.b.t. voorgraadse uitvoerende sangstudente. Hierdie studie is daarop gemik om die tendense in resonansie by individue uit die laasgenoemde groep te ondersoek, soos dit blyk in die twee uitvoerende kontekste: ensemble en solo. Agt sangstudente (twee per SATB stemgroep) is geselekteer om aan die studie deel te neem. Die deelnemers het hul stempartye individueel en in volle ensemble gesing, en is by beide geleenthede opgeneem. Deur die solo-opnames te meng, kon vergelykings van die spektrale inhoud gemaak word tussen die solo en ensemble konteks. 'n Steekproef (n=4) is uit elke stuk onttrek vir spektrale analise. Die STER en SPR gemiddeldes was eweredig vir beide stukke. Resultate toon dat die sangers beduidend hoër vlakke van spektrale energie in die F_s omgewing produseer in ensemble konteks as in solo konteks vir een stuk ($p<0.05$), terwyl die bevindinge vir die tweede stuk nie beduidend was nie. Die bevindinge van hierdie studie kan belangrik wees vir die pedagogiese benadering tot stemopleiding, en lewer empiriese basis vir gesprekke oor die betrokkenheid van sangstudente in die ensemble bedryf.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

CI – Confidence Interval

EFR- Extended Frequency Reinforcement

ENS – Ensemble

ER – Energy Ratio

F_s – Singer's Formant Cluster

GM – Great Minds against Themselves Conspire (piece 1)

IB – Immortal Bach (piece 2)

LTAS – Long-Term Average Spectrum

SATB –Soprano, Alto, Tenor and Bass

SM – Soloist Mix

SNR – Signal-to-Noise Ratio

SPR – Singing Power Ratio

STDEV – Standard Deviation

STER – Short-Term Energy Ratio

SU – Stellenbosch University

VOT – Vocal Task

VT – Voice/Vocal Training

CHAPTER 1: INTRODUCTION

1.1. Background

A noticeable amount of disparity seems to exist among vocal and choral practitioners regarding the effect group/ensemble/choir singing has on the singing technique of individuals when they sing as soloists (Bragg, 2012:*abstract*). The same reasoning is inferred across all contexts of group singing, whether amateur choir, professional choir, vocal ensemble or community choir, and irrespective whether the group is singing *a cappella* or accompanied: singers participating in ensemble music making are likely to suffer an ill effect to their vocal technique

Undoubtedly, a myriad of variables can be isolated as influencing factors to the singing manner of an individual singing in solo context or group context. Some universally accepted factors – that have been researched empirically – are: blend and ensemble formation (Daugherty et al., 2012); vibrato (Bragg, ibid.); conducting gesture (Fuelberth, 2004; Manternach, 2011); and, spacing phenomena like the 'Self-to-Other Ratio' (Ternström & Karna, 2002:273/4).[1] However, when all these factors are isolated – which they can – one main aspect is further underscored as rendering a characteristic difference between the two modes: it is argued that the use of and prevalence of the singer's *head resonance* could lie at the core of the contention (Detwiler, 2008:4; Weiss et al., 2001).[2] Head resonance (Christy, 1977:49), 'forward placement' (Miller, 1993:72), 'ring' (McCoy, 2004:47) or 'twang' (a term more commonly used by American and Australian *vocologists*[3], e.g. Titze, 2001; Kenny & Mitchell, 2006) [4], is pedagogical jargon for the acoustic phenomenon properly referred to as the *singer's formant cluster* [5] (Sundberg, 1987:119). It is descriptive of slight resonatory vibrations in the head and facial areas as a manifestation of the proper use of the singer's formant cluster (Carlsson & Sundberg, 1992; Titze, ibid.).

[1] See, for example Brenda Fauls' (2008) dissertation titled *A Choral Conductor's Reference Guide to Acoustic Choral Music Measurement: 1885 to Present.* She elaborates on a host of relevant studies.
[2] This contention has taken centre-stage since the 1920s. Ford (2003) cites Witherspoon (1925) as exclaiming that "probably not even the question of breath has caused more dire confusion and uncertainty, not to speak of faulty emission of voice, than this comparatively new bugaboo, RESONANCE!".
[3] Weiss et al. (2001) explain that this term denotes several types of practitioners, including: voice scientists, clinicians, and trainers.
[4] This is a personal observation.
[5] See *Section 2.2* for a more concise explanation of this phenomenon.

This demarcates a central topic resonating[6] with me. I have been studying solo singing as a performance major for nearly a decade (during high school and tertiary study years) and have concurrently participated in choir and ensemble activities. As such, I have been presented with the contention surrounding the use of head resonance. To avoid digression along the lines of deeply entrenched 'studio-isms'[7], I am interested in *objectively* investigating the way a singer conducts himself/herself as a soloist singing in group context[8]. This is then the main premise of this study.

One of the biggest challenges presented is that the scope of this research is defined by acoustic parameters; it exits the domain of the humanities and social sciences and enters the realm of the physical sciences, as it involves musical acoustics and psychoacoustics, according to Lindsay's (1964) 'Wheel of Acoustics'.[9] I reflect here on Johann Mouton (2005: 138), who states that "the most distinctive feature of scientific [research] is that the scientist selects phenomena from…[everyday life]…and *makes these into objects of inquiry*". In the light of the fact that there exists so much uncertainty around the acoustic differences between 'singing *modes*',[10] I aim to follow scientific methods of *inquiry* to find clarity amid presumptuous debates as "the search for truth or truthful knowledge is the overriding goal…" (ibid.).[11] Concomitantly, the choice of research design and supported methodology was chosen carefully.[12]

[6] Pardon the pun.
[7] As Gwendolyn Detwiler (2008: 24) denotes.
[8] All musical ideals in this study are modelled on examples that would typify Western Art music. In terms of vocalisation, ideal models would be 'operatic singing' and 'classical singing'; thus, exclusive from styles that would be representative of Western Contemporary styles such as rock, 'pop', jazz etc. Similarly, when referring to 'ensemble', we only present research relevant to groups that perform Art music, or strive to reach musical goals epitomised in Art music. A central reason for stressing this point is that individual singers and ensembles are able to present music from the Art music canon, or repertoire, but in a style that is contrary to Art music ideals. Ternström and Kalin (2007) presented a case study of a barbershop ensemble. They intimated that these singers "are arguably the most acoustically aware of [all] ensemble singers", and it is commonly known that their intonation standards are exemplary. However, they seek a different type of resonance than described in this dissertation, by means of varied vowel production between individual singers. This is contrasted with vowel shaping unification goals in 'traditional' Art music ensemble singing modes – irrespective of repertoire choice. Other examples could be given, but it is not necessary as I am certain the point is made clearly.
[9] This research follows in the discipline of acoustic research, which is very much a science as it typifies the "science of sound, including its production" (Fauls, 2008: 101).
[10] A term propagated by the likes of Ekholm (2000), Reid et al. (2007) and Detwiler (2008:iv), which – if left undefined – could cause confusion. Thus, for the duration of this thesis the term 'mode' is analogous with 'context', 'style' or even 'genre', and not a type of technique, or phonation style. Hence, 'choral mode', 'ensemble mode' and 'solo mode' could be interchanged with 'choral context', 'ensemble genre' and 'solo style', vice versa, etc.
[11] Incidentally, it should be mentioned that the prospects of interdisciplinary studies between the fields of music practice and acoustic research has long been deemed as having potential mutual benefits. Fauls (2008: 36)

1.2. Aim and Rationale

I do not propose addressing a long entrenched debate on whether solo-trained singers should sing in ensembles, or choirs, or not. Debates along these lines are qualitative in nature. I share very similar sentiments with Ford, who refers to the near exact contention I have pointed out (2003):

> Such controversy often appears to have a life of its own, quite apart from scientific findings.

Ekholm (2000) states, categorically, that there are definite benefits for soloists to participate in choral endeavors. She lists "music reading, ear training…and performance practices" as several integral skills that are developed in this 'mode' of singing. Ekholm's sentiments still resonate with the likes of Reid et al. (2007).

A paradox surrounding the debate has, however, arisen by means of experimental procedures. Some researchers found that singers – both amateur and trained – tend to adapt their head resonance detrimentally when singing in choral mode (Goodwin, 1980; Rossing et al., 1986 & 1987). Furthermore, Ekholm (2000) and Ford (2003) showed, by means of scientific method, that auditors from a relatively broad range of musical backgrounds preferred a choral music sound that is 'less resonant'. Such findings could be used to justify the argument of those that are bent on *proscribing*[13] a singer's involvement in ensemble music practices. However, such a stance would be conflicting with theoretical principles, which dictate that "singers who are unable to negotiate the differences between choral mode and solo mode fail to benefit from the advantages that choral singing offers" (Detwiler, 2008: iv).

To confound matters, Letowski et al.'s (1988, cited in Fauls, 2008) study provided somewhat ambiguous findings that fit both sides of the argument. The authors found that trained singers singing in 'choral', or ensemble, mode harnessed less energy in the singer's formant cluster

quotes Knutson: "…musicians as a group know far too little about acoustics, and acousticians know even less about music". This quote stems from 1987; since then acousticians such as Johan Sundberg, Sten Ternström and Ingo Titze have contributed vastly to our understanding of the acoustic parameters of the singing voice. Yet, Knutson underscores a quintessential methodological approach that supports the reasoning behind choosing a scientific design for this research project. See Chapter 3: *Methods.*

12 See *Section 1.4*.

13 Note: NOT *prescribing*

region, whereas – under the same conditions – untrained singers produced vocal timbre characteristic of resonant voice – a perceptual correlate of the singer's formant cluster[14].

Reid et al.'s (ibid.) study was aimed at determining how the vocal performance of members of Opera Australia differed acoustically between solo and chorus modes. The researchers worked from the premise that some of these "artists stated that they are unsure whether they naturally vary their singing technique for solo versus choral roles". Findings proved to be unprecedented: results indicate clearly that the professional operatic singers produced more spectral energy in the higher end of their vocal spectra – around the region of the singer's formant cluster, or head resonance – when singing in chorus. The authors noted, crucially:

> Singing teachers' advice to their solo singing student not to perform in choral music due to difficulties with combining two conflicting vocal techniques does not seem to hold.

The critical contribution of the above mentioned study is the finding that professional singers produced more head resonance in chorus mode. It would be a farce to extrapolate from these results that all professional singers, notwithstanding amateur singers, would adapt their voice production along similar lines. However, Richard Miller (1993: 76) articulates that "in the well trained…voice the singer's formant cluster remains constant". In other words, for a professional singer, spectral energy displayed in the singer's formant cluster region for any voice task should remain consistent; implying that Reid et al.'s (ibid.) finding should be followed-up promptly through further empirical studies.

1.3. Research Question

Many studies have used similar approaches to that of Reid et al. (ibid.), including Rossing et al. (1986 & 1987), Letowski (1988, cited in Reid et al., ibid.) and Hunter et al. (2006). These authors focused their research on liturgical choirs, untrained and semi-trained singers, and untrained and trained singers, respectively. Yet, a critical group is left unspoken for: undergraduate voice majors. Detwiler attempted to provide for this group in her 2008 PhD dissertation entitled: *Solo Singing Technique & Choral Singing Technique in Undergraduate Vocal Performance Majors: A Pedagogical Discussion*. As you may observe, the

[14] Please refer to *Section 2.2.4* for a discussion on 'resonant voice'.

methodological approach followed is a 'pedagogical discussion' – no results were obtained through experimentation. And thus, it has not been shown conclusively how these singers adapt their head resonance during solo and choral/ensemble singing tasks in *one* juxtaposed study. I am especially interested in this target group as it is postulated (Mendes et al., 2003; Mürbe et al., 2007) that it is during this developmental phase that a singer acquires the ability to harnesses resonant voice consistently during any given phonatory task.

Against this background, the central research question for this study is whether undergraduate voice majors produce more energy in the singer's formant cluster region when singing in ensemble mode as compared to solo mode? From this the hypothesis for this study is framed as follows:

Undergraduate voice majors produce more energy in the singer's formant cluster region when singing in ensemble mode as compared to solo mode.

1.4. Research design

Mouton (2005: 57) sketches a typology for research design types, classifying the proposed study well within the arena of empirical study, as the main focus will be on the assimilation of numerical and – to an extent – statistical data, both of which are primary data. The data will be captured under laboratory conditions, which offer a high degree of control (ibid., pp. 145). A fitting framework for the choice of research design can thus be defined along four dimensions. Table 1 relates aspects of the proposed study to the framework suggested by Mouton (ibid., pp. 146):

Table 1.1 Proposed typology and framework correlated with designated research design.

Dimension	Mouton's framework	Aspects pertaining to proposed study
1	Empirical study	Existing data being referred to in the literature of the broader field of acoustic research – of which this study is representative – has no bearing on the rationale of this research project if it is qualitative by nature. In fact, when referring back to the background of this study it is clear that the intention, from the outset, is to add scientific backing to qualitative perceptions regarding singing modes.
2	Primary data	All stages of the research process will be experimental by

		nature, necessitating the yielding of data under specific conditions. Resultant data will be representative of a strict experimental design.
3	Nature of data	The data in mention is related to various acoustic measurements of the vocal spectrum. Frequency measurements and power spectrum data, for example, are quantifiable. Nothing more can be read into garnered data from a qualitative point of view.
4	Degree of control	All data will be observed and captured under rather strict laboratory conditions, with limiting factors or variables being isolated and controlled. Hence, the degree of control over experimental processes is quite high. Results can only be obtained, in this experiment, through the optimal functioning of all measuring and data-capturing equipment, with strong reference to experimental designs of other, similar experiments that have been conducted successfully before.[15]

After referring to all the design types put forward by Mouton (ibid., pp. 148-164), it is clear that the above tabulated framework fits a standard 'experimental design' type (ibid., pp. 155).[16] Mouton describes this design as related to studies that are "causal [studies] of a small number of cases under highly controlled conditions" (ibid., pp. 155)[17]. The participation of auditors in the experimental process[18] would imply elements of a 'participatory action research design' type (ibid., pp. 150). However, such elements would not subvert the overall experimental design.

1.5. Thesis structure

This thesis follows a traditional outline for masters' theses. Following the present introductory chapter, the second chapter serves as a composite literature study. Here a broad discussion on vocal spectra lays a firm foundation for delving into more central topics related to vocal resonance. Some essential definitions regarding spectral analytical methods and tools are also presented, so as to contextualize the reader during the methods chapter. Chapter 3 describes in detail all empirical procedures followed, including the recording session, subjects profiling, vocal tasks required, the auditing process and spectral analytical methods.

[15] Smith et al. (2004), Reid et al. (2007), Rossing et al. (1986 & 1987) and Ford (2003)
[16] See section *1.2* (i.e. here I cite Mouton, 2001: 138).
[17] In Chapter 3: *Methods,* it is explained that a small target group of only 8 individuals will be used for all experimental conditions. The smallest case unit will be one individual who will be following strict instructions before undertaking phonatory tasks. This method is very much in line with Mouton's reasoning (ibid.).
[18] See Chapter 3: *Methods.*

The fourth chapter presents all results garnered from the spectral analyses, as well as descriptive results including statistical relevancies. An objective discussion about the results of the spectral analyses is given in Chapter 5, and the results are compared to findings from similar studies that served as premise for this study; followed by a concluding section.

CHAPTER 2: LITERATURE STUDY

The purpose of this literature study is to explore the main topic of this research project: the singer's formant cluster. A sufficient discussion on the nature, function and means of quantifying the singer's formant cluster is necessitated. The central hypothesis of this research project requires investigating singer's formant cluster strength of singers. Hence, focus placed on vocal pedagogical or psychoacoustic aspects will not be excessive. This approach implies that, in accordance with the hypothesis, the literature being consulted will predominantly be quantitative in nature.

This literature review comprises three main sections: 2.1.) Vocal Spectrum; 2.2.) The Singer's Formant Cluster; and 2.3.) Spectral Analyses. The first section broadly defines the acoustical principles and formant ranges of the vocal spectrum. The discussion in section 2 follows a 'funnel' methodology[19] that gravitates toward defining the phenomenon of the singer's formant cluster. It serves as the pivotal acoustical feature of our scientific enquiry[20]. Section 3 gives an overview of several spectral indicators relevant to this research project. A short concluding section draws together all the relevant aspects of the singer's formant cluster as pertaining to this research project.

[19] As described by Hofstee (2010: 96).
[20] See *Section 1.4*.

2.1. Vocal Spectrum and Formant Ranges

2.1.1. The Vocal Spectrum

Humans actually perceive a note/tone/sound emanating from a vocal source, as a rather complex amalgam of many softer, linear sounds (Alldahl, 2008: 50). In fact, these linear sounds exist as a series of physical, sinusoidal waves and are often referred to as *partial frequencies,* as they "[carry] a partial characterization of the whole sound" (Loy, 2006:29). They are positioned in very specific mathematical proximity to one another relating to their frequency reading in Herz (Hz) (Alldahl, 2008: 51). The first partial, with the lowest frequency, is referred to as the *fundamental frequency* (F_0)[21]. Each partial has a given amplitude, which means that different partials vary in intensity and thus in their perceived 'loudness' (Loy, 2006:37). The human brain perceives the original, complex sound. It then proceeds to process it so that, as an extraction of the entire spectrum, we only 'hear' the F_0 as it always boasts the largest amplitude in the vocal spectrum. (Goodwin, 1980)

A synonym for partial is *harmonic*. The series of harmonics above the F_0 is known as the *harmonic series*, or *overtone series* (Levitin, 2008:42). For clarity's sake: the reader might be familiar with the term *overtones*, which can actually only be used as the collective noun for all the partial frequencies in the harmonic spectrum above the F_0 (Alldahl, ibid.). In a vocal tone the frequency readings of the partials – as one ascends the harmonic spectrum – would always increase as an integer multiple of the F_0 frequency reading. (Loy, 2006:29; Levitin, ibid.). As an example, if the note A4 (the fourth 'A' on a conventional piano), with a frequency of 440 Hz, were to be sung the first partial frequency after F_0 (i.e. 440 Hz) would have the value of 880 Hz (i.e. 2 x 440 Hz); thereby rendering the next partials 1320 Hz, 1760 Hz, 2200 Hz…etc. (Goodwin, ibid.)

Alldahl (2008:50) also explains that – due to the integer multiple relationship of the harmonic series – the partials in a harmonic spectrum succeed each other in a very specific intervallic series. The series evolves as follows from the F_0: 8ve, perfect 5th, perfect 4th, major 3rd, minor 3rd, minor 3rd, major 2nd...etc., as seen in Figure 1. If the harmonics of a tone are not integer multiples of the F_0, then the intervals rendered would not be as consonant. An example of such a series could be that the F_0 has a frequency of 200 Hz, but the next partial

[21] Hereafter 'fundamental frequency' will be abbreviated as such, aligned with Sundberg (2001) and Mürbe et al. (2007).

lies at 410 Hz, followed by 605 Hz and a continuous series of imprecise partial doublings. Although the increment is roughly 200 Hz, the readings are not exact multiples of the F_0 (i.e. 200 Hz, 400 Hz, 600 Hz…etc.). Such a series is an *inharmonic series*. These sounds are not always conducive to be implemented in the carrying of melody (Loy, ibid.).

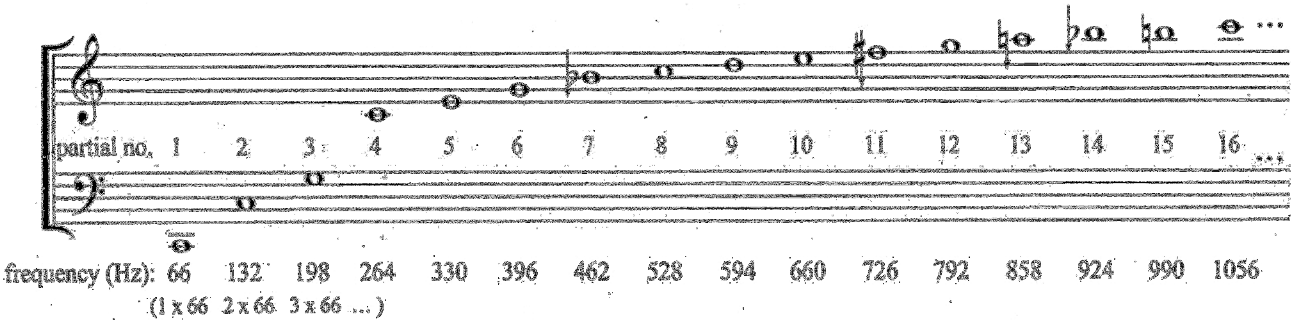


Fig. 2.1. First 16 partials in the harmonic series of F_0 = 66 Hz (C2) (from Alldahl, 2008:50).

Examples of instruments with inharmonic series' are bells and percussive instruments, including the piano[22] (Levitin, 2008:44). A typical aural implication of an inharmonic sound is that it is difficult to ascertain a definite pitch of a note played on the implicated instrument. For example, when a drum, bell or gong – or any cymbal – is struck[23] the inharmonic arrangement of the partial spectrum masks the pitch of the F_0; often the 'fundamental pitch' of the sound is perceived to morph as the sound degenerates over time. (Loy, 2006:31). With this said, Alldahl (ibid.) wittily remarks that "[to] sing 'as clear as a bell' could be regarded as an insult, since it can be tricky to determine which tone in the sound of a bell is the fundamental".

Succinctly stated, the human voice would always produce tones that are characterised, in the acoustic domain, by harmonic spectra. From afore mentioned one could deduce that the standard criteria for such a postulation would pivot on the fact that the phonated note would have a definite, perceived fundamental pitch – in spite of the dynamic, and articulatory, context and timbre of the sung note.

[22] The piano is classified as a percussive instrument due to the action of generating sound from a string by hitting it with a hammer. Sometimes the overtone series of a note played on the piano have partials in the spectrum that are not exact integer multiples of the F0 (Levitin, 2008:44).
[23] These instruments have overtone series that are severely non-uniform, in terms of the frequency distribution of the partials.

2.1.2. Formant Ranges

2.1.2.1. Occurrence and Nature of Formants

Some of the partials in the overtone series lie in specific frequency bandwidths that allow the partials to be amplified marginally (Sundberg, 2003). This is a characteristic acoustic phenomenon of all music instruments; with regards to the human voice, it relates to vocal resonance (McCoy, 2004:29). Amplification occurs due to the partials in mention being resonated as a function of certain favourable physiological conditions[24] (Sundberg, ibid). Such amplification areas are, therefore, resonant peaks within the harmonic spectrum of the vocal sound, referred to as 'formants' (Miller, 1993: 71-72). Formants do not produce any sympathetic partial frequencies; they are only able to enhance already present partials in a given sung tone (Ternström & Karna, 2002: 281).

The utility of formants strengthening and attenuating certain partial frequencies could further describe formants as 'acoustic filters'. This is a significant characteristic of the human sung sound. Manifested as gradual peaks (i.e. not spiked peaks) in the spectrum, formants are also visually wave-like. Correlatively formants are defined as covering a relatively defined bandwidth; having a central, peak frequency within the bandwidth; and having amplitude. (Ferguson et al., 2010)

There are typically about 5 formant ranges potentially available to a singer. The ones that are concerned with auditory perception are located between 200 Hz and 3600 Hz (Goodwin, 1980; Sundberg, 2001). These exist in the following, approximate spectral bandwidths (Miller, ibid: 74):

- 1^{st} Formant: 200 Hz – 600 Hz
- 2^{nd} Formant: 700 Hz – 1300 Hz
- 3^{rd} Formant: 1500 Hz – 1800 Hz
- 4^{th} and 5^{th} Formants: 2000 Hz – 4000 Hz (These formants are listed together as their ranges overlap greatly, and are often indistinguishable.)

It would seem that the over-arching range of all the formants is very limited (i.e. 200 Hz – 3600 Hz), relative to what is commonly known to be the maximum capacity of human

[24] These 'conditions' are not relevant to this study, which is the reasoning behind their being omitted. The physiological conditions of the 3^{rd}-5^{th} formant cluster *are,* however, relevant – to an extent – and will be elaborated on in *Section 2.2.5*.

auditory perception at 20 Hz – 20 kHz (Fauls, 2008: 12). However, with the first formant bandwidth starting at 200 Hz, it implies that the first partial above the F_0 has a frequency of 200 Hz. Therefore, F_0 will have to be 100 Hz (according to the multiple integer principle). It is very seldom that the F_0 of melodic tones of a vocal soloist would be lower than 100 Hz, or A2. Regarding the upper limit (3600 Hz) of the formant ranges, Ternström (2008) clarifies that there are several formant peaks in the harmonic spectrum above 3600 Hz. One formant clustering between 5 kHz – 10 kHz has a marginal influence on the timbre of sung tones in males, and is correlatively negligible as an important formant peak. Ternström (ibid.) also states that most partials, in sung tones, above 11 kHz are near inaudible; accordingly these also have no specific qualitative bearing on the perceived tone.

2.1.2.2 Certain Articulatory Functions of Formants

The articulatory potential of formant frequencies has enjoyed much attention in vocological related empirical studies (Fauls, 2008:10-11). It is a fascinating phenomenon that some harmonics can be amplified given their specific position within the vocal spectrum.

Detwiler (2008: 22; also corroborated by Fauls, 2008: 10; Ford, 2003 and McCoy, 2004:40/41) expounds on the function of each formant. The first two formants (and to an extent the third) are closely associated with the way in which vowels are shaped and their specific timbre. Most tones are projected through vowels; which means that formants 1 and 2 are basically ever-present in vocal production. The upper three formants are linked to the timbre of the vocal sound. Singers have a special interest to harness the resonance potential of these formants as a collective. Given the close proximity of formants 4 and 5, their clustering – together with the 3rd formant – allows a singer to generate immense spectral power through the amplification of partials that lie in this area. This resonance cluster is termed the singer's formant cluster.[25]

Another interesting articulatory function of formant ranges is the theory of 'formant tuning' (Sundberg et al., 2013). It is described as a strategy that is employed by singers (both trained in Western classical or Western contemporary styles) involving tuning the lower harmonics of their voiced sounds to the proximal center frequency of the first or second formant of the spectrum (Sundberg et al., ibid.). Carlsson and Sundberg (1992) explain that the potential

[25] Please refer to *Section 2.2* for a comprehensive discussion on this formant cluster.

benefit to singer could be that the spectral energy of the partials in the 1st and 2nd formants would be boosted, which would increase the projection of the overall vocal sound. However, concomitant to their empirical findings, the authors expressed doubt as to the artistic relevance of formant tuning in Western classical singing[26]. They distinguished as such, because there is evidence that formant tuning has a functional place in Tibetan chanting (a type of 'drone' singing) and a projection method, sometimes used by stage actors, referred to as the *Lessac call*. (Carlsson & Sundberg, ibid.)

Sundberg (2003) stated that he was concerned about the methodological approach some researchers (up to the time of his publication) were employing to ascertain validity for the use of formant tuning strategy in Western classical singing. The author (ibid.) listed several reasons why partial to formant matching seems problematic. Firstly he mentioned that some researchers espoused a method of isolating a partial garnered from a note sung by a singer on a recording. According to Sundberg the amplitude level of harmonics can differ by as much as 70% in different placements in a room relative to the recording source. This principle would have immense bearing on the measurement of the isolated partial. It seems as though the researchers did not carry knowledge of the acoustic parameters associated with the isolated audio samples that were used. Sundberg further held the opinion that any matching between harmonics and formant frequencies would be a chance occurrence, but that rigorous statistical analyses could make a better case for the applicability of the phenomenon. Lastly, the author mentioned that formant tuning would undermine generally accepted vowel shaping ideals of a Western classical singer.

All the while researchers have kept their sights on the apparent formant tuning strategies of singers as a major theme in research (Sundberg et al., 2013). Sundberg et al. (ibid.) recently completed a series of studies following stringent methodological approaches relating to the same theme. Both studies showed empirically that "professional singers did not show any tendency to tune F1 or F2 to a harmonic partial". It must be stated that this theory is plausible in accordance with acoustic laws. However, to date it is yet to be proven to hold fast under empirical conditions and demonstrate meaningful uses for Western classical singing. (Sundberg et al., ibid.)

---

[26] Please refer to *Section 2.2.6* for a discussion on projection through the use of the F_s.

2.2. The Singer's Formant Cluster (F_s)

2.2.1. Background

The significance of the singer's formant cluster (F_s)[27] as an essential characteristic of classical singing is demonstrable in the vast scope of scientific studies that have placed it at the center of empirical inquiry. Ford's (2003) publication and Fauls' (2008) dissertation paint a colourful picture of how this inquiry has manifested in a comprehensive body of relevant literature. It would be expedient for the reader to consult Fauls (ibid.) for a comprehensive overview of related studies throughout the 20th century. Here a brief overview is given with the aim to emphasise the historical significance of the F_s, corroborated by Weiss et al. (2001):

> Questions as to its [(the F_s)] necessity, requirement, frequency, and variability have been raised in professional voice journals and meetings for more than half a century.

It was as far back as 1863 that the first writing appeared about the aural observation of a slight 'tinkling' in the frequency range of singers during choral singing, at 2640-3168 vibrations per second by Helmholtz (Ford, ibid.). How this gentleman came about such a precise frequency range wasn't clear; however, it is strikingly similar to the F_s center frequency, as the reader would see later in this chapter. Further observations of resonance manifestations through the subsequent decades amounted to Fillebrown's notion, in 1911, that scientific based inquiry is necessitated to explain the observed phenomenal resonance (ibid.).

During the early 1920s the first major empirical findings on formant ranges were published by the likes of Paget, in 1922 (ibid.). In 1934 Bartholomew published the first findings on the F_s by noting a high amount of spectral power in the range of 2400 Hz – 3200 Hz (in Sundberg, 2003; in Weiss et al., ibid.). Winckel reported similar findings, over several studies ranging from 1952-1956, around the 3000 Hz area of sung tones (Sundberg, ibid.). Rzhevkin followed with his 1956 publication listing spectral peaks at 2200 Hz – 2800 Hz (Fauls, 2008:19; Ford, ibid.).

[27] Hereafter the 'singer's formant cluster' will be abbreviated as such, aligned with Weiss et al. (2001) and Fauls (2008:14).

By the 1960s the use of 'formant' (when referring to spectral peaks) as terminology among voice researchers – such as Delattre, in 1958 (Fauls, 2008:20; Ford, ibid.) Gunn, in 1960 (Ford, ibid.); Albert, in 1960 (ibid.); Arment, in 1960 (Fauls, 2008) and Vennard and Irwin, in 1966 (Ford, ibid.) – already became entrenched. Researchers seemed to have a fanatical urge to find this "holy grail" of vocal resonance, which is the F_s (Weiss et al., ibid.). Fauls (2008:21-86) goes on to document – at length – studies that emerged right up until the turn of the century, validating Weiss et al.'s (ibid.) statement earlier in this section. Research on divergent aspects related the F_s had spread as far afield as China – in Wang's studies from 1983-1986 (Sundberg, ibid.) – and India – in Sengupta's (1990) publication.

Up until the 1970s, however, the term 'singer's formant cluster' had no yet been used to describe the spectral peak in the high-end of vocal spectra. Authors referred to this along the lines of "an area of spectral reinforcement" (Weiss et al., ibid). Apparently some of Magill and Jacobson's (1978) contemporaries alluded to it along more technical lines. They called it the "2800 factor", given that peaks in spectral energy was commonly observed around the region of 2800 Hz. Among other studies in the early 1970s, by the likes of Fant (Fauls, 2008:20; Sundberg, ibid.), it was the 'pioneering'[28] research of Sundberg, starting in 1968 (Sundberg, 2003), that burgeoned more systematic approaches and methods by researchers to isolate and define the F_s (Ford, ibid.).

Sundberg (1974) somewhat inadvertently coined the term 'singing formant' in his publication titled *Articulatory Interpretation of the 'Singing Formant'*. It was also in this study that it was showed, for the first time, that the F_s was an embodiment of the clustering of the 4^{th}, 5^{th} and partly the 3^{rd} formants in the vocal spectrum (Sundberg, ibid.). The term was embraced readily among his contemporaries and ever since the term, adapted as 'singer's formant', became standard terminology. In recent years Sundberg (2003) proposed that it would be more technically accurate to substitute 'singer's formant' with 'singer's formant cluster'[29]. He presented a very valid explanation, substantiating his reasoning with technical acoustic examples. (Sundberg, ibid.) The author has since promulgated the revised term in his publications (e.g., Sundberg et al., 2012 & 2013), with some of his peers adopting it (e.g., Ternström, 2008; and Ferguson et al., 2010) and others keeping with the initial terminology

[28] As Reid et al. (2007) write. It is but one example of Sundberg being venerated among current researchers world-wide. He (i.e. Sundberg) continues to contribute to a legacy of paramount vocal acoustic research that he embarked on more than 40 years ago.
[29] Aligned with Sundberg (2003), the term 'singer's formant cluster' is used for the entirety of this study.

(e.g., Lister, 2009:10; Howard, 2009; Hunter et al., 2006.). It would be safe to say that the F_s has been well defined and the terminological differences have not brought about disparity among researchers and vocologists[30], nor allowed for confusion as to its quintessential definition.[31]

2.2.2. F_s Inherently Associated with Western Classical Singing

By this point the reader would have become acquainted with the fact that divergent aspects of the F_s have been highly researched. The F_s proved to be somewhat enigmatic in its physical manifestation. Accordingly, although much empirical findings have been published on it (see for example Sundberg's, 1987 volume), up until about a decade ago the F_s was still regarded by many as a theoretical feature (Weiss et al., 2001) of singing. Acoustic definitions for the theory seemed to become well defined, and yet researchers were still considering the F_s as a 'phenomenon' (Millhouse et al., 2002; Weiss et al., ibid.). Its prevalence could be noted in various styles of phonation across the world (Sengupta, 1990; Kovačić et al., 2003), however scientist are convinced that it is becoming increasingly apparent that the F_s finds its epitomical application in Western Art music, or classical singing (Millhouse et al., ibid.; Sundberg, 1987:118).

Even in styles very similar to Western Art music genres, such as the Chinese opera art, where singers sing in a fashion that resembles Western 'operatic' singing, it has been showed that singers don't produce the F_s at all (Sundberg et al., 2012). Some construed that the ability of a singer to harness the F_s ideally is a function of classical vocal training (Magill and Jacobson, 1978). Mendes et al. (2003) and Weiss et al. (ibid.), among others, have recently echoed this statement. Reports of high levels of F_s are mainly found with professional classical singers (Reid et al., 2007; Sundberg, 2001; Millhouse et al., ibid.; Ternström, 2008; Mendes et al. ibid.), hence the exclusive association with Western Art music.

2.2.3. F_s Center Frequency Variability

It has been shown previously that the F_s is a merging of the 3^{rd}, 4^{th} and 5^{th} formants into a rather narrow range within the vocal spectrum (McCoy, 2004:47).[32] As with singular

[30] This term is defined in *Section 1.1.*
[31] This is my opinion based on knowledge gained through reading of a large part of scientific publications that represent the corpus of F_s definitive literature.
[32] Please refer to *Section 2.1.2.*

formants, the F_s also has a specific 'center frequency' (Ferguson et al., 2010) that functions optimally in the 2.4 kHz – 3.6 kHz region (Millhouse et al., 2002; Ferguson et al., ibid.). Although encapsulating a general spectral bandwidth between 2 kHz – 4 kHz at its absolute extremes (Ford, 2003), Dmitriev and Kiselev showed in their 1979 that the F_s can vary for different voices, according to their classification (in Ferguson et al., ibid.). It also depends on the physiology of the vocal tract (Nair, 1999:46). For basses, the center frequency is generally from 2.3 kHz – 3 kHz, with 3 kHz – 3.8 kHz being the general region for tenors (Sundberg, 1987:119). Sundberg confirmed (2001) that the center frequency was indeed dependent on the pitch range of a specific voice type. He also set out to find which voice types produced the most amounts of spectral energy in the F_s region. Results showed that baritones generally produced the highest mean spectral levels, with basses and tenors being, on average, in the region of 3dB less than the mean of the baritones, and altos producing, on average, about 9dB less than this mean. The 4^{th} and 5^{th} formants in the sopranos were not approximated as one formant cluster, which means that they didn't really use the F_s. (Sundberg, ibid.)

The prevalence and significance of the F_s in soprano voices has been a main point of investigation (Weiss et al., 2001). Sundberg (ibid.) found that in soprano voices the 4^{th} and 5^{th} formant peaks are too distinct to be considered as a F_s. Weiss et al. (2001) showed that sopranos produced large amounts of spectral energy beyond the maximum range of the F_s, up to about 4.7 kHz. Findings along these lines have prompted some to infer that sopranos do not have the F_s; however, Sundberg (2003) states that this "issue is a matter of definition". At F_0 in the lower end of the soprano range, using the F_s is very probable, as would often be the case for mezzos. However, in the higher end, especially at low dynamics, spectral energy is distributed much more broadly than the confines of the F_s. Of this finding, Weiss et al. (ibid.) had the following to say:

> The broadband high harmonic reinforcement phenomenon typical of strongly projecting soprano voices could simply be called "extended frequency reinforcement" or EFR for want of a better term to distinguish this from the F_S.

The term EFR has not been used by other researchers to describe the spectral energy in the high end of the soprano's vocal spectrum. Some researchers have avoided complications by focusing research on the F_s on male subjects only (Ferguson et al., ibid.). However, others have continued to include sopranos alongside other voice groups in studies where the measurement of the F_s was a main outcome. One such study is that of Reid et al. (2007). These researchers did not hold the opinion that sopranos should be disqualified from tasks where spectral analysis was conducted to compare F_s readings in various singing contexts.

2.2.4. Resonant Voice

Having shown that F_s is a trait of Western art music, it is important to grasp with which particular aspect of classical singing it is associated. It has become specifically synonymous with good vocal resonance in Western classical singing, as it involves the selective resonance of partial frequencies.[33]

The significance of resonant voice is evidenced by the fact that it is the target outcome in various fields, including music, drama and voice pathology (Smith et al., 2004). Therefore, resonant voice is a target for both speech and singing purposes. Clinicians use the rehabilitating potential of resonant voice in speech therapy (Smith et al., ibid.) as it is produced easily and has relatively strong projecting power (Howard, 2009) due to the amplification of spectra content (Titze, 2001). Optimal formant frequency amplification in the speaking voice occurs around 3300 Hz, as shown in recent studies by Leino (2009), and Leino et al. (2010) (Leino, and Leino et al., cited in Sundberg, 2012). The projecting power of resonant voice makes it an attractive feature to pursue among actors, and especially classical singers, in order to combat vocal fatigue (Carlsson & Sundberg, 1992; Titze, ibid.). Miller (1993:76) asserts that the "singer's formant [cluster]…assures carrying power and proper resonance".

2.2.5. Physiology and Training Associated with F_s Production

In the light of the fact that it took so many decades for the concept of the F_s to be elucidated empirically (Weiss et al., 2001), there has been a fair amount of misinterpretation associated with it. Titze (2001) explains that "sensory perception of resonant voice involves head vibrations". Indeed, the term 'head resonance' has become a popular pedagogical correlate for resonant voice in classical singing, i.e. the F_s (Sundberg, 1987:119; Christy, 1977:49). This description is not problematic if one understands the true manifestation of resonance by means of the F_s. However, some have interpreted 'head vibrations' as the F_s being made

[33] Here the reader is reminded of the resonance potential of formants, as discussed in *Section 2.1.2.*

manifest in the head, which is untrue (Titze, ibid.). The author recognises part of the dilemma being that some confuse the proper 'ring' (Magill & Jacobson, 1978) quality of the F_s with the 'honky' quality of nasal murmur, which is contrary to ideal resonance (Titze, ibid.). Nasal murmur involves propagation of vocal energy through the nose; examples are consonants such as /n/, /m/ and /ŋ/. The resonant quality of F_s would still be present even with a constricted nasal passage, whereas nasal murmur cannot function without a clear nasal passage. (ibid.)

Amplification of partials actually occurs in, and as a function of the shape of, the vocal resonator (Howard, 2009; Nair, 1999:46). Smith et al. (2004.) add that the F_s functions optimally when there is a ratio of 6:1 between the volume of the pharynx and laryngeal, or *epilarynx* tube (Titze, ibid.), openings. Originally this ratio was hypothesised by Fant (Sundberg, 2003); but, Sundberg later (1974) defined it. Sundberg (2003) intimates that this is "somewhat of a sacred ratio", and is thus important in the understanding the propagating nature of the F_s. Detweiler (1994) challenged his pharynx-larynx ratio hypothesis through vigorous empirical studies, but Sundberg (ibid.) later rebutted her findings, showing that Detweiler's methods were scientifically unfounded.

The development of adequate resonance in the F_s region rests largely on the pre-disposition of a suitable vocal mechanism (Howard, ibid.). In other words, some individuals would be able to produce the F_s easier, due to a more conducive, natural, laryngeal and pharyngeal configuration. But, Howard (ibid.) goes on to mention that, generally, proper resonance is subject to the following, dynamic chain of physiological events: "the power source [breath control], the sound source during voiced or pitched sounds (the vibrating vocal folds) and the sound modifiers (the vocal tract)". Therefore, proper resonance involves this whole chain However, for classical singing it is the last step that is essential, the establishment of the 6:1 vocal tract ratio (Nair, 1999:47), which allows for the functioning of this area as an acoustic filter (Ferguson et al., 2010). Ferguson et al. (2010) explain that – when configured in the mentioned ratio – the pharynx and larynx amplify and attenuate the correct partials in the vocal spectrum, resulting in the characteristic spectrum peak in the region of the F_s.

In an early study done by Magill and Jacobsen (1978), the authors presented findings that pointed to a positive correlation between the amount of training singer's had with the spectral strength of the F_s. Training experience was not quantified absolutely, according to numbers of years, but rather just categories of experience (Magill and Jacobsen, ibid.). Even in a recent longitudinal study, by Mendes et al. (2003) the authors were not able to track a consistent development pattern of the F_s among voice students. The authors stated that after two years, no statistically relevant data could be garnered to see a marked difference between beginning of studies and the end of the 4th semester; however, the F_s was advanced noticeably by the end of the 4th year. This means that a specific development-plan for the development of the F_s can't be postulated. However, as with Magill and Jacobsen (ibid.), researchers must deduce that F_s is still a function of training, due to definite differences in spectral strengths between F_s of singers-in-training and professionals (Weiss et al., ibid.).

Being a function of vocal training, the development of the F_s – as was alluded to by Howard's (ibid.) 'dynamic chain' – goes hand in hand with learning how to manipulate the movements of vocal articulators, such as the jaw, lips and tongue, in a relaxed way; and establishment of healthy bodily posture (Bragg, 2012:27). Nair (ibid.) summarises that the production of spectral energy in the F_s region depends on "a well-trained voice".

2.2.6. Projection Function and Aural Significance of F_s

The quintessential importance of the F_s is its phenomenal utility to project the singing voice over live accompaniment without any electronic amplification (Omori et al., 1996; Lundy et al., 2000; McCoy, 2004:48). If one has had any exposure to live opera, or concert performances where a solo singer is juxtaposed against a large instrumental accompaniment, one would experience this function of the F_s in its archetypal form. That the singer is audible above the collective sound of an orchestra is a remarkable acoustic feat. (Nair, 1999:46)

The first time this enigma was empirically researched was in 1972, by Sundberg (Ford, 2003). He produced a phonograph superimposing the spectral slopes of an orchestral sound, normal speech and a sung tone, see Figure 2. From the spectrograph, one can clearly see that the orchestra's spectrum makes a prominent power spike around 400 Hz – 450 Hz along with the sung tone; however, the orchestra's spectrum makes a dramatic, uniform downward curve after this. Notice that that of the singer makes several more spikes, with the more prominent one being in the region of 2 kHz – 3.5 kHz. Last mentioned spike represents the F_s. At its highest point the harmonic content of the singer, in the region of the F_s, is about 15dB louder than that of the orchestra. According to acoustic principles this 15dB difference makes the spectral energy in this region of the vocal spectrum 2.5 times louder than that of the orchestra, considering that an increase of 10dB doubles volume (Detwiler, 2008:18).

Sundberg's discovery proved that it is the F_s that provides the singer the capacity to project his, or her, voice over the sound of an orchestra (Sundberg, 2001), since it occupies a niche spectral region that fills a 'hole' in the spectrum of an orchestra (Weiss et al., 2001; Sundberg, ibid.; Ferguson et al., 2010). This principle has been used by vocologists, universally, to substantiate the development of the F_s and its essential role in aiding vocal projection in Western Art singing (Nair, 1999:47; McCoy, 2004:49).

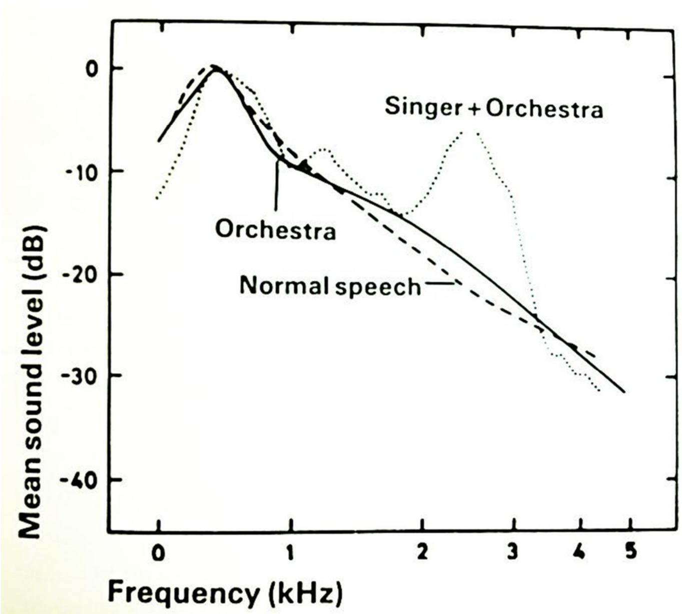


Fig. 2.2. Spectrograph of the power spectra of an orchestra, speech and classical singing (Sundberg, 1987:123). The solid line represents the orchestra, the dotted line the F_s and the dashed line speech. The F_s peak of the singer rises high above the downward sloping spectra of the orchestra and speech.

Sundberg (1987:123) notes that the F_s makes the projection of a professional singer's vocal production more directional. He explains that frequencies produced in the low range of the vocal spectrum radiate in an omnidirectional way from the singer. However, higher frequencies – especially those concentrated in the region of the F_s – project more acutely at right-angles to the sides and toward the front of the singer. By using the F_s, the singer ensures that the maximum amount of spectral energy is diverted straight toward the audience. (Sundberg, ibid.)

The directional projection function of the F_s subverts substantiation for the use of formant tuning as an alternative projecting technique[34]. Formant tuning would amplify partials around

[34] Please see *Section 2.1.2* for a discussion on formant tuning.

the first two formants, as they lie at relatively low ranges in the spectrum. Although this would make the sound seem 'bigger' in the vicinity of the singer, due to Sundberg's explanation of omnidirectional radiation (ibid.), the sound would not project very well.

Detwiler (2008:16) explains an acoustic principle that is a succinct application of resonant voice quality. It relates to the harmonic make-up of a vocal tone. She explains that research has shown that a tone displaying, or evidencing, a 'rich' array of harmonics (i.e. harmonic partials that are amplified by formant regions)[35] would be perceived by the listener to be louder than a tone sung at the same volume, but with less harmonic partials being resonated. This could appear to be an obvious fact, but note that the F_0 of both tones are equal in volume (Detwiler, ibid.). It is further substantiation for employing the F_s as a means to curb vocal fatigue.

From a pedagogical point of view, it would be expedient for a singer in training to understand that harnessing the use of the F_s should not be perceived as a wholly foreign task. The frequency of the F_s, with the +- 1.2 kHz bandwidth that it envelopes, lies at nearly the exact same bandwidth in the vocal spectrum as the 3rd and 4th formants of spoken tones. In healthy speech the partials around the F_0 would be amplified in conjunction with the partials in the 3rd and 4th formants. This is conducive to resonant voice in speech, as described in *Section 2.2.4*. In very much the same way, during a healthy classical vocal tone, the partials approximating the F_0 are amplified together with those in the F_s region. Figure 3 presents this spectral similarity between classical singing and healthy speech. One would note the similar amplification strategies of both types of vocalisation. The respective scenarios signify the characteristic timbre of speech and classical singing. Given the high similarity, it shows that classical singers are really pursuing a sound that is not foreign, but rather founded in the unique timbre of the human voice. (Sundberg, 1987:118)

[35] See S*ection 2.1.2* for a discussion on formants.

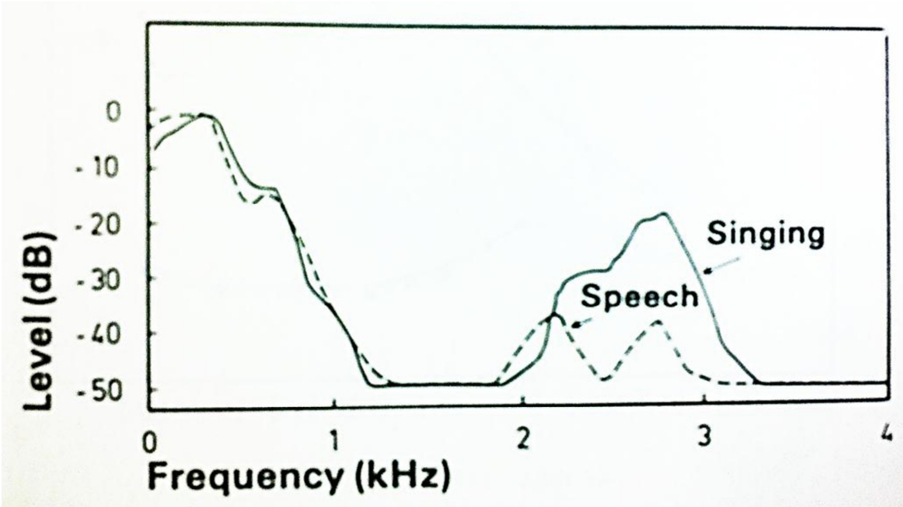


Fig. 2.3. Spectrum plot of spectral energy against frequency of spoken (dashed line) and sung (solid line) phonation (Sundberg, 1987:118). One notices similarity in the topography of the two spectra, only with clearly identifiable gain in spectral energy in the F_s region from about 2.4 kHz – 3.2 kHz during singing. The irregular peak shape of the F_s represents the clustering of several formants.

A particularly fascinating trademark of the center frequency of the F_s is that it bizarrely approximates the most sensitive part of human hearing, at around 3 kHz – 4 kHz (Weiss et al., ibid.; Hunter et al., 2006; Barnes et al., 2004). It seems more than coincidence, as Ternström (2008) explains. He shows in Figure 3 that the F_s not only attenuates at 4 kHz, but general spectral energy in the sung spectrum between 4 kHz – 5 kHz makes a drastic, characteristic dip; spectral energy between 5 kHz and 20 kHz generally stays 40dB – 50 dB lower than levels at the F_s (Ternström, ibid.). So, from the audience's perspective, the vocal sound of the singer would not only be projected straight toward them (Sundberg, 1987:123), audible over the accompaniment (Thorpe et al., 2001.); but, as discussed here, it would also excite the most excitable part of their hearing (Millhouse et al., 2002). Hunter et al. (ibid.) mention that studies have shown that the development of the F_s holds further aural incentive as professional singers use at least 7% more of their hearing range than amateur singers, and incrementally more than non-singers. The increase in the hearing range is directly correlated with the development and use of the F_s. As classical singers develop the use of the F_s as part of their vocal technique over time, their aural facility, therefore, becomes especially advanced. (Hunter et al., ibid.)

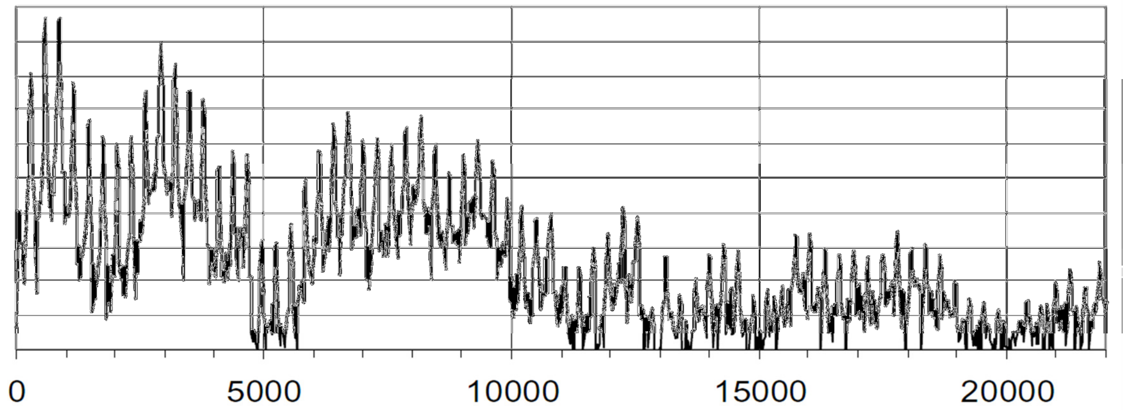


Fig. 2.4. A plot of spectral energy (in increments of 10dB) against frequency (in Hz), of F_0 = 291 Hz; sung by a male singer on the /a:/ vowel (Ternström, 2008). Note the relative spectral peak in the F_s bandwidth from about 2 kHz – 3.5 kHz, in the sensitive part of human hearing. It is followed by a drastic drop in spectral energy between 4 kHz – 5 kHz, hitting an almost zero at 5 kHz.

2.2.7. Qualitative Function of F_s

More effective resonance in the F_s region also has an important *qualitative* bearing on the singing voice. A resonant singing voice is described as having a characteristic 'ring' (Nair, 1999:47; McCoy, 2004:47)[36], 'ping' (Miller, 1993:72) or 'twang' (Titze, 2001; Thorpe et al., 2001) to it. These are actually a perceptual descriptions of harmonic content in the F_s region being amplified; making the produced sound more vibrant (Christy, ibid.; Titze, ibid.). The F_s is commonly referred to as *head resonance* (Christy, 1977:49), or *forward placement* (Miller, ibid.) in the singing studio.[37]

The term 'twang' is almost a paradoxical analogy. I, for one, had to stop myself from interpreting this description, of a type of pleasant vocal quality, as analogous with the poignant expression on my face after biting into a lemon: 'sour', or 'tangy'. But, Mesaros and Astola, (2005) elucidate what a 'twangy' sound could entail spectrally; which clarifies Titze's (ibid.) and Thorpe et al.'s (ibid.) analogy. Mesaros and Astola, (ibid.) state that "harmonics 7, 11, 13 [and] 14 have a negative influence on timbre, making the sound harsh, [whereas]…large amplitude of harmonics 2 to 6 gives a warm sound".

[36] That the page numbers for Nair and McCoy are identical is coincidental, and not an editing oversight.
[37] We will continue to use the term 'singer's formant cluster', as the empirical part of this study pivots on acoustic analysis, and not on a pedagogical discussion.

The reader is reminded that, according to Alldahl (2008:50), partials 2 through 6 create more consonant intervals, including the 8^{ve}, 5^{th}, 4^{th}, major-and minor-3^{rds}, relative to the F_0. Partials 7, 11, 13 and 14 represent more dissonant intervals with the F_0 as the comparative 'pitch melody' of the partials slope toward a plateau, refer back to Figure 2.1. Note, in this figure, that these partials create the following intervals relative to the F_0: flat minor-seventh, flat augmented-fourth, flat octave and flat minor-second. In this light, Mesaros and Astola's (ibid.) use of the terms 'negative influence', 'harsh', and 'warm' equates to dissonance and consonance, respectively. They do not imply, qualitatively, that the overall sound is 'beautiful' or 'unpleasant'.

Bearing this in mind, when extrapolating this principle to the functioning of the F_s, it implies that partials 7, 11, 13 and 14 would be amplified if they lie within the F_s bandwidth of optimal resonance at 2.4 kHz – 3.6 kHz (Millhouse et al., 2002). If these partials, as Mesaros and Astola's (ibid.) describe, imbue a 'harsh' quality to the tone, the 'harshness' would be increased due to the amplification potential of the F_s. However, only a specific region of fundamental frequencies would be influenced by this factor. By applying simple algebra one would find that the lowest F_0 that could have one of these dissonant partials in the F_s region would be F3, with a frequency of 174.61 Hz (i.e. 174.61[F_0 reading] x 14 [highest dissonant partial number] = 2444.54 Hz). Conversely, the highest F_0 with such a partial in the F_s region would be B4, at 493.88 Hz (i.e. 493.88 x 7 [lowest dissonant partial number] = 3457.16 Hz).[38] Hence, fundamental frequencies between F3 – B4 could have dissonant partials (7, 11, 13 & 14) that lie in the F_s region. To the F_s resonance, these notes could be perceived by the listener as having a slight dissonance in the harmonic make-up, which could be described as infusing a certain amount of 'twang' into the vocal tone.

Ford (2003) writes that Bartholomew (mentioned previously, in section 2.2.1.) was the first to allude to the qualitative function of the F_s, stating that Bartholomew (1934, cited in Ford, ibid.) found that the "upper resonance in the singer's formant range is one of the components that is preferred for a solo quality". In more current vocal acoustic literature this attribute is commonly accepted (Watts et al., ibid.). Lundy et al. (2000) write that the "classically trained singer strives to develop a singer's formant [cluster] that enhances the overall richness and

[38] The frequency values for music-alphabetical correlates F3 and B4 were garnered from the 'Piano Pitch ~ Hertz Chart' in Fauls (2008:161).

ringing in his or her voice”. Mendes et al., (2003) adds that the F_s imbues “clarity” and “timbral differentiation” to the vocal sound. This ‘differentiation’ is what is said to be a distinguishing factor between the vocal sound of trained and untrained classical singers (Nair, ibid.; Howard, 2009).

The distinguishing function of the F_s served as the main premise of Reid et al.’s (2007) study. Their ground-breaking study rested on the following hypothesis: “to determine the acoustic differences, if any, between opera chorus members performing in the chorus versus solo singing mode.” The authors found that, when singing in solo contexts, the opera chorus members inadvertently produced less spectral energy in the F_s region than during singing tasks in chorus contexts. Up until the time of this study, it was nearly universally accepted (within the realm of Western Art singing) that the opposite would be the case. (Reid et al.’s, ibid.) In this regard, their findings progressively contrasted those of Goodwin (1980); Rossing et al. (1986 & 1987) and Letowski et al. (1988, cited in Fauls, 2008).

Reid *et al*. (ibid.) presented an interesting corollary to this distinguishing function of the F_s. They proposed that the F_s can be employed as an unequivocal discriminating factor when adjudicating between various voices. This would present an objective way of making qualitative distinction between not only amateur and professional singers, but especially between professional singers. Thereby, as the authors reason:

> Indeed, it may be that the use of a ‘‘singer’s formant [cluster],’’ or at least increased energy in the singer’s formant [cluster], by an opera chorus…would result in a richer sound quality and larger dynamic range. If this was the case, opera chorus members could be [hired] according to their ability to sing with increased energy in the singer’s formant region.

Here, the two defining features of the F_s are underscored, once again: “richer sound quality” (qualitative significance) and “larger dynamic range” (projecting potential). What the authors are highlighting is the potential of incorporating digital measures of F_s, in the produced sound of a singer, as a means of developing an objective system to discriminate between voices

But, the suggestion of Reid et al. (ibid.) above actually echoes a similar, broader issue pertaining to the perceptual assessment of voice quality. This, too, has been the subject of debate and main premise of many studies for many years prior. Apparently, it has been very difficult for researchers to correlate perceptual preferences of voice quality among human auditors with demonstrated scientific findings related to energy in the F_s. Succinctly stated, researchers weren't able to definitively ascertain whether a human assessor would deem a voice to be 'good' even if instrumentation show a prominent F_s. (Kenny & Mitchell, 2006)

Kenny and Mitchell (ibid.) hypothesised the dilemma and brought it under specific empirical scrutiny. The authors asked 15 highly experience voice teachers to perceptually rate recordings of singers singing with good and sub-optimal vocal phonatory techniques. Ratings were cross-correlated with measured values of spectral correlates of the F_s within the recorded sound. A relevant finding was that some auditors rated voice quality of some recordings quite low, in terms of the perceived quality, yet the actual measures of energy in the F_s region were relatively high. It would appear that the authors held to a predilection for consistency among human auditors; essentially, contrasted to Reid et al.'s (ibid.) proposed digital system of voice quality assessment. Such a stance is paradoxical. On the one hand the authors (Kenny & Mitchell, ibid) present actual F_s readings that are high for a given audio sample; on the other hand, the auditor perceives the same sample as representing low quality. Incongruity appears to be more dependent on human error, in this regard. However, the authors generalised that, although levels of F_s can be obtained using several different spectral correlates, the measurement protocol needed to be refined to take into consideration more nuanced and subtle changes in vocal production.

Advances in in spectral correlates have, concomitantly, been made. Certain technical aspects, brought into question by Kenny and Mitchell (ibid.), were modified by Ferguson et al. (2010). Last mentioned authors implemented their modified spectral correlate of the F_s and showed that it could hold fast as an effective, objective tool to measure levels of F_s; and, if readings are interpreted thoroughly, it could be used successfully to quantify aspects of voice quality.

2.2.8. Resonant Voice is not Synonymous with Vibrato

Resonant voice in classical singing is not synonymous with vibrato; although it would not be incorrect to reason that a voice demonstrating vibrato also has a 'resonant quality'. And it would be beneficial to clarify briefly what vibrato entails, in order for one to make the clear distinction between it and the F_s. Mendes et al. (2003) clearly delineate vibrato and the F_s as two exclusive acoustic characteristics of the classical singing voice, that manifest independently of one another. Where the F_s is a formant in the vocal spectrum, vibrato is defined as such:

> A rhythmic modulation of the fundamental frequency that…is perceived as a pulsation of pitch, and it is usually accompanied by synchronous pulsations of loudness and timbre, giving "grace" to the tones.
>
> (Mendes et al., ibid.)

It takes time for a singer's voice to display consistent high levels of vibrato as vocal technique is developed (Mendes et al., ibid.). Mendes et al. (ibid.) undertook a longitudinal study of the effect that vocal training had on several features of the vocal sound of university students, over four years. Among others, he observed the development of vibrato and F_s. At the end of the fourth year of vocal training, results showed that – on average – the singers' vibrato rate didn't show a notable increase. The rate remained between 5.45Hz-5.65Hz from first year through fourth year. (Mendes et al., ibid.) This is nearly in line with the optimal rate of professional singers being around 6.7Hz (Bragg, 2012:29; Sundberg, 2012). However, the extent of amplitude of the vibrato of the singers was considerably lower than that of professional singers, with readings for the singers' vibrato at 2.2dB-3.28dB compared to 2dB-10dB of professionals' (Mendes et al., ibid.).

Vibrato in the sung tone could indicate that the sound 1) is anchored in good bodily posture, (Bragg, 2012:07) 2) is supported by proper breath control in modern research referred to by its corollary *appoggio*, (Bragg, ibid) – and importantly 3) emanates from a relaxed larynx (McKinney, cited in Bragg, 2012:05). In this sense, vibrato and the F_s are similar in that they are complimentary qualitative features of a well-produced classical singing tone. And, functionally, we learnt that that the F_s aids in projection[39]; whereas vibrato helps to stabilise a F_0 (Howard, ibid.); although, both could be harnessed beneficially for achieving good intonation (Sundberg, 1994) and combatting vocal fatigue (Bragg, ibid; Titze, 2001). Voice researchers are clear about the relationship between F_s and vibrato. Weiss et al. (2001) state that the presence of vibrato in the vocal sound of singers wouldn't affect their study on the F_s;

[39] See *Section 2.2.6*.

they did not put in measures in place to isolate vibrato "since it is a natural manifestation of trained voice production."

2.3. Spectral Analyses

In the previous sections we learnt how interest in understanding the F_s among scientists and vocologists increased since Bartholomew's publication in the 1930's (Ferguson et al., 2010). It is now in order to note what means are available to objectively inform a researcher about the 'presence' of the F_s in the spectrum of a singer.

Sound is a physical manifestation as it relies on the propagation of physical energy through a medium, which, in the context of singing, is air (Fauls, 2008:02; Nair, 1999:21, 23). Yet, these tiny waves – although abundant and often palpable[40] – are not visible. In order to analyse sound, the invisible physical, or analogue, wave property of sound needs to be converted into its digital correlate of many spectral components. Doing so gives a researcher and casual observer a visual means of experiencing sound. (McCoy, 2004:51)

This process of analysis is referred to as spectral analysis, acoustic analysis or voice analysis. Much like Form analysis in music theory, spectral analysis allows one to define the parameters, limits and other quantifiable properties of sound in an objective way. (McCoy, ibid.) Spectral analysis serves inter-disciplinarity in this research study. It is the communicative bridge between Acoustic Science and the musical practice, as "language is quite limited in its ability to describe sound" (Nair, 1999:207).

There are many methods (Hunter et al., 2006; Lamesch et al., 2012; Ferguson et al., ibid.), measuring tools/spectral indicators (Choi et al., 2012; Moon et al., 2012; Ternström et al, 2006, Lundy et al., 2000) and equipment (Fauls, 2008:138-151) available for spectral analytical purposes regarding voice science. However, when homing in on the F_s, some methods have been standardised through the implementation of very specific spectral indicating tools. A discussion of the most relevant spectral indicators used for spectral analysis in this research project is presented here.

[40] Ponder, for instance, the physical vibrations one experiences when standing in front of a subwoofer at high loudness levels; or the venue-shaking resonance of an extremely low not produced in the 32-foot pipe of an organ. Both these experiences leave one with the perception that one can *feel* the sound.

2.3.1. Definitions of Specific, Relevant Spectral Indicators

The *Long Term Average Spectrum* (LTAS) is an averaged display of several shorter spectra (Watts et al., 2006). A *spectrum* is the conversion of a waveform into its spectral components using the mathematical configuration *Fast Fourier Transform* (FFT) (McCoy, 2004:51); the end result of which is a graph of spectral energy (in dB) against frequency (in Hz) (Fauls, 2008:131). The LTAS ignores the time domain (Ferguson et al., ibid.) to plot "an overall impression of an entire excerpt by identifying certain consistent features contained in the sound over time" (Kenny & Mitchell, 2006). For vocal analytical purposes, the LTAS exhibits the formant characteristics of a sung tone or phrase (Ferguson et al., ibid.). Thereby, one could visually track the spectral contours of the formant ranges. This is useful for perceptual inferences regarding the musical input (Watts et al., ibid.).

Singing Power Ratio (SPR) is a spectral indicator, derived from the LTAS, that measures the difference between the highest energy peaks in the low (0 kHz – 2 kHz) and high (2 kHz – 4 kHz) bandwidths of the vocal spectrum (Ferguson et al., ibid.). Essentially, it measures the relative energy of the F_s (Reid et al., 2007; Fauls, 2008:129) and is therefore said to be "a quantitative measure of the resonant quality of the singing voice" (Watts et al., ibid.). Although it is termed as a difference, SPR is measured in Decibels, which essentially means it is a ratio between the peaks (Omori et al., 1996).

Energy Ratio (ER), also a spectral indicator derived from the LTAS, is measured in an analogous fashion to SPR. ER also focusses on the low and high bandwidths; however, it includes the overall energy in these bandwidths, and not just the highest spectral peaks (Ferguson et al., ibid). Thus, ER is a ratio of the difference between the total energy in the 0 kHz – 2 kHz and 2 kHz – 4 kHz regions (Kenny & Mitchell, ibid.). Both SPR, and ER, readings are attained from the LTAS – rendered with FFT – of a vowel sample, or 'vocalic segment' (Watts et al., ibid.; Barnes et al., 2004). They are also interpreted in the same way in that, for both indicators, low readings indicate high levels of spectral energy in the F_s region, whereas the opposite is true for higher measures of SPR and ER (Reid et al., ibid.). ER is usually calculated in conjunction with SPR as it reflects directly upon spectral changes in the Fs region and is also accepted as a tool for quantitative evaluation of the F_s (Ferguson et al., ibid.). ER and SPR measures are therefore spectral correlates of the F_s, in a similar way that 'head resonance' (Sundberg, 1987:119) is a pedagogical correlate, and 'ring' is a perceptual correlate (Nair, 1999:47).

Short-Term Energy Ratio (STER) is near identical to ER and is derived using a very similar method and calculation process. It was developed in an attempt to assimilate more accurate readings of distribution of spectral energy than from the traditional LTAS (Cabrera et al., 2011; Ferguson et al., ibid.). For standard ER, calculations are based on LTAS of the entire audio sample under observation, whether it be 1s or 60s long. STER differs from ER in that a sliding window of about 21ms in length is moved across the audio sample. It performs a FFT, followed by a measure of ER for that window, at a consistent interval of 10ms – thus there will be an overlap of 10ms for every reading. Individual ER readings for all the windows across the given sample are collated as time series data. These readings more accurately represent the distribution of spectral energy over the span of an audio sample. Time series data is averaged, rendering an average ER reading for the audio sample. (Ferguson et al., ibid.) To the date of compiling this thesis, STER has been successfully implemented for vocal acoustic analyses by Cabrera et al. (ibid.).

2.3.2. Implementation of Spectral Indicators SPR, ER and STER for Vocal Analytical Purposes

Omori et al. (1996.) were the first to use, and suggest the further use of, Singing Power Ratio (SPR) as a 'quantitative' measure of F_s. The authors postulated that singing students could benefit from the development of the F_s because of its ability to amplify the voice over orchestral accompaniment. Thus, they aimed at finding a tool to consistently monitor the development of the F_s. As the F_s functions in the extreme bandwidth of 2 kHz – 4 kHz, and the first two formants – functioning between 0.2 kHz – 2 kHz – are ever-present in sung tone, the authors reasoned that measuring the difference between the highest peaks in these bandwidths would provide an objective measure of the strength of the F_s. Results indicated that SPR reading for professional singers with more than 4 years training were considerably higher than for singing students. An obvious finding, but it showed that SPR could indeed be used to monitor training, and be adopted as a standard spectral indicator. Perceptual tests were also incorporated, using experienced vocal instructors as auditors. Perceptual results also showed that SPR could possibly be used as quantitative means to measure vocal quality. In their opinion they produced successful results.

Lundy et al. (2000) sought to replicate certain of Omori et al.'s (ibid.) findings; essentially testing the validity of SPR. Some results of Lundy et al.'s (ibid.) study contrasted to those of Omori et al. First mentioned authors were not convinced that SPR could readily be used as a quantitative measure of F_s, and stated that future studies using SPR are necessitated. However, they concluded that SPR holds great potential as a spectral indicator and aligned their opinion with that of Omori et al. (ibid.), stating that "SPR analysis may provide an objective tool for monitoring the student's progress [at developing resonance]".

Lundy et al.'s (ibid.) concluding statement seemed to have acted as a green light for researchers to start using SPR as a spectral analytical tool. Watts et al. (2006) constructed a study that was geared toward using SPR to determine whether untrained singers possessed talent. SPR measures would be used by the authors as quantitative correlates for voice quality. They were able to use SPR results to differentiate between two groups of singers, viz. talented and non-talented. Finally, the authors reasoned that singing talent could be dependent upon resonance in the F_s region, which influences perceived voice timbre; this could be measured successfully by SPR.

It was shown earlier that increased energy in the F_s region benefits projection (McCoy, 2004:48).[41] Given that quantitative means arose to evaluate the F_s, Thorpe et al. (2001) designed an investigation based on the hypothesis that certain breath control patterns underpinned vocal projection in the high frequency range of the F_s. The authors aimed at using a similar spectral indicator as SPR, however, they actually calculated the ratio of total energy in the low and high energy bandwidths. Using this spectral measure, results clearly delineated differences in spectral energy between singer participants, which allowed the investigators to make informed interpretations about the way breath control related to spectral projection. This was essentially the first time Energy Ratio was successfully used to measure vocal spectral energy differences, although the authors didn't actually refer to it as such. They merely provided calculations.

The term 'Energy Ratio' was first coined, in reference to Thorpe et al.'s (ibid.) study, by Barnes et al. (2004). These authors formalised the calculation of Energy Ratio and

[41] Please refer to *Section 2.2.6.*

abbreviated it as 'EnR', although afterwards it has always been abbreviated as 'ER' by their contemporaries (Ferguson et al., 2010). They used this indicator, alongside SPR, to gain more understanding regarding high frequency spectral characteristics of soprano voices. Results exhibited striking correlations between SPR and EnR, proving that they are directly proportional. Low readings in SPR and EnR – which indicated and ranked projecting power of singers – correlated very well with ranking of participants, in terms of their experience. It substantiated the belief that spectral indicators SPR and EnR are two-beans-in-a-pod and could be used effectively for measurement of the strength of the F_s in future studies.

Ever since, there has been little doubt that SPR and ER are the most conventional indicators of relative energy in the F_s (Ferguson et al., ibid; Kenny and Mitchell, 2006). However, from a qualitative perspective, researchers have been reluctant to use it as means to quantify voice quality, as Omori et al. (ibid.) initially suggested. Kenny and Mitchell (ibid.) published findings on research they carried out in much the same vein as Watts et al. (ibid.); however, their findings stood in stark contrast. Unlike Watts et al. (ibid.), Kenny and Mitchell (ibid.) provide findings that show a considerable discrepancy between SPR and ER rankings and auditors' perception of vocal quality. SPR and ER readings correlated significantly, as showed by Barnes et al. (ibid.) and the authors stated that these readings were calculated legitimately. However, auditors appeared to perceive 'optimal' or 'sub-optimal' vocal quality apart from high and low SPR and ER readings of F_s energy. The authors concluded that the discrepancy could lie in the fact that ER is measured from LTAS. Due to the fact that this averages several consistent spectral features over the given time, the authors argued that LTAS misrepresents the exact spectral content in the high and low bandwidths. Consequently, there would be dissociation between SPR and ER rankings and perceptual judgements.

Reid et al. (2007.), however, employed SPR and ER in their assessment of vocal power in the F_s, in a comparative study of opera singers singing in solo and chorus modes. They did not aim to correlate SPR and ER with perceptual ratings, but to use raw readings as absolute representation of the strength of the F_s in two different modes. This use of SPR and ER is corroborated with Kenny and Mitchell (ibid.), who stated that "both SPR and ER enable effective comparison between individual singers and between groups of singers". Reid et al.'s (ibid.) analyses produced results, resting on SPR and ER readings that led them to conclude that professional singers were able to produce more spectral power in the F_s

region during ensemble tasks than during solo vocal tasks. Findings asserted the validity for SPR and ER to be associated with F_s, quantitative assessment.

Considering Kenny and Mitchell's (ibid.) dissatisfaction with LTAS, Ferguson et al. (ibid.) looked at an alternative methodological approach in garnering ER to address the *problematique*. The authors developed the short-term energy ratio in response.[42] As an 'update' – in present-day terms – of ER, STER, alongside SPR, could continue to serve as the most accurate means to quantify F_s.

2.3.3. Relevance of Spectral Indicators in This Research Project

It would be expedient to use SPR and STER in any further spectral analytical studies pertaining to the F_s; given the efficacy of SPR – as shown in methodologically sound empirical studies by Watts et al. (2006), Kenny and Mitchell (1996) and Reid et al. (2007) – and STER – as shown by Ferguson et al. (2010) and Cabrera et al. (2011) – in providing quantitative information on the relative strength of spectral energy in the F_s region.

Kenny and Mitchell (ibid.) vouched for the effective use of SPR and ER in comparative studies between individuals and groups of singer, and STER is an enhanced version of ER (Ferguson et al., ibid.). Thus, as the main methodological approach necessitated by my hypothesis is to measure the comparative strength of the F_s between solo and ensemble modes, it seems clear that using spectral indicators SPR and STER would be most effective in garnering relevant spectral data.

2.3.4. Conclusion

The singer's formant cluster (F_s) is almost exclusively associated with classical singing; and is said to be the characterising trait of a professional singer's spectral output. Spectral energy in the F_s region of a professional singer is expected to stay consistent for various vocal tasks, where no detrimental adjustment to vocal technique is made. The literature revealed that comparisons of spectral output in the F_s professional singers singing in solo vs. ensemble modes have been carried out successfully. Even though it has been argued that sopranos do not have a F_s, not enough evidence has been given to disallow their participation in studies where a main focus is the measurement of the F_s. It is not clear at what point, in their

[42] Please see *Section 2.3.2* for the definition of STER as developed by Ferguson et al. (ibid.).

technical development phase, singing students are able to willingly control the consistent functioning of the F_s. However, it has been shown that by the end of the 4^{th} semester of voice training there could be marked improvement in a student's ability to produce spectral energy in the F_s region. The spectral indicators Short-Term Energy Ratio (STER) and Singing Power Ratio (SPR) have been shown to objectively measure the amount of energy any singer harnesses in the F_s region during Spectral Analysis. Therefore, a comparison of the amount of spectral energy in the F_s of a classical singing student – across all voice categories – singing in solo and ensemble can be made by means of Spectral Analysis using STER and SPR.

CHAPTER 3: METHODS

The over-arching framework that typifies the processes associated with this research project is an experimental design type. The project comprised several phases: gaining understanding of the nature of vocal acoustic data and empirical processes associated with spectral analyses through literature; garnering audio samples under experimental conditions; consulting auditors for selection of recorded tracks; crystallising data of relevant spectral indicators through spectral analyses of audio; and scrutinising findings by means of statistical analyses. Established empirical approaches were used as models. Some models were treated as exclusive, whereas others were more conducive to be adapted into more inclusive procedures. A composite description of the methods employed to pull divergent processes together is given here.

3.1. Target Group

3.1.1. Constraints

The pivotal element of this design is to build a study around undergraduate vocal majors; thus, involving the appropriate individuals for the project was crucial. Resonant voice is but one of many teaching outcomes; therefore I sought guidelines to determine constraints for an appropriate target group. Mürbe et al. (2007) investigated certain acoustic trends in the vocal development of 22 professional singers at the onset and completion of their 3 year professional vocal studies at a university. One relevant finding was that the singers applied more thorough pitch control during slow and legato vocal tasks (VOT). It is vital to minimise the effect of limiting factors in the present research, in so doing producing more accurate results accorded to *resonance* tendencies. Accordingly, I shall follow several similar methods to that employed by Mürbe et al. (ibid.):

1. Using slow tempi and legato singing as major points of instruction for VOT to maximise pitch accuracy – this is vital, given that all VOT will be done *a cappella.*
2. Selecting undergraduate singing students who are nearing the end of their undergraduate studies – as it was shown that they have harnessed certain acoustic/aural skills only to a *modest* extent and are therefore situated very well between the domains of amateur singers and professional singers.

In reference to the second point, Mendes et al.'s (2003) study –focusing on a student target group – showed similar outcomes to that of Mürbe et al.'s (ibid.). Students proved to produce inconsistent readings of several key factors related to the solo-trained voice. After 4 semesters of voice training (VT) these factors did not seem to have been developed sufficiently[43].

3.1.2. Singer-Subjects: Recruitment and Profiling

Eight singing students (here after referred to as 'subjects') were invited to participate in the study. All of these were registered for full-time music degree courses at Stellenbosch University at the time that the invitation was extended. Six subjects entered into their 6th semester of VT. Two subjects had entered into the 8th semester of VT. Ages ranged from 20.83-24.25, with the mean age being 22.47.

Two essential parameters underpinned this study:

1. The singer-subjects involved should all intend to follow a professional career as soloists in classical singing.
2. The combined voice classification of all eight subjects had to represent a four-part Soprano, Alto, Tenor, Bass (SATB), mixed choir

The eight subjects that were invited were perceived by the main researcher to be in line with both of the above mentioned criteria. However, confirmation by each subject was warranted. Accordingly, all subjects underwent verbal screening in terms of their respective voice classification and their personal goals for studying singing. Each singer confirmed their classification within four general categories: soprano, mezzo, tenor and bass/baritone. Two women classified themselves as sopranos; two women classified themselves as mezzos; two males classified themselves as tenors; and two men classified themselves as bass/baritones. All subjects stated that they intended to follow careers as professional classical singers.

In light of above mentioned, the target group comprises a full SATB choir (i.e. 2 x voices per voice group), according to Ford's (2003) method. Eckholm (2000) and Libeaux *et al.* (2007, cited in Fauls 2008: 84) were also of the opinion that 8 voices – with the above mentioned voice-group classification – constituted a legitimate choir.

[43] It firmly places our research target group between that of Rossing *et al.* (ibid.) – i.e. amateur choristers – and that of Reid *et al.* (ibid.) – i.e. professional singers. This specific phase of vocal development has been reasoned to be fitting as a target phase for our research – please refer to *Section 1.3*.

For ethical reasons, the subjects were also expected to give consent to the following factors:

1. they were to be recorded in solo mode and as part of a choir
2. all audio recordings could be anonymously scrutinised by auditors
3. all audio could undergo spectral analysis
4. personal details of all subjects, and any audio or video recordings associated with them, would be kept anonymous

All eight subjects gave their consent to the ethical factors mentioned above. The participation of the singer-subjects was cleared by the Departmental Ethics Screening Committee of the Music Department of Stellenbosch University (SU) and granted by the Research Ethics Committee of SU. At confirmation of all subjects' involvement, the group was sent a preliminary time-table with possible time-slots for solo and ensemble recording sessions in both venues. A suitable time-frame was confirmed among all participants for a recording session at both venues. A schedule was drawn-up for each session. Apart from a 1 hour slot at the outset of each recording session, the subjects were allocated 20 minutes for their solo recordings and scheduled according to their availability within one 3 hour time-frame.[44] At the completion of drawing-up a time-table, the subjects were asked to confirm plausibility of their allocated slots. Consequently, the recording schedules for each recording session was finalized.

3.2. Music used for Recordings

Two pieces were selected for use as the performance material for all VOT. They are 1.) *Great Minds Against Themselves Conspire*, the penultimate chorus from Henry Purcell's (1659-1695) opera *Dido and Aeneas*, written for SATB, (see Appendix 1); and 2.) *Immortal Bach*, originally a chorale (*Komm, süßer Tod)* by J. S. Bach (1685-1750), arranged by Knut Nystedt (1915-) for SATB, (see Appendix 2).

These pieces would usually be distinguished along several divergent lines. Piece no. 1 represents a choir piece from the opera genre, with piece no. 2 being representative of the general genre of *a cappella* choir music. They differ in character, as piece no. 1 would mostly

[44] The subjects were granted this option as the scheduled time-frame fell within a full morning – from 07:40-12:40 – on an academic day, where some individuals had prior academic commitments.

be performed by an opera chorus; whereas piece no. 2 would usually be performed only in a chamber concert. The performance approach of these would also differ in that piece no. 1 lends itself to dramatic expression (associated with operatic performance practice, including acting), contrasted to the much more reserved approach necessary for the execution of good intonation in piece no. 2 (where choristers would – nearly as a rule – be expected to perform less animatedly, focusing on implied expression, in the character of chamber music recitals). However, performance practice is subject to performance preferences of individuals, who – implicitly – have subjective interpretive penchants. Often this would dictate certain technical approaches required of singers.[45] To discount performance subjectivity as an influencing factor in our experiment, subjects were referred to the main experimental instruction: all subjects were to perform as a soloist.

3.3. Recording Session 1

3.3.1. Venue Set-Up and Data Acquisition

Venue: Endler Concert Hall, Music Dept., *Konservatorium*, Stellenbosch University

It was previously requested of each subject to arrive at the recording session fully warmed-up, as to their own discretion. Each of them agreed to this request. The subjects were expected to arrive at 07:40 on the day of the first scheduled recording session. They were briefed regarding the outline for the entire recording session. After all subjects, the main researcher and the present sound engineer were in consensus as to the lay-out of the schedule for the session, the main researcher gave the performance instructions for all recordings to follow in that session. The performance instructions were as follows:

1. During VOT subjects were to perform as soloists; this applied to all ensemble and solo VOT.
2. The researchers necessitated 3 usable recordings, at the end of the recording session, for any given VOT, i.e. 3 x ensemble recordings and 3 x solo recordings/singer were needed.

The only qualitative instruction was that subject should strive to perform all VOT ‘artistically’. Thus, no instruction was given regarding choir ‘blending’, vibrato extent, exact

[45] Once again, mainly resting on qualitative debates, controversies such as straight tone singing and reduced vibrato come to mind.

dynamic range and pronunciation. All performance responsibility thus rested with the singer-subjects.

At 08:00 the recording of the ensemble VOT ensued. One hour was set aside to garner three good ensemble recordings. During this session the choir stood in a pre-determined semi-circle formation.[46] Before each take a reference chord was played on a well-tuned Bösendorfer Concert Grand piano to ensure a consummate entry from the ensemble, given that both pieces were to be performed *a cappella*. A video camera was placed on the center-axis of the semi-circle, in between the 4^{th} and 5^{th} subjects. The camera recorded a conductor conducting the ensemble during each ensemble take. In this way, it was ensured that the subjects all sang in the same tempo. The conductor did not create any interpretative expectations through his conducting gesture, so as to leave all artistic interpretation as the responsibility of the singer. This was also important so as to rule out the conductor's influence as a limiting factor. The conductor made slight gestural references at break-off points, in order to minimise scatter of final consonants. Furthermore, the conductor conducted in a very flowing style. This would aid *legato* singing and be conducive to breath-flow. The conductor strove to conduct in a near-identical fashion during all recordings of piece no. 1. Similarly, he strove to maintain conducting consistency during all recordings of piece no. 2. The subjects confirmed that the conducting style employed by the conductor throughout the recording session did not pose any threat to healthy vocalisation and did not contradict the classical vocal technique harnessed by any of the subjects.

By 09:00 three audio recordings – with concomitant video recordings of the conductor – of the ensemble, for both pieces, were successfully made. Before commencement of the first solo recording take, myself and the recording engineer needed to choose a video and audio pair that would be implemented for playback as pitch reference (audio) and time reference (video) for all solo recordings. The 3^{rd} recording for each piece was selected at random. Both pieces were performed at a rather moderate tempo, with piece 1 at roughly q = 82, and piece 2 at roughly q = 44-48. A reference chord for each piece was also recorded and added to the respective reference tracks.

[46] Refer to *Section 3.3.2.2.* for a discussion on placement of subjects.

During solo takes, subjects would sing along to an ensemble take played back through single-sided headphone feedback, with the concomitant video of the conductor projected onto a screen in front of the subject. The audio and video of the 3rd takes of the ensemble sessions for both pieces were used for the respective solo recording sessions. Although this choice does not align with the preferred choice of the auditors for spectral analysis, using these tracks as references would not negatively affect the output of the subjects during solo recordings. All three ensemble tracks available for auditing were of high quality and standard. There were only slight differences in the amount of head resonance produced between tracks. This is, essentially, what the auditors were expected to distinguish.[47] Correlatively, the choice of the 3rd recorded ensemble tracks as reference tracks for solo takes was discounted as a limiting factor to vocal output of subjects.

After the video and audio for playback were synchronised in Pro Tools by the recording engineer, solo recordings followed as scheduled between 10:00h and 12:40h. Subjects returned to the recording venue and took-up their demarcated positions on the stage. The subjects were each given the opportunity to get accustomed to the audio and video playback procedure. They could listen to the ensemble recording played back once through the headphones, with video being projected onto the screen. Subjects were welcome to sing along to the playback to get used to the balance between the playback and their own singing. All subjects commented, having being asked, that the reference chord, conductor cues and ensemble feed through the headphones was clear and easily comprehensible. At the conclusion of the recording session three recordings for both pieces for each subject were available for auditing.

[47] *Appendix 3*shows the auditors' track selection for analysis.

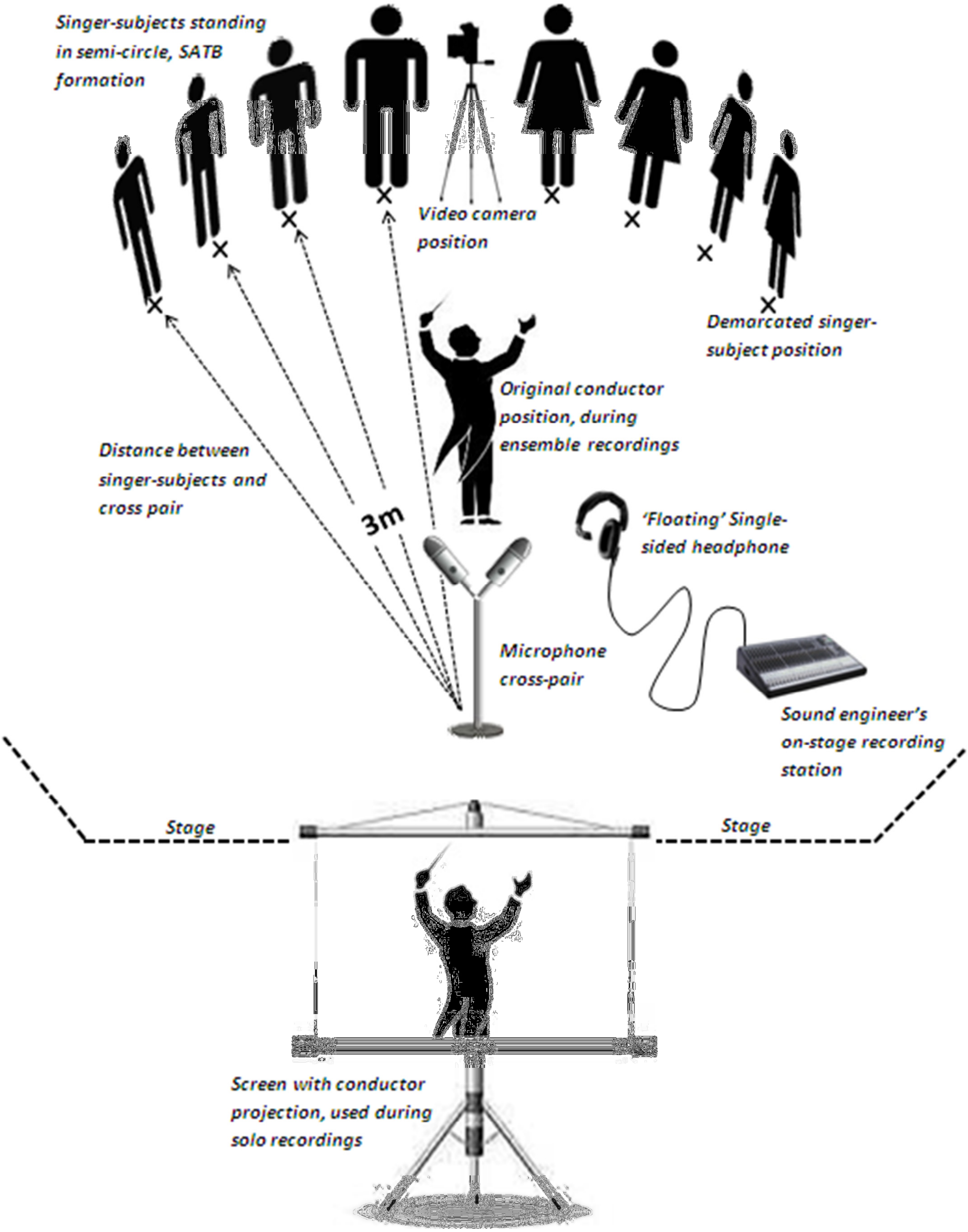


Fig. 3.1 Stage and recording set-up for recording session 1 in venue 1, Endler Hall, *Konservatorium*, Music Department, SU.

3.3.2. Recording Set-Up (See Fig. 3.1)

3.3.2.1. Recording Apparatus

Two *Sennheiser* MKH 8040[48] were arranged in an XY set-up in front of the subjects. The audio interface was the Avid HD Omni[49] with 35 dB gain on the built-in microphone pre-amplifiers. Recordings were captured in Pro Tools 10 HD. Audio was sampled at 96 kHz, and saved through 24 bit linear encoding. Beyerdynamic DT 102 - 400 Ohm[50] single-sided headphones were used for playback during all solo VOT.

3.3.2.2. Placement of Singer-Subjects

At the outset of the first recording session subjects were asked to stand in a semi-circle formation, grouped in voice categories. From the audience's perspective, basses stood to the left, tenors center-left, mezzos center-right and sopranos to the right. Given the width of the ensemble, subjects were spaced equidistant at a 3 meter radius in front of the microphone pair for optimal recording. Once a neat, comfortable and equally-spaced formation was achieved strips of masking-tape with the names of the respective subjects written on it were placed on the position where each subject stood.

It was important to demarcate a position for each subject to create consistency between recordings. With the microphones recording in stereo every subject's output is picked-up in a directional fashion relative to the subject's position to the microphone pair. Thus, by fixing the subjects' position the researchers could make sure that the subjects stood in the same position relative to microphones for all ensemble and solo recordings. Consequently, the output level of the subjects – other than whether the subjects subjectively sang differently – between solo and ensemble takes could be ruled out as a limiting factor.

3.3.2.3. Conductor Projection

A white screen was erected ± 2 meters in front of the stage, in the audience seating, on the center-axis of the venue. A projector set-up on the edge of the stage projected the video of the conductor back onto the screen. The screen was placed at this specific position so as to be at a comfortable viewing position relative to the sight of each subject. The projection of the

[48] See specifications: http://en-de.sennheiser.com/condenser-microphone-cardioid-guitar-acoustic-bass-brass-mkh-8040.
[49] See specifications: http://www.avid.com/US/products/hd-omni.
[50] See specifications: http://europe.beyerdynamic.com/shop/dt-102.html.

conductor onto the screen was bright and clear – to the satisfaction of all subjects. Singers did not note any discomfort as a result of viewing the screen. Acoustic interference of the screen was negligible, due to the directional pick-up characteristics of the microphones.

3.4. Perceptual Auditing of Recorded Audio

In order to maintain anonymity of the singer-subjects, unique codes were assigned to each recording. These codes were used for track-listing purposes for the auditing process. The tracks were arranged in a playlist at random. Firstly, the order of singer-subjects was chosen at random; secondly, the order of the three recordings per subject and ensemble were listed at random. This was necessary in order to give equal weight to each recorded track and avoid the possible assumption – on the part of the auditors – that there was a natural progression of standard between the first and third recording for each recording group.

Two experienced senior singing lecturers at a South African university were asked to participate as perceptual auditors in choosing the one track for each ensemble and solo recording context to be used for spectral analysis. Two hours were required for the auditing of all audio. At the time of auditing, the average amount of years that the auditors have lectured at tertiary level was 19.25.

The auditors were invited to come to the Music Dept. of the University of Stellenbosch, where the auditing of the recordings would take place. They were expected to follow one main instruction for all auditing tasks: preferences for the chosen recordings were only to be based on perceived *strength of the subjects' and ensemble's singer's formant cluster*. No other adjudicating factors were to take precedence over this instruction. In other words, irrespective of auditors' exception to qualitative or technical anomalies evident in the vocal output of individual subjects, or the ensemble, the track that displayed the best use of head resonance were to be put forward for analysis. Nine tracks for each piece, amounting to a total of 18 tracks, were required to be set aside for analyses at the conclusion of the auditing phase.

The auditing procedure took place in the recording studio. Tracks were played back through Pro Tools 10 HD over an optimal set-up of Bowers & Wilkins 805 Diamond Speakers[51]. The auditors were free to ask for playback levels to be adjusted to volumes that suited their best aural observation of singer-subjects' use of F_s for all VOT.

For practical reasons only two individuals were asked to participate in the auditing process. However, this fact steered the way in which tracks were set apart for analysis. As there were only two auditors, it was difficult to expect to get unanimous decisions; for, at any point the auditors could prefer two different tracks for one recorded group (i.e. auditor 1 could prefer the 2nd take, whereas auditor 2 could prefer the 1st or 3rd take). For this reason the auditors were asked to put forward their preference; in the event that there were different preferences, the auditors would discuss their opinions and be given the opportunity to listen to the three recorded options once again. Much of the time the auditors unanimously chose the same takes. For takes that needed to be discussed consensus was reached very speedily. It is believed that the auditing process, with the positive attitude and expert adjudicating skills of the auditors, yielded the best choices of tracks to be used for spectral analytical purposes.

3.5. Mixing and Editing of Audio Data

All tracks were trimmed – in the time domain – so that all audio for a given piece would be the same length. For piece no. 1 (Great Minds) lengths were trimmed to 42s, with piece no. 2 (Immortal Bach) being trimmed to 35s.

Our methodological approach would not be legitimate if we were to compare the solo recordings with the ensemble. It would be incorrect to merely divide STER and SPR means of the ensemble by 8 as a means to compare with the unique solo recordings. An average of the ensemble would comprise spectral information of all four represented voice categories (i.e. Bass, Tenor, Alto, and Soprano). In order to 'compare apples-with-apples', we needed to represent the solo recordings in a way that amalgamated the four voice categories. By mixing the solo tracks, associated with a specific ensemble template, into one track we then had a 'soloist mix' ensemble that could stand proxy for the soloists to be compared with the authentic ensemble recording. These mixes were performed in Pro Tools 10 HD. In the end,

[51] See specifications: http://www.bowers-wilkins.com/Speakers/Home_Audio/800_Series_Diamond/805-Diamond.html.

our conclusions would be based on a comparison of STER and SPR values between the ensemble template and the soloist mix.

For spectrum analyses concerned with *resonance* strategies of classically trained singers it is vital to delimit audio that only contains vowel spectra. All other miscellaneous spectra components – such as silence, breathing points and consonants – are not associated with resonance. Given that comparisons were to be made between different recordings (i.e. ensemble and solo recordings) we needed to garner audio samples that were aligned between all solo tracks and their respective ensemble template recording. Thereby we could ensure that all samples for a specific point in a piece contained comparable spectra information.

Following the standardising of the time component for all tracks of the two pieces, and mixing the soloist track for each piece, four sample time-frames were identified in both pieces. Table 3.1 displays the time-frames of the four samples within each piece.

Table 3.1 Sample lengths (ms) and number of 20ms windows (sliding across the sample face at 10ms intervals) associated with the spectral analysis of each sample.

Piece	Sample 1		Sample 2		Sample 3		Sample 4	
	Length (ms)	Windows analysed	Length (ms)	Windows analysed	Length (ms)	Windows analysed	Length (ms)	Windows analysed
Great Minds (Purcell)	400	38	500	48	300	28	1000	98
Immortal Bach (Nystedt)	1000	98	1000	98	500	48	1000	98

One notices that some time-frames for samples in *Great Minds* are relatively short. In the spectral domain, there is ample information to be gathered from such samples. Also, the analysis sliding-window employed in the spectral analysis phase is only 20ms long, and moves at intervals of 10ms[52]. Thus, for a 300ms time-frame, we garner 28 STER and SPR readings, with as many as 98 of these readings for samples that are 1000ms in length. Apart from sample 3, being 500ms in length, all frames in *Immortal Bach* are 1000ms long. From

[52] Please refer to the next section for details surrounding the analytical process.

this piece, then, we get 48 and 98 readings, respectively. The samples were extracted *en mass*, per piece, in Pro Tools 10 HD.

The four samples extracted from each piece would serve as the four sample groups needed for analyses. At this point we had 80 individual samples (2 pieces x 4 samples x 10 solo/ensemble/soloist-mix tracks). All recordings were made at 96 kHz, in stereo. Thus, for each sample there were two channels available for analytical purposes. It would be superfluous to analyse separate channels for individual recordings. All audio was downsampled to 48 kHz (for ease of processing in MATLAB), and the stereo tracks were summed into mono. For this conversion we used a bash script invoking the command line tool FFmpeg.[53] The editing stage reached its completion with the normalising of the samples in Sound Exchange (SoX), invoked with another bash script.[54] The normilasing amounted to an effective gain increase of 5.11dB across all analyses samples. In both processes batch editing was applied to all tracks of the same piece.

3.6. Spectral Analyses: Design and Implementation of Analytical Tools

The literature showed that for our analysis purposes it would be most expedient to employ Singing Power Ratio (SPR) and Short-Term Energy Ratio (STER) to ascertain the strength of the F_s as harnessed by the singer-subjects during any given VOT.[55]

MATLAB was used to program tools, for spectral analyses.[56] A graphical user interface (GUI) was created for simple navigation in MATLAB, easy importing of audio files and quick rendering of data. Ferguson et al.'s (2010) method for estimating STER was followed. A Hanning window, 20ms in length, sliding across the wave file of the audio sample at 10 ms intervals, was designed. It would perform simultaneous FFT, STER and SPR on any given audio sample that was imported into the tool. Traditionally SPR would be calculated from the LTAS (Reid et al., 2007), as the highest points of the low- and- high-ends of the vocal spectrum would lie on the contour of the LTAS (Ferguson et al., ibid.). However, given that

[53] *Appendix 4* provides details of the FFmpeg script used for conversion (see http://www.ffmpeg.org/).
[54] *Appendix 4* provides details of the SoX script used for conversion (see http://sox.sourceforge.net/).
[55] In *Sections 2.3.1 & 2.3.2* thorough definitions and implementation for spectral analytical purposes, respectively, of these spectral indicators are given.
[56] See *Appendix 5* containing the MATLAB codes used for all tools.

Ferguson et al.'s (ibid.) method involves calculating STER from very short (20ms in length) LTAS readings in the first place, it is most beneficial to use the upper extremes of the low- and- high bandwidths calculated for this process as readings for SPR. Thus, we have incorporated the measurement of SPR and STER as part of the same process. The means of the multiple STER and APR readings for one sample is later calculated, giving an STER and SPR means for the entire sample. In this way the Long Term Average Spectrum (LTAS) is negating and a more accurate means is given for any given sample – as endorsed by Kenny & Mitchell (2006).

3.7. Preparation of Data for Descriptive Analyses and Statistical Analyses Procedure

The main finding of this study is expressed in terms of final STER and SPR 'means' for the ENS and SM for both pieces. Calculating these means could follow one of two averaging processes: an arithmetic averaging process, or a geometric averaging process. First mentioned is a standard linear averaging process, expressed by the following formula:

$$\frac{x_1 + x_2 + x_3 \ldots x_n}{n}$$

Generally the geometric means serves the purpose of providing a more accurate means of comparison, than the arithmetic averaging process, between data sets that are not entirely homogenous. It achieves this by weighting several outlying measures in a data series differently than would be the case during linear averaging.[57] The geometric mean is represented by the n^{th} root of the product of all the values in the series,

$$\left(\prod_{i=1}^{n} x_i\right)^{1/n} = \left[\sqrt[n]{x_1 x_2 x_3 \ldots x_n}\right] = exp\left(\frac{1}{n}\sum_{i=1}^{n} \log x_i\right)$$

Through the use of the sliding window during the spectral analytical process the window 'sees' a different spectral image with every reading, due to dynamic inter-actions of the spectral content. In order to find one mean that appropriately represents the shifts between windows the geometric means would be the most fitting. In the context of our data, the geometric mean serves the purpose of regulating a vast spread of STER, or SPR, means yielded by any one given sample. The samples extracted from the two pieces are also slightly incongruent in that, especially for GM, they all render varying amounts of windows. Thus the

[57] See WIKIPEDIA: http://en.wikipedia.org/wiki/Geometric_mean.

geometric function creates more homogeneity between the samples in order for us to appropriately calculate the final STER and SPR values for each piece from their respective samples groups.

The final stage of the research project dealt with the crucial aspect of corroborating the findings of the spectral analyses statistically. In order to determine significant differences between singers performing in ensemble mode and solo mode, the final STER and SPR means of these modes, for both pieces (*Great Minds* & *Immortal Bach*) were analysed. A Repeated measures Analyses of Variance (ANOVA) test was applied to the STER and SPR means of the four samples comprising the final means. Five per-cent significance levels (i.e. $p<0.05$) were used to determine significant differences.

3.8. Recording Session 2

3.8.1. Details Regarding Cancelation of Session

Venue: Recording studio, Music Dept., *Konservatorium*, Stellenbosch University

The recording studio was selected as the venue for the second recording session. Initially it was envisaged to reproduce the recording method employed in the first recording session. A comparative analysis of spectral findings between the two venues was to be undertaken. Findings could elucidate voice students' phonation tendencies, with specific reference to the F_s, in various venues.

Correspondence with the ensemble, divulging details for the second session, was finalised well before the set date. A recording schedule for the second session was drafted and corroborated with the subjects. The entire session was subverted when, in the space of several hours prior to the start of the second recording session three singers informed me that they were unable to sing due to health reasons. Unfortunately the session had to be postponed.

Despite my best efforts, I was unable to find a further date that suited all the subjects for the second recording session before the end of the academic year. Given that a comparative study of the same ensemble singing in different venues were to be done, it was not possible to schedule the recording session for the next semester as the two students that were in their 8^{th} semester of VT were to complete their music studies at the end of that semester. Both moved away from the proximity of the recording venue. This was the first confounding factor. Furthermore, their involvement in any further recordings after the conclusion of their studies would disregard a main premise of this study, in that the target group should comprise undergraduate voice-majors.

3.8.2. Implications and Reassessment

Due to social constraints, exacerbated by circumstances out of my control, it was not possible to undertake the second recording session, scheduled to take place in the recording studio. Accordingly, I only garnered data from one recording session in the concert venue. If any, data acquired in the concert venue would suffice for the spectral analytical method employed. This is corroborated by the experimental methods followed by Reid et al. (2007); who thought it of critical importance to make recordings – also with the aim to measure levels of SPR and ER – in a venue that subjects were familiar with. In our case, the most comfortable venue for subjects might be the vocal studios where they receive instruction. However, these would not be large enough for ensemble recordings. The concert venue would be the next most comfortable venue, as it boasts good acoustic properties, for feedback purposes. Subjects also regularly perform recitals and undertake vocal examinations in this hall.

Barnes et al. (2004), further intimate that singers should "be recorded in surroundings more conducive to the feeling of performance, [which would make] the recorded samples more reliably representative of the subject's professional sound". In our study the recording studio might have proven to be too clinical, and not conducive to natural phonation by the subjects, especially given the developmental phase that they were at, at the time of the recording session. It is thus argued that the absence of data from recording in the studio does not limit findings, apart from the fact that the acquired data represents a small sample size and only one performance scenario. Findings would only be applicable to the concert venue milieu; it would be conjecture to extrapolate results to other performance environments, such as recording, practice and teaching studios.

In retrospect, any attempts to replicate procedures in divergent venues, such as acoustically modified recording studios, especially anechoic chambers[58], would lead to complications that would have been beyond the scope of this project. The acoustic differences between the

[58] During the conceptualising phase of this project, an anechoic chamber was considered as an option for recordings, alongside the concert hall and recording studio.

venues would lie at the heart of such challenges. Such difficulties were evidenced in Ternström's (1989) comparative studies of several choirs. He found that boys' choirs and adult choirs adapted their vocal techniques as a result of varying feedback brought about by divergent acoustic properties of several venues used in the experiments. Although it would conflict with some pedagogical constructs that require *professional* singers to maintain their optimal vocal technique in any given situation (Miller, 1993: 76), the vocal technique of our target group – singers-in-training – might also be affected (Mürbe et al., 2007; Mendes et al., 2003) by acoustic dissimilarities between different venues.

CHAPTER 4: RESULTS

Two pieces, Purcell's *Great Minds against Themselves Conspire* and Nystedt's *Immortal Bach*, were selected as material for Vocal Tasks (VOT). A selected ensemble recording, for each piece, served as template to which all solo recordings were matched (in terms of tempo and pitch). From each piece the four samples that were extracted *en mass*, for each item – including the manual Soloist Mix (SM) – served as sample repetitions. Thus, for both pieces four sample data groups were available for spectral analyses. However, the four sample groups were time-dependent, as we could only select samples that displayed a homogenous vowel, and did not contain any other non-resonatory spectra components.

Spectral analyses of the audio, and the subsequent statistical analyses of data, revealed divergent, yet noteworthy trends between the two musical excerpts. For piece one, *Great Minds*, singers generally produced more relative spectral energy in the F_s region during ENS mode than during their individual solo recordings.[59] The exact opposite seems to be true for *Immortal Bach*, yet statistical analyses did not verify the significance of the results for this piece.

The values displayed in Figure 4.1 are the final STER (a) and SPR (b) means and concomitant 95% confidence intervals (CI), indicated by the error bars, for ENS and SM for *Great Minds*. Findings show that singers produced significantly more F_s spectral energy when singing in ENS than in solo mode ($p<0.05$). STER and SPR means were lower in ENS as compared with SM, see Table 4.1 for final means and standard deviation (STDEV).

In Figure 4.2 the final STER (a) and SPR (b) means and CI for ENS and SM for *Immortal Bach* are juxtaposed. Here STER and SPR of the SM was lower than that of the ENS, see Table 4.1. Findings were, however, not significant for both STER as well as SPR means.

[59] From the outset the reader is reminded of the correct interpretation of STER and SPR values: for both indicators *lower* means indicate *higher* relative energy in the singer's formant cluster (F_s); *higher* means indicate the *opposite* resonance trend.

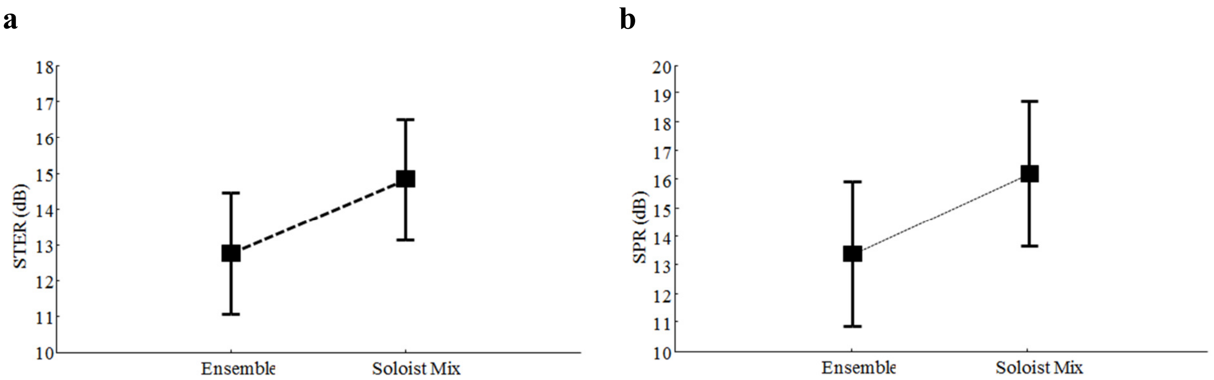


Fig.4.1 Graph showing mean and standard deviation of STER (a) and SPR (b) for *Great Minds* (GM). For STER F(1, 3)=23.05 and p=0.02. For SPR F(1, 3)=23.54 and p=0.02.

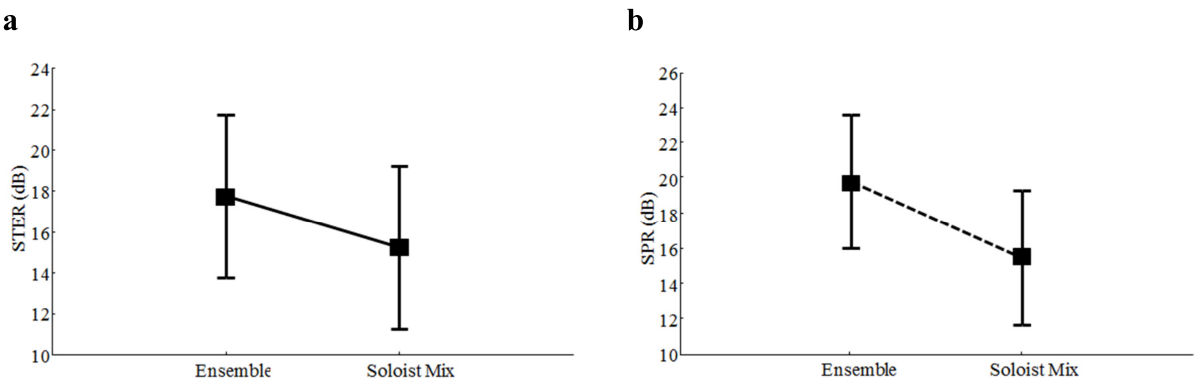


Fig.4.2 Graph showing mean and standard deviation of STER (a) and SPR (b) for *Immortal Bach* (IB). For STER F(1, 3)=3.79 and p=0.15. For SPR F(1, 3)=10.198 and p=0.05.

Individual means and standard STDEV for STER and SPR for all four samples in *Great Minds* (GM) are delineated in Figure 4.3. The means of these samples contribute to the final mean seen above in Figure 4.1. The same details, regarding *Immortal Bach* (IB), are provided in Figure 4.4.

The graphs express marked differences of the characteristics of the spectral content in the F_s region between the two musical excerpts. As described before, one notices that the STER and SPR sample groups of the ENS in Figure 4.3 (GM) generally lie lower on the Y-axis than that of the SM. It asserts that singers were generally singing with more F_s spectral energy in ENS mode. The opposite is suggested by the values in Figure 4.4 (IB), where the SM samples generally lie lower on the Y-axis than that of the ENS. But, in IB several overlaps between means of the ENS and SM are observed. The STER of sample 3 in IB overlaps with that of samples 1 and 4 in SM mode. SM STER means for sample 4 also overlaps with all other ENS STER means. The SPR means of sample 4 of the SM for IB overlaps with sample 2 and 3 of the ENS. There are no overlaps of this sort between ENS and SM in GM.

Table 4.1 Mean and standard deviation of STER & SPR for *Great Minds* (GM) and *Immortal Bach* (IB).

Group/Individual Label	STER GM	SPR GM	STER IB	SPR IB
Ensemble	12.8 ± 0.87	13.4 ± 1.41	17.7 ± 2.10	19.8 ± 1.04
Soloist Mix	14.8 ± 1.23	16.2 ± 1.75	15.2 ± 2.84	15.5 ± 3.23
B1	14.0 ± 2.47	15.2 ± 2.21	17.9 ± 5.09	19.3 ± 4.55
B2	15.8 ± 4.10	17.9 ± 4.59	20.0 ± 6.43	22.4 ± 6.02
T1	15.4 ± 3.39	15.6 ± 4.05	8.5 ± 8.96	8.6 ± 9.38
T2	12.0 ± 4.52	11.8 ± 4.66	19.6 ± 3.46	21.8 ± 5.17
A1	18.0 ± 2.53	19.2 ± 2.17	21.1 ± 10.79	22.8 ± 11.04
A2	17.4 ± 3.83	18.1 ± 3.60	20.0 ± 7.68	22.3 ± 7.87
S1	16.6 ± 2.30	16.5 ± 2.61	19.8 ± 5.34	20.5 ± 6.33
S2	18.5 ± 2.81	19.0 ± 2.81	23.4 ± 6.83	24.2 ± 7.34

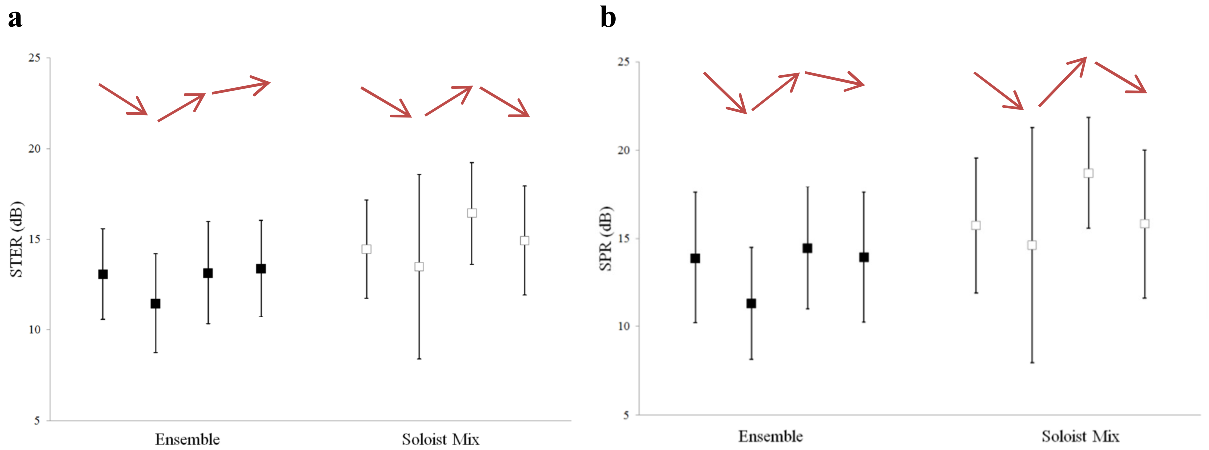


Fig.4.3 Graph showing mean and standard deviation of STER (a) and SPR (b) for each of the four samples that were used to calculate the mean STER & SPR for *Great Minds* (GM). Red arrows articulate trends between sample groups.

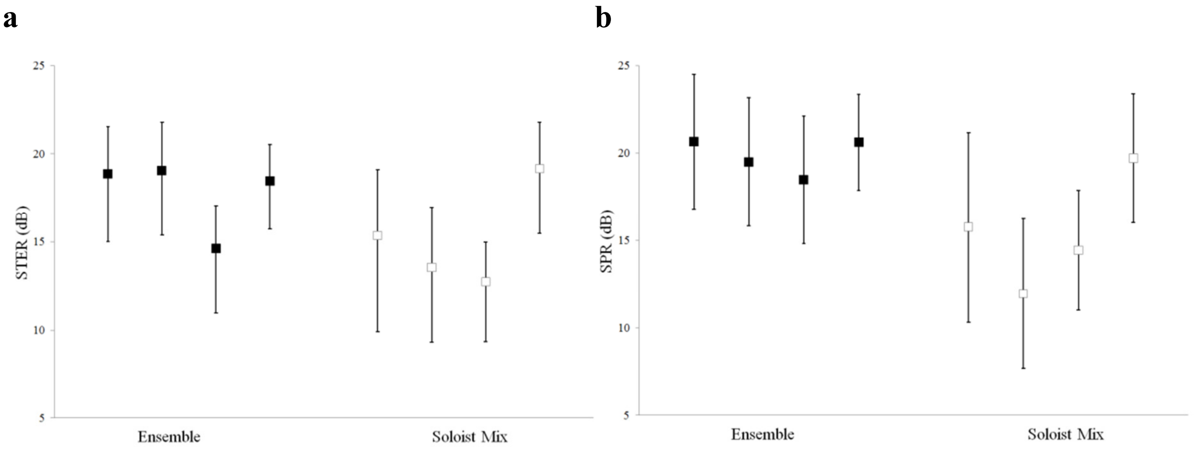


Fig.4.4 Graph showing mean and standard deviation of STER (a) and SPR (b) for each of the four samples that were used to calculate the mean STER & SPR for *Immortal Bach* (IB).

Another major difference is in the general spectral energy difference between GM and IB. One notes that means generally lie higher on the dB scale in IB than in GM. Table 1 further illustrates this observation. STER means for both ENS and SM are mostly below 15dB. In IB one sees that the same indicators lie above 15dB.

Singers produced more consistent spectral energy in ENS mode than in solo mode, when considering STDEV. The error bars for sample 2 of the SM in Figures 4.3, and that of sample 1 in Figure 4.4 indicate the greatest STDEV. That the ENS shows more consistent energy levels is reflected in Table 1. We see here that the final standard deviations (i.e. STDEV over the four samples), for both sets of spectral indicators, are consistently lower for the ENS than for SM.

The graphs in Figures 4.3 & 4.4 further illustrate resonance tendencies of the singers between the two pieces. The trend between STER means of the ENS in GM is mirrored, albeit with increased magnitude, in the SM groups. Succinctly stated, increase and decrease in STER and SPR means follow very similar trends (except for the discrepancy in the trend between STER samples 3 & 4 of the ENS and that of the SM) between the ENS mode and SM, see the red trend lines in Figure 4.3. These trends indicate that the singers were consistent in the resonance strategies they employed for specific phrases within GM. There are no consistent trends in IB.

In this study a direct comparison was made between the ENS and SM modes. The means of the SM is implicitly proportional to the means of the solo samples, as the SM comprises the eight solo samples. However, from one SM mean one is not able to discern STER and SPR trends between individuals. By analysing the data garnered from the solo recordings, one is able to note resonance characteristics of individual voices, and groups of singers.
Figure 4.5 show STER (a) and SPR (b) means and STDEV for all four samples of *Great Minds* for the eight soloists, as a series. The data are grouped in voice classifications. Thus, from the left to right of each graph the four-part ensemble is displayed with basses on the left, through tenors and altos, to sopranos at right. [60] A quick scan of both plots conjures a strong visual correlation between the STER and SPR means. The general trends from B1 through S2 for both plots are highly similar, save for the proportionally higher SPR means of B2.

a **b**

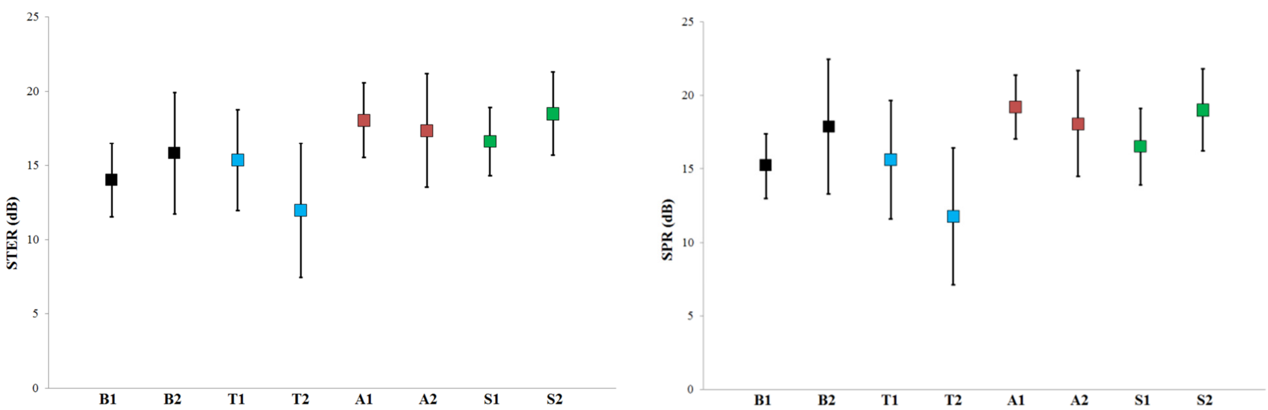


Fig.4.5 Series of mean and standard deviation of STER (a) and SPR (b) for individual singers, over four samples, for *Great Minds* (GM). Values are grouped according to voice classification, in BTAS (Bass, Tenor, Alto, and Soprano) format from left to right. Colours aid in distinguishing between voice-groups.

Several interesting trends within voice-groups, as well as between gender groups are highlighted. Generally, the female voices produced less spectral energy in the F_s region than the males. This is truer when considering the STER means; whereas the SPR plot shows an overlap between B2 and S1. Intra-group means provide further meaningful trends: a marked

[60] The use of numbering was simply employed as a way to distinguish between two data sets within the same voice classification. It does not imply any other classification, or qualitative ranking.

overlap exists between B1 & T2 and B2 and both tenors; the altos both overlap with S1. These overlaps are pronounced in the SPR plot.

Although the female voices produce generally higher means of STER and SPR than the men, their STDEV – represented by the error bars – are smaller in GM, but slightly higher in IB. This difference is poignantly represented between the SPR of T2 and A1. Both voices are outliers, in relation to the other 6 singers. T2 has the lowest SPR, with A1 having the highest. But, A1's STDEV is much lower. The large deviation of T2 calls into question the significance of his mean. Similar to A1, B1 and S1 show less deviation.

A series of means and STDEV for STER & SPR of all soloists for IB is plotted in Figure 4.6. The topographies of the STER (a) and SPR (b) plots are similar to that of GM in two ways: 1.) all the means lie within a narrow band of ±10dB, except for one extreme outlier (T1) at about 10dB lower than the second lowest singer (B1), for both indicators; and, 2.) a general trend indicates that males, once again, produced more spectral energy in the F_s band, compared to female singers. With this said, one would note that the general values for the female singers do not differ so starkly from that of the males, in comparison to the readings of GM.

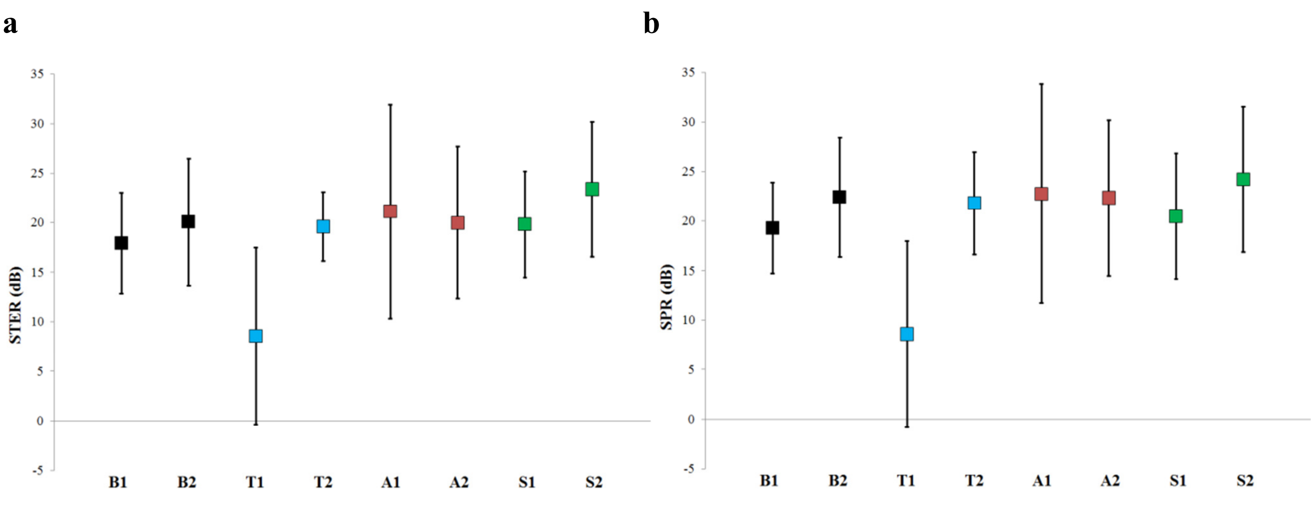


Fig.4.6 Series of mean and standard deviation of STER (a) and SPR (b) for individual singers, over four samples, for *Immortal Bach* (IB). Values are grouped according to voice classification, in BTAS (Bass, Tenor, Alto, and Soprano) format from left to right. Colours aid in distinguishing between voice-groups.

A major observation for both spectral indicators is that the general STDEV is larger in IB than in GM. In fact, apart from T2's STER reading and B1's SPR reading, all other STDEV ≥ 10dB. It evidences that, for the most part, the singers produce more inconsistent readings of spectral energy during this piece. Also, the narrow ±10dB means bandwidth lies at 15dB-25dB. This is 5dB higher than for GM.

Several findings are significantly different in comparison to GM. Firstly, the schism between T2 and A1 has been reduced to several dB and STDEV was inverted. In GM A1 had among the lowest STDEV and T2 the highest. In IB T2 provides the most consistent readings, as opposed the A1 that displays the greatest STDEV for both pieces. Secondly, the altos produced more spectral energy in the higher spectral bandwidth (indicated by lower lying STER means around 20dB & SPR means around 23dB), relative to the basses, in this piece. This aids in attenuating the degree of difference between the male and female voice parts. Lastly, the highest mean spectral energy was produced by T1, as compared to T2 in the first piece. The STER and SPR readings for T1 are also lower than that of T1 in GM.

The values for T1 are of special interest in this piece. One would note that the STDEV for both dips below zero. This is due to the fact that he produced several negative values for both indicators over Samples 2-4 of IB. The minimum range for SPR extended as low as -5.75dB; STER had a minimum range of -4.59dB. In other words, for the samples that dipped below the zero dB threshold, T1 was actually producing more spectral energy in the higher band (2 kHz-4 kHz) than in the lower band (0 kHz-2 kHz).

One of the most significant secondary findings is that final STER and SPR paired-means for both pieces proved to be directly proportional to one another. The scatterplot in Figure 4.7 demonstrates this important relationship. Analyses points to strong correlations in both GM (R^2 = 0.92) and IB (R^2 = 0.97).

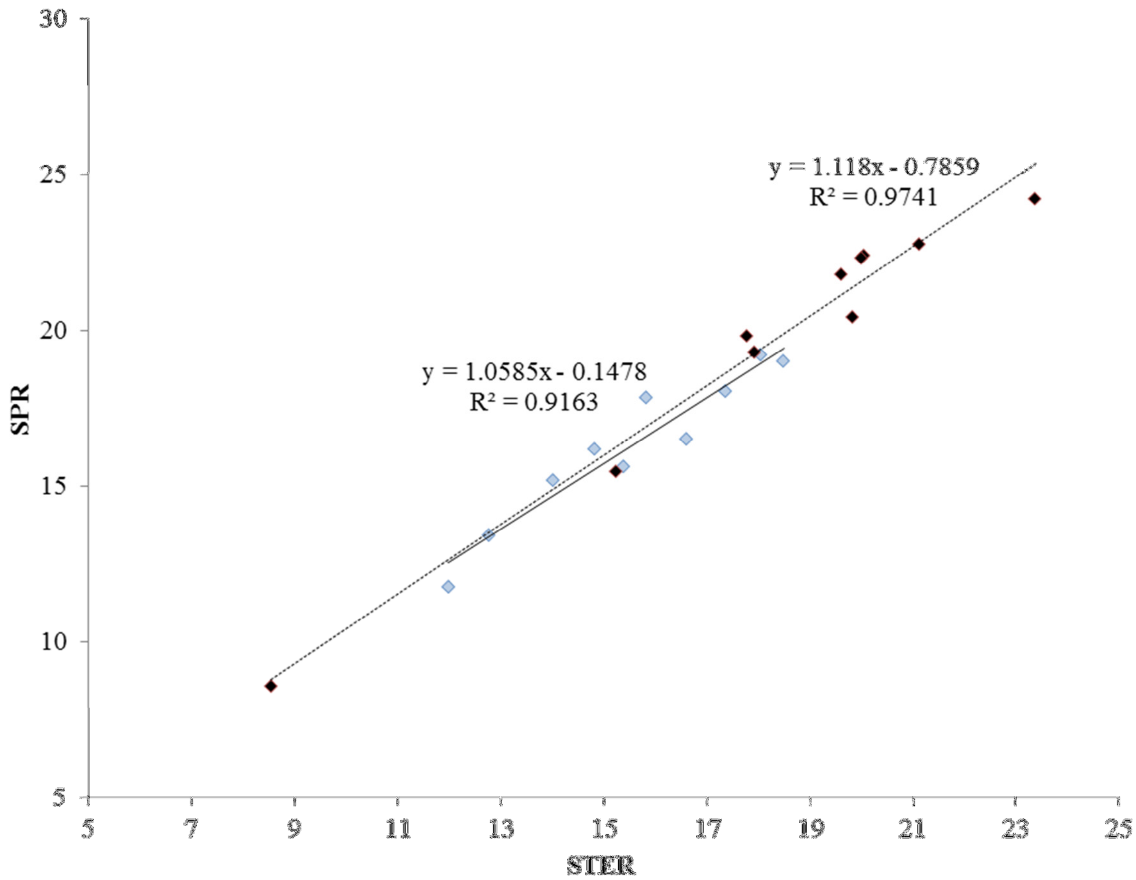


Fig.4.7 Relationship of all ensemble and solo STER and SPR pairs. Light-blue markers represent *Great Minds* (GM) pairs, accompanied by a solid trend line; black markers indicate *Immortal Bach* (IB) pairs, with the dotted line displaying the trend. Minimum values for both axes have been set at 5dB for ease of viewing.

In line with the central, significant findings ($p<0.05$ for STER and SPR means) – that the singers produced more F_s spectral energy when singing in ENS mode, compared with solo mode, in GM – Table 4.2 presents correlating data. Here the differences of the STER and SPR means between the ENS and correlative SM and solo recordings are given. If the difference produces a negative value, the ensemble had a lower means for that specific comparison. The negative relationships are all highlighted. There are only 7 values (importantly, including the SM STER and SPR values for IB) that are not highlighted. For these values, the STER, or SPR, means of the ENS was higher than that of the related SM or solo means.

Table 4.2 Mean STER & SPR difference between Ensemble and Soloist Mix/individual for *Great Minds* (GM) and *Immortal Bach* (IB). All values less than 0 (i.e. less than ensemble value) are highlighted. A negative value indicates that singers produced more spectral energy in the F_s region during Ensemble mode than during solo recordings for the given piece.

Group/Individual Label	STER GM	SPR GM	STER IB	SPR IB
Ensemble	0	0	0	0
Soloist Mix	-2.06	-2.83	2.53	4.35
B1	-1.25	-1.80	-0.16	0.53
B2	-3.06	-4.49	-2.30	-2.60
T1	-2.60	-2.22	9.20	11.24
T2	0.78	1.65	-1.86	-1.97
A1	-5.28	-5.81	-3.37	-2.93
A2	-4.60	-4.67	-2.23	-2.51
S1	-3.85	-3.11	-2.08	-0.65
S2	-5.73	-5.60	-5.63	-4.41

CHAPTER 5: DISCUSSION

5.1. Main Findings

5.1.1. Presentation of Main Findings

The main objective of this study was to ascertain the efficacy of an individual singer's resonance strategy in the spectrum bandwidth of 2.4 kHz – 3.6 kHz, during two performance modes: solo and ensemble (ENS). The main findings confirmed our hypothesis that *undergraduate voice majors produce more energy in the singer's formant cluster region when singing in ensemble mode as compared to solo mode*. In one piece (*Great Minds*) the singers produced significantly higher amounts of F_s spectral energy in ENS mode ($p<0.05$), whereas in the other piece (*Immortal Bach*) results indicate the opposite; however, differences were not significant.

For *Great Minds* (GM) the 95% confidence intervals for STER and SPR (inclusive of both ENS and solo modes) are lower than for *Immortal Bach* (IB). In more descriptive terms, the true STER and SPR means – for both singing modes – falls within a range of 5dB in GM. The range for IB is higher, at 8dB. These measures are implicitly associated with the significance of the differences between means of the ENS and soloist mix (SM) in both pieces. Thus, not only do we know that differences of the indicators between ENS and SM in GM are significant, we also see now that their means lie within a narrower CI range than in IB.

The strength of the performance – in terms of head resonance – of the solo singers during ENS mode in GM is seen in the robust STER and SPR means of the four analysed samples. Results showed that there were no overlaps between STER, or SPR, means of ENS and SM samples. Two sets of standard deviation (STDEV), on both a macro and micro scale, were also telling in this regard: 1.) The STDEV of the four ENS samples (macro scale) of GM, for STER and SPR, was smaller than for the same measures in SM; and, 2.) the STDEV of the individual samples (micro scale) for the ENS of GM, for STER and SPR, was generally lower than that of the SM.

We saw similar trends of comparative STDEV, on a macro and micro level, between the ENS and SM modes in IB, albeit that final means were lower for SM mode. Thus, these trends make a strong case, verified with robust findings, for the notion that singers have the ability to produce F_s spectral energy in ENS mode in excess of the amounts produced as soloists.

5.1.2. Consistency of Main Findings with Previous Research

The main premise of this study found its impetus in the study done by Reid et al. (2007). These authors were the first to show that professional singers produced significantly more F_s spectral energy in ENS mode than in SM mode. Their findings provided a framework for our hypothesis. However, our application differs from theirs in that our target group comprised undergraduate vocal performance majors.

Importantly, our study generated findings that proved to be consistent with those of Reid et al. (ibid.). Thus, although our main research question and the concomitant findings were not wholly new, our study did provide empirical data for a previously unexplored target group; herein rests the innovation of our findings.

Our findings would need to be corroborated through further investigations involving the same target group. But, it is clear that, by being consistent with those of Reid et al. (ibid.), our significant findings further substantiate the hypothesis that classically trained singers do inadvertently produce more head resonance singing in ensemble mode, as compared with solo mode, when following the same performance instructions, and if all other limiting agents are kept constant.

5.2. Secondary Findings

5.2.1. Divergence between Sexes

Several interesting secondary findings were revealed through analyses of our results. Firstly, we saw that males generally produced more head resonance than females. Apart from several SPR overlaps between one of the basses and some female voice, and one soprano SPR mean dipping below that of a bass and tenor in IB, the male STER and SPR means are consistently lower than that of the female voices. This trend is unequivocally presented in the series of STER and SPR solo means for GM. In IB the distinction is much narrower.

The general topography of the STER and SPR means does not seem to correspond very well with the F_s ranking model implied by Sundberg (2001) and Millhouse et al. (2002). It is implied that F_s strength varies from strongest to weakest in baritones, tenors, altos/mezzos and sopranos (in that order). This model is grounded in the postulation that the F_s is more a feature of the male classical singer's resonance technique than of higher female voice classifications (Barnes et al., 2004; Sundberg et al., 2012). In our target group both basses mentioned that they would be classified as baritone/bass. Ideally we should expect the trend described by Sundberg (2001) to be evident in our findings.

The ranking model is not explicitly represented in our results. We saw a definite distinction between the results of the male and female voice groups, however, it was only pronounced in GM. Numerous factors led to disassociation with the expected upward rising trend. Higher STER and SPR means in the men – relative to woman – in IB; generally low spectral energy in the altos; consistently high energy displayed by one soprano – relative to the rest of the woman, and some men; and major overlaps between the tenors and basses, subverted the expected trend.

One major similarity with the ranking model is that – apart from erratic activity in the resonance strategies of the tenor voices – one bass consistently produced the lowest STER and SPR means, with his antithesis in one soprano, who – save for one alto SPR mean – consistently produced the highest STER and SPR readings. Its exclusivity does not imply that this is an insignificant finding. It is in line with the generalisation that baritones produce the highest amounts of spectral energy in the F_s region (Sundberg, ibid.), as opposed to sopranos that are argued to produce the least (Weiss et al., 2001).

5.2.2. Correlation between STER and SPR Means

In both pieces the topography contours created by juxtaposing the solo STER and SPR means in a series showed the strong correlation between STER and SPR pairs. These indicator pairs are highly proportional to one another. A decided gap was evidenced in the STER and SPR pair of the Ensemble in IB. This gap stands in stark contrast to the very small margin between the STER and SPR means of the Soloist Mix in IB – evidence for the singers producing more F_s spectral energy in solo mode. Fascinatingly, this finding is undercut by the fact that results for IB were not significant. However, it certainly shows that, in this way, the solo recordings showed a level of consistency that was lacking in ENS mode during the performance of IB.

This relationship between the STER and SPR pairs emphasises the value of using these spectral indicators in conjunction with each other. In a way, the one corroborates the relevance of the other. For example, the SPR means highlight the highest peaks in the low and high spectrum of a given window. However, these peaks are not necessarily underpinned with high general spectral energy, measured by means of the STER. Peaks with very steep rises and drop-offs could, theoretically, have low 'base energy'. A high peak in the low band could have low energy content, compared with a lower peak in the high band that has very higher energy content. In this situation a very high SPR reading would be calculated, accompanied by a very low STER reading. Which one of these reading properly portrays the singer's use of high spectral energy? The ideal would be to have both indicators lying in proximity so as to validate spectral energy information.

Our findings show that STER and SPR readings were directly proportional. Due to the strong correlation, this finding validates the parallel use of STER and SPR in this study. In the broader scope of research it serves to be consistent with Barnes et al. (ibid.), as well as with Kenny & Mitchell (2006) who stated that they "expected" to find "a very highly significant relationship between ER and SPR".

The relationship between STER and SPR spoken of here should not be confused with resonance trends. We are reminded that the soloists produced less consistent levels of high band spectral energy during the performance of IB – the incongruent trends between the four samples of the SM points to this fact. We learn from this that the energy content of the singers' spectra were characterised by balanced STER and SPR means, yet, at the four different points of measurement in each piece, the general energy that was depicted by these means were unbalanced.

5.2.3. Relevance of Divergent Aspects of Target Group in Reference to Target Groups from Similar Studies

In this study the target group consisted of undergraduate performance majors that were nearing the completion of their vocal performance studies. The solo STER and SPR means garnered in this study provided us with a means to draw comparisons of their performance with that of other target group individuals from similar studies.

Regarding the females in our study, Kenny & Mitchell's (ibid.) is the closest comparative study. These authors focused their research only on female voices. Their target group consisted of 3 x sopranos and 3 x mezzos. Their level of development was slightly higher than the females from our target group, with all 6 having recently received their first degree in vocal performance. Mean SPR and Energy Ratio (as compared to STER in our study) were calculated. SPR values ranged between 13dB – 25.7dB with mean SPR at 19.7dB. Their ER values lay between 11.1dB – 22.3dB, and the mean was 16.4dB. In our study females produced STER means ranging between 16.6dB – 23.4dB. The STER means is 19.4dB, which is 4dB higher the Kenny & Mitchell STER means. Our target group had a SPR mean at 20.3dB, ranging from 16.5dB – 24.2dB. Here our females produced very similar SPR means to that of the comparative study.

The results of the male voices in our target group can most appropriately be compared to that of the singing students in the study done by Fergusson et al. (2010). The researchers called together 9 male baritones, 3 of which were students at a similar level of development than ours. Only STER readings were calculated, though. The STER range for these 3 singers was between 10dB – 16dB, with a means at about 13.3dB (exact mean were not provided). Our males produced STER means from 8.5dB – 20dB and the mean was at 15.41dB.

Lundy et al. (2000) measured the SPR of 21 senior undergraduate vocal students (among others), and found the mean to be 10.9dB. Their target group was inclusive of male and female students; thus, we can't draw a direct comparison to their results. However, generally our male and female voice-groups seemed have higher (i.e. less spectral energy) SPR means at 16.6dB and 20.3db (respectively) in relation to Lundy et al.'s (ibid.) group.

Comparisons drawn between the resonance strategies of individuals within our target group and those from target groups are, at most, interesting. It seems that for both male and female groups in our study STER and SPR data were highly similar to data provided from their comparative studies. However, in all comparative cases it is almost certain that the variables that were not constant between our and other studies, such as repertoire choice, target group details and sample details, would subvert direct comparison. Neither did we perform statistical analyses on the STER and SPR data of individuals. But, these inconsistencies do not totally detract from drawing superficial comparisons.

5.2.4. Phenomenal Findings in Reference to Similar Studies

The STDEV of singer T1 was indicative of the fact that he had produced negative STER and SPR values. It is an unprecedented example of the F_s region displaying more energy relative to the low bandwidth that contains the first two – very strong – formant bandwidths around the fundamental frequency. After the initial conception of SPR by Omori et al. (1996), and STER by Thorpe et al. (2001), the use of SPR and STER as spectral indicators has only recently risen to prominence as a way to quantify F_s spectral content. In very few of these studies do we see similar, negative, ratios as those of T1 in our study. One singer in Omori et al.'s (ibid.) target group, described as an amateur baritone singer with about four years' vocal training experience (i.e. comparable to our singer), produced a negative SPR mean. Of all the studies incorporating SPR, STER, or both, only Fergusson et al. (ibid.) found that several advanced and professional baritones from their target group produced negative STER.

The negative STER and SPR of T1 are evidently not common measurements, as seen above. It is not clear what conclusion to draw from these results. On the one hand the correlation with the experienced singers in Fergusson et al.'s (ibid.) creates a very healthy impression of T1's resonance strategy during the implicated performance. Also, STER and SPR means of T1 compare very interestingly with those of the two tenors in Reid et al.'s (2007) study. In last mentioned study the tenors produced identical means, with the energy ratio at 14.4dB and SPR at 16.0dB. T1 had STER means at 15.4dB (GM) and 8.5dB (IB), and SPR at 15.6dB (GM) and 8.6dB (IB). In the context of our experiment, T1 ranks among the professional tenor soloists from Reid et al (ibid.). T2 also showed highly competitive results in this regard in GM. On the other hand, the contrasting results between the two performed pieces for both T1 and T2 bring into question their general consistency, as we spoke of earlier. Their competitive STER and SPR means could be viewed to be undermined by a lack of consistency between the two pieces.

5.3. Reliability of Findings

5.3.1. Statistical Robustness

This study was severely constricted in that descriptive trends of data at lower-lying strata could not be defended statistically. We were concerned with the way singers' head resonance varies – if indeed at all – between two singing modes. Our findings made a meaningful and significant contribution to understanding 'intra-ensemble' resonance tendencies of undergraduate voice majors, but we did not track the trends of specific individuals, as Reid et al.'s (ibid.) study showed. Significant findings are applicable in the broader field, but our relatively small sample-size (n=4) does limit the strength of statistical data.

On the opposite side of the spectrum, we did not establish the impact factor of broader aspects such as repertoire. We sampled data from only two pieces. Although it provided us with enough data to measure resonance strategies, it would be inappropriate to draw any conclusions on the actual influence of the repertoire. It could be safely stated that repertoire exerted a causal force on our findings. It is widely known that music representative of different stylistic compositional approaches calls for a myriad of varying performance approaches. In retrospect it would have served our study well to investigate music from one specific genre, or even one stylistic period.

The size of our target group was adequate, satisfying the definition for ensemble according to Ford (2003). Reid et al. (ibid.) recorded only 12 singers (i.e. one singer more, per voice category). However, the impact of individuals on the final STER and SPR should not be negated. In our relatively small target group several outlying voices would have a telling bearing on the ensemble measures. We do not necessarily believe this to have been the case, as we found only one singer in the solo mode of each of the two pieces to be severe outliers (ironically, these were at very low STER and SPR levels). With this said, it was beyond our capacity to measure outliers in the ENS, as we were presented with an essentially amalgamated wave-file for this mode. It is the corollary of the fact that we could not track individual voices in ENS mode.

Research conducted in previous years report large subject-participation sizes: Omori et al. (ibid.) = 37; Watts et al. (2006) = 39. Our numbers pale in comparison. Kenny & Mitchell (ibid.) had a target group of only 6 singers and Fergusson et al. (ibid.) included only 9. But, in both last mentioned studies these groups represented the sample sizes (i.e. n=6 and n=9, respectively); implying that our sample size (n=4) was – once again – limited. The significance of findings still stand in the light of the fact that our data is associated with 8 undergraduate singers, compared with only 6 in Kenny & Mitchell's (ibid.) and 4 students in Fergusson et al.'s (ibid.) study.

5.3.2. Reliability of Methods

The use of single-sided headphones could not be ruled out as a causal factor in our study.[61] It served the purpose of providing the individuals with a pitch reference, concomitant to the tempo reference in the conductor projection, during solo recordings. It could be argued that it compromised the aural facility of the soloists; despite the individual soloists communicating confidence in their ability to provide convincing performances whilst wearing the headphones. Alternative means of providing singers with on-going pitch references (considering that they were singing *a cappella*) was beyond the technical scope of this study.

Comparing our findings with those from similar studies would be based on the *proviso* that our microphone distance was invariably different to that of the comparative study. Our microphone-pair was placed equidistant at 3 meters in front of the singers. Typical recording techniques see microphones placed several centimeters from the mouth as the microphone is often head-mounted (Reid et al., ibid.; Fergusson et al., ibid.; Watts et al., ibid.). Our increased distance may have contributed a low signal-to-noise ratio (SNR) to the recorded audio (Fergusson et al., ibid). However, our methodology employed to measure STER and SPR means from many windows provides an accurate measurement of spectral energy. The spectral peaks that we are concerned with would negligibly conflated by the ambient noise-floor of the recording venue; and – at most – the influence of SNR would be constant over all solo and ensemble recordings.

5.4. Significance of Study

The implications of this study are two-fold. Firstly our study provides the first empirical-based research on undergraduate performance majors as the main target group of a study of this nature. Previous studies have explored the resonance strategies of untrained singers and professional singers singing in ensemble mode. Our study is the first to be geared toward this

[61] See *Section 3.3.2.1.*

evidenced gap, and findings proved to be significant. Therefore, from an academic point of view this study is doubtless one the first of its kind in South Africa, and (due to our target group) our findings are relevant internationally.

Secondly, our findings could contribute to a growing body of empirical research in striving for a more informed pedagogical approach, among voice teachers and ensemble, or choral, practitioners, pertaining to undergraduate voice students. Concomitantly, this study shows that undergraduate performance majors optimise their head resonance when singing in an ensemble setting. It suggests that this mode of singing could benefit a singer-in-training further than complementing his/her aural facility or interpretative intelligence: now it is postulated that a singer's phonation and resonance technique is advanced inadvertently.

It is clear that the choral music practice has much to gain through the participation of voice students in any choral venture. However, a blatant lack of responsibility has been evidenced in the choral practice at large, in that there is wide-spread ignorance regarding the treatment of especially developing singers. Acoustic and voice pedagogical literature unequivocally articulate that the epitomical manifestation of the F_s, or head resonance, is grounded on healthy vocal technique. If the choral practice is to gain from this acoustic feature – which this study has shown it could –choral practitioners could show initiative by regarding vocal pedagogy more readily.

But, the biggest investors in this regard are the voice students themselves. By being accommodative of empirical acoustic research choral practitioners and voice pedagogues would allow for more transparent, and necessary, discourse such that a didactical approach rendering free reign could be founded. This free reign would allow developing singers to benefit from both solo and corporate music making practices, if the singer has reached the appropriate stage in his/her development.

5.5. Suggestions for Future Research

5.5.1. Effect of Repertoire

Reid et al. (ibid.) restricted extrapolation of their results to the confines of a choir serving the function as a professional opera chorus. Their material only included operatic music. Our significant findings are associated with the opera chorus that was used as performance material. It is still uncertain what the causal effect of a wider scope of repertoire would be on resonance tendencies of soloists between solo and ensemble mode. Future study in this line is necessary.

5.5.2. Target Group Size

Mendes et al. (2003), Mürbe et al. (2007) and Fergusson et al. (ibid.) have explored divergent factors pertaining to the technical development of undergraduate vocal students. Still, there is yet to be compiled a composite literature base to provide concrete evidence for the resonance strategies of undergraduate vocal performance majors singing in ensemble as compared with solo modes. We believe that increasing target group sizes (thereby increasing sample sizes and diversifying data) would strengthen statistical trends between singing modes, and further strengthen the general understanding of the F_s resonance strategies of this pivotal target group.

5.5.3. Longitudinal Studies

This phase in the career of aspiring singers is characterised by divergent and inconsistent development of inter-dependent factors (Mürbe et al., ibid.). It would be expedient to track F_s spectral energy changes of undergraduate singers in further longitudinal studies. Comparative studies conducted between a control group that is comprised of vocal students that do not participate in ensemble ventures and a vocal student group that regularly participates in ensemble singing, of the entire course of undergraduate performance studies, would serve well in this regard. Such a study could further be bolstered by the crucial research question whether the usage of F_s in ensemble mode could develop any aspect of a singer's technical facility further than in the voice studio. Findings would point to decision making regarding future individuals' participation in ensemble ventures, or projects at an undergraduate performance level.

5.5.4. The Need for One Variable to Represent F_s

Spectral indicators STER and SPR are two similar quantitative measures of the way a singer harnesses spectral energy in the F_s region. They were developed independently (Omori et al, ibid.; Thorpe et al., ibid.), but have since proved to be complimentary when their means are juxtaposed. A fascinating feature of their relationship is that their values are approximate; and it goes without saying that their means are directly proportional (Kenny & Mitchell, ibid.).

Yet, the difference between paired STER and SPR means still provides space for various interpretation and it has not yet been ascertained which one of these indicators are the more appropriate indicator, with regards to their application. Future research could explore new physical and mathematical models incorporating both STER and SPR in one variable that is the quantitative representation of an individual's F_s.

5.5.5. Ambiguity surrounding Spectral Energy in Higher Female Voices

We have learnt elsewhere that the presence of the F_s in the spectrum of high female voice classifications is a highly debated issue.[62] In their empirical studies, where measures of energy ratio or SPR were employed to measure spectral energy levels in the F_s, Reid et al. (ibid.) included sopranos in their target group – in spite of the debate – and Fergusson et al. (ibid.) excluded the participation of mezzos and sopranos. In this light, future empirical research could point voice scientists toward a way of defining that characteristic spectral feature of mezzo and soprano ranges that is analogous with the F_s associated with male voice groups. Appropriate spectral indicators would be necessitated to quantify this feature and aid in further studies where target groups include cross-spectrum voice categories.

5.6. Conclusions

The initial impetus that gave rise to this study was that there is much dispute about the participation of classically trained singers in choral or ensemble ventures. What was unearthed is that it is actually the use, and in a sense the preservation of, a singer's ability (correlated with his/her technical grounding) to maintain high levels of spectral energy in the singer's formant cluster (F_s) that adds the most meaning to such discussion. This provided the premise for our study, where we sought to investigate this phenomenon-like, resonatory feature of a trained singer's vocal output. Literature proved to be laden not only with references to the F_s, but a rich scope of empirical research has provided us with means to understand the acoustic manifestation of the F_s.

We noticed that the extant literature did not shed much light on empirical data regarding the F_s among singers that are situated in the phase between being classified as amateur and professional. Aligning this study with that of others, we hypothesised the following:

[62] See *Section 2.2.3*.

undergraduate voice majors produce more energy in the singer's formant cluster region when singing in ensemble mode as compared to solo mode.

Further delving into the literature showed that Singing Power Ratio (SPR) and Short-Term Energy Ratio (STER) measures are, as yet, regarded to be the most adequate way to quantify F_s spectral energy. Following spectral analytical methods that are modelled on very recently published approaches, we garnered SPR and STER means of the performance of singer-subjects during two juxtaposed modes: solo and ensemble. Our main findings would rest on a comparison between the performance of the group of singers in ensemble and solo modes.

Our main finding was that the singers in our study indeed produced significantly higher levels of spectral energy in the F_s when performing in the ensemble, than when performing as soloists. That the hypothesis was confirmed was an astonishing finding, given that literature suggested that it was unpredictable what the resonance tendencies of singers at this stage of their development would be. Also, our main finding is aligned with published research focussing on professional singers.

Our data further provided very interesting insight into the acoustic nature of the spectral domain of the performance of our target-group singers. We saw glimpses of individuals producing unprecedented levels of energy in the higher spectral bandwidth. Several trends provided evidence for the applicability of certain theoretical principles, governing vocal acoustics, in this study. One such trend is that male singers generally did produce higher levels of F_s spectral energy than female singers, including the fact that it was a baritone/bass that generally produced the most energy. These are constructive findings, reflecting very positively on the current state of the technical development of individuals in the implied target group. In this vein, however, we also highlighted descriptive trends that indicate inconsistencies in the way some singers phonate. But, such findings could very well be afforded to the bridging phase that singers are entering into at present, between being developing singers and earning professional status.

A very relevant finding is the strong correlation between SPR and STER means in both pieces. Such findings strengthen the notion that these spectral indicators should be readily employed as a paired-measure indication of F_s spectral energy. To date SPR and STER have not been used alongside each other, as has been the case before with Energy Ratio (ER) and

SPR. We certainly endorse their use in future studies that aim to quantify high-band spectral energy. Further research could see the amalgamation of these measures into one representative indicator.

The singer-subjects of the target group in this study have definitely displayed the potential for developing singers to optimise F_s spectral energy output when singing as soloists, irrespective of the performance mode being ensemble or solo. This indicates that developing singers can maintain consistent resonance in response to the instruction of singing in a soloistic fashion. Thus, in terms of the preservation of resonance technique, it seems possible that this instruction would aid in mitigating resonance discrepancies of voice students performing in ensemble mode – an observation that could be of interest to any individual playing an instructive role in the development process of a voice student.

The main contribution of this study is that it added valuable empirical data to an underrepresented body of literature, with regards to the acoustic knowledge base of voice students. No previous studies have attempted a comparative study of undergraduate performance majors performing in ensemble and solo modes. Having elucidated the head resonance strategies the subjects in our target group followed, it would be expedient to replicate these findings in future studies. We have not so much as uncovered the tip of the ice-berg, which are the divergent aspects of undergraduate performance majors' technical development. More empirical data is necessitated to locate the exact phase in young singers' development where they start harnessing full control over the resonatory function of their voices.

LIST OF REFERENCES

Alldahl, P. 2008. *Choral Intonation* [Trans. Carrol, R.]. Stockholm: Gehrmans Musikförlag.

Barnes, J. J., Davis, P., 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 pp. 530–538.

Bragg, K. A. 2012. Development and Utilization of Vibrato and Non-Vibrato Singing in Choral Music. *Unpublished Masters Dissertation.* University of Central Missouri.

Cabrera, D., Yadav, M. & Kenny, D. T. 2011. Linearity in the loudness envelope of the *messa di voce*. In: *Proceedings of ACOUSTICS '11. Gold Coast: Australia*; 2011: 1–5.

Carlsson, G. & Sundberg, J. 1992. Formant Frequency Tuning in Singing. *Journal of Voice* 6: 3 pp. 256–260.

Choi, S. H., Lee, J. Y., Sprecher, A. J. & Jiang, J. J. 2012. The Effect of Segment Selection on Acoustic Analysis. *Journal of Voice* 26: 1 pp. 1–7.

Christy, V. A. 1977. *Foundations in Singing: A Basic Textbook in the Fundamentals of Technic and Song Interpretation*. 3rd ed. Dubuque, IA: WM. C. Brown Company Publishers.

Daugherty, J. F., Manternach, J. N. & Brunkan, M. C. 2012. Acoustic and perceptual measures of SATB choir performances on two types of portable choral riser units in three singer-spacing conditions. *International Journal of Music Education* DOI: 10.1177/0255761411434499.

Detweiler, R. F. 1994. An Investigation of the Laryngeal System as the Resonance Source of the Singer's Formant. *Journal of Voice* 8: 4 pp. 303–313.

Detwiler, G. 2008. Solo Singing Technique & Choral Singing Technique in Undergraduate Vocal Performance Majors: A Pedagogical Discussion. *Unpublished Doctoral Dissertation.* University of Cincinnati.

Ekholm, E. 2000. The Effect of Singing Mode and Seating Arrangement on Choral Blend and Overall Choral Sound. *Journal of Research in Music Education* 48:2 pp. 123–135.

Fauls, B. K. S. 2008. A Choral Conductor's Reference Guide to Acoustic Choral Music Measurement: 1885 to Present. *Electronic Theses, Treatise and Dissertations.* Paper 4492. [http://diginole.lib.fsu.edu/etd/4492]

Ferguson, S., Kenny, D. T. & Cabrera, D. 2010. Effects of Training on Time Varying Spectral Energy and Sound Pressure Level in Nine Male Classical Singers. *Journal of Voice* 24: 1 pp. 39–46.

Ford, J. 2003. Preferences for Strong or Weak Singer's Formant in Choral Tone Quality. *International Journal of Research in Choral Singing* 1: 1 pp. 29–47.

Fuelberth, R. J. V. 2004. The Effect of Various Left hand Conducting Gestures on Perceptions of Anticipated Vocal Tension in Singers. *International Journal of Research in Choral Singing* 2: 1 pp. 27–38.

Goodwin, A. 1980. An Acoustical Study of Individual Voices in Choral Blend. *Journal of Research in Music*. 28:2 pp. 119–128.

Hofstee, E. 2010. Constructing a Good Dissertation: A Practical Guide to Finishing a Masters, MBA or PhD on Schedule. Johannesburg, South Africa: EPE.

Howard, D. 2009. Acoustics of the Trained Versus Untrained Singing Voice. *Current Opinion in Otolaryngology & Head and Neck Surgery* 17 pp. 155–159.

Hunter, E. J., Švec, J. G. & Titze, I. R. 2006. Comparison of the Produced and Perceived Voice Range Profiles in Untrained and Trained Classical Singers. *Journal of Voice* 20: 4, pp. 513–526.

Kenny, D. T. & Mitchell, H. F. 2006. Acoustic and Perceptual Appraisal of Vocal Gestures in the Female Classical Voice. *Journal of Voice* 20: 1 pp. 55–70.

Kovačić, G., Boersma, P. & Domitrović, H. 2003. Long-Term Average Spectra in Professional Folk Singing Voices: A Comparison of the *Klapa* and *Dozivački* Styles. In: *Proceedings of the Institute of Phonetic Sciences, University of Amsterdam* 25 pp. 53–64.

Lamesch, S., Doval, B. & Castellengo, M. 2012. Toward a More Informative Voice Range Profile: The Role of Laryngeal Vibratory Mechanisms on Vowels Dynamic Range. *Journal of Voice* 26: 5 pp. 672.e9–672.e18.

Levitin, D. 2008. *This is Your Brain on Music: Understanding a Human Obsession*. London, UK: Atlantic Books.

Lister, M. C. 2009. Male Choristers' Perceptions of and Preferences for Choral Formations Based on Individual Singer Placement within the Ensemble. *Unpublished Doctoral Dissertation.* Ball State University, Muncie, Indiana.

Loy, G. 2006. *Musimathics: The Mathematical Foundations of Music.* Volume 1. USA: The MIT Press.

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 pp. 490–493.

Magill, P. C. & Jacobson, L. 1978. A Comparison of the Singing Formant in the Voices of Professional and Student Singers. *Journal of Research in Music Education* 26: 4 pp. 456–469.

Manternach, J. N. 2011. The Effect of Varied Conductor Preparatory Gestures on Singer Upper Body Movement. *Journal of Music Teacher Education* DOI: 10.1177/1057083711414428.

McCoy, S. 2004. *Your Voice: An Inside View – Multimedia Voice Science and Pedagogy*. Princeton, New Jersey: Inside View Press.

Mendes, A. P., Rothman, H. B., Sapienza, C. & Brown, Jr., W. S. 2003. Effects of Vocal Training on the Acoustic Parameters of the Singing Voice. *Journal of Voice* 17: 4 pp. 529–543.

Mesaros, A. & Astola, J. 2005. Inter-Dependence of Spectral Measures for the Singing Voice. In: *International Symposium on Signals, Circuits and Systems, ISSCS 2005.* 1 pp. 307–310. *IEEE.*

Miller, R. 1993. *Training Tenor Voices*. USA: Schirmer Books.

Millhouse, T., Clermont, F. & Davis, P. 2002. Exploring the Importance of Formant Bandwidths in the Production of the Singer's Formant. In: *Proceedings of the 9th Australian International Conference on Speech Science & Technology* pp. 373–378.

Moon, K. R., Chung, S. M., Park, H. S. & Kim, H. S. 2012. Materials of Acoustic Analysis: Sustained Vowel Versus Sentence. *Journal of Voice* 26: 5 pp. 563–565.

Mouton. J. 2005. *How to succeed in your Master's and Doctoral Studies.* Paarl, South Africa: Van Schaick Publishers.

Mürbe, D., Zahnert, T., Kuhlisch, E. & Sundberg, J. 2007. Effects of Professional Singing Education on Vocal Vibrato – A Longitudinal Study. *Journal of Voice* 21: 6 pp. 683–688.

Nair, G. 1999. *Voice Tradition and Technology – A State-of-the-Art Studio*. USA: Delmar Cengage Learning.

Omori, K., Kacker, A., Caroll, L. M., Riley, W. D. & Blaugrund, S. M. 1996. Singing Power Ratio: Quantitative Evaluation of Singing Voice Quality. *Journal of Voice* 10: 3 pp. 228–235.

Reid, K. L. P., Davis, P., Oates, J., Cabrera, D., Ternström, S., Black, M. & Chapman, J. 2007. The Acoustic Characteristics of Professional Opera Singers Performing in Chorus Versus Solo Mode. *Journal of Voice* 21: 1 pp. 35–45.

Rossing, T. D., Sundberg, J. & Ternström, S., 1986. Acoustic Comparison of Voice Use in Solo and Choir Singing. *Journal of the Acoustical Society of America*, 79: 6 pp. 1975–1981.

Rossing, T. D., Sundberg, J. & Ternström, S., 1987. Acoustic comparison of soprano solo and choir singing. *Journal of the Acoustical Society of America*, 82: 3 pp. 830–836.

Sengupta, R. 1990. Study on Some Aspects of the "Singer's Formant" in North Indian Classical Singing. *Journal of Voice* 4: 2 pp. 129–134.

Smith, C. G., Finnegan, E. M. & Karnell, M. P. 2004. Resonant Voice: Spectral and Nasendoscopic Analysis. *Journal of Voice* 19: 4 pp. 607–622.

Sundberg, J. 1974. Articulatory Interpretation of the 'Singing Formant'. *Journal of the Acoustical Society of America* 55: 4 pp. 838–844.

Sundberg, J. 1987. The Science of the Singing Voice. DeKalb, Illinois: Northern Illinois University Press.

Sundberg, J. 2001. Level and Center Frequency of the Singer's Formant. *Journal of Voice* 15: 2 pp. 176–186.

Sundberg, J. 2003. Research on the Singing Voice in Retrospect. *TMH-QPSR, KTH,* 45: 1 pp. 011–022.

Sundberg, J., Gu, L., Huang, Q. & Huang, P. 2012. Acoustical Study of Classical Peking Opera Singing. *Journal of Voice* 26: 2 pp. 137–143.

Sundberg, J., Lã, F. M. B. & Gill, B. P. 2013. Formant Tuning Strategies in Professional Male Opera Singers. *Journal of Voice* 27: 3 pp. 278–288.

Ternström, S. 1989. Long-Time Average Spectrum Characteristics of Different Choirs in Different Rooms. *TMH-QPSR, KTH,* 30: 3 pp. 015–031.

Ternström, S. & Karna, D. 2002. Choir. In: *The Science & Psychology of Music Performance*. [eds. Parncutt, R. & PcPherson, G.]. USA: Oxford University Press.

Ternström, S., Bohman, M, & Sodersten, M. 2006. Loud speech over noise: Some spectral attributes, with gender differences. *Journal of the Acoustical Society of America* 119: 3 pp. 1648-1665.

Ternström, S. & Kalin, G. 2007. Formant Frequency Adjustment in Barbershop Quartet Singing. In: *19th International Congress on Acoustics. Madrid: Spain*; 2007:1–6.

Ternström, S. 2008. Hi-Fi voice: observations on the distribution of energy in the singing voice spectrum above 5 kHz. In: *Proceedings of ACOUSTICS '08. Paris: France*; 2008:3171–3176.

Titze, I. R. 2001. Acoustic Interpretation of Resonant Voice. *Journal of Voice* 15: 4 pp. 519–528.

Thorpe, C. W., Cala, S. J., Chapman, J. & Davis, P. J. 2001. Patterns of Breath Support in Projection of the Singing Voice. *Journal of Voice* 15: 1 pp. 86–104.

Watts, C., Barnes-Burroughs, K., Estis, J. & Blanton, D. 2006. The Singing Power Ratio as an Objective Measure of Singing Voice Quality. *Journal of Voice* 20: 1 pp. 82–88.

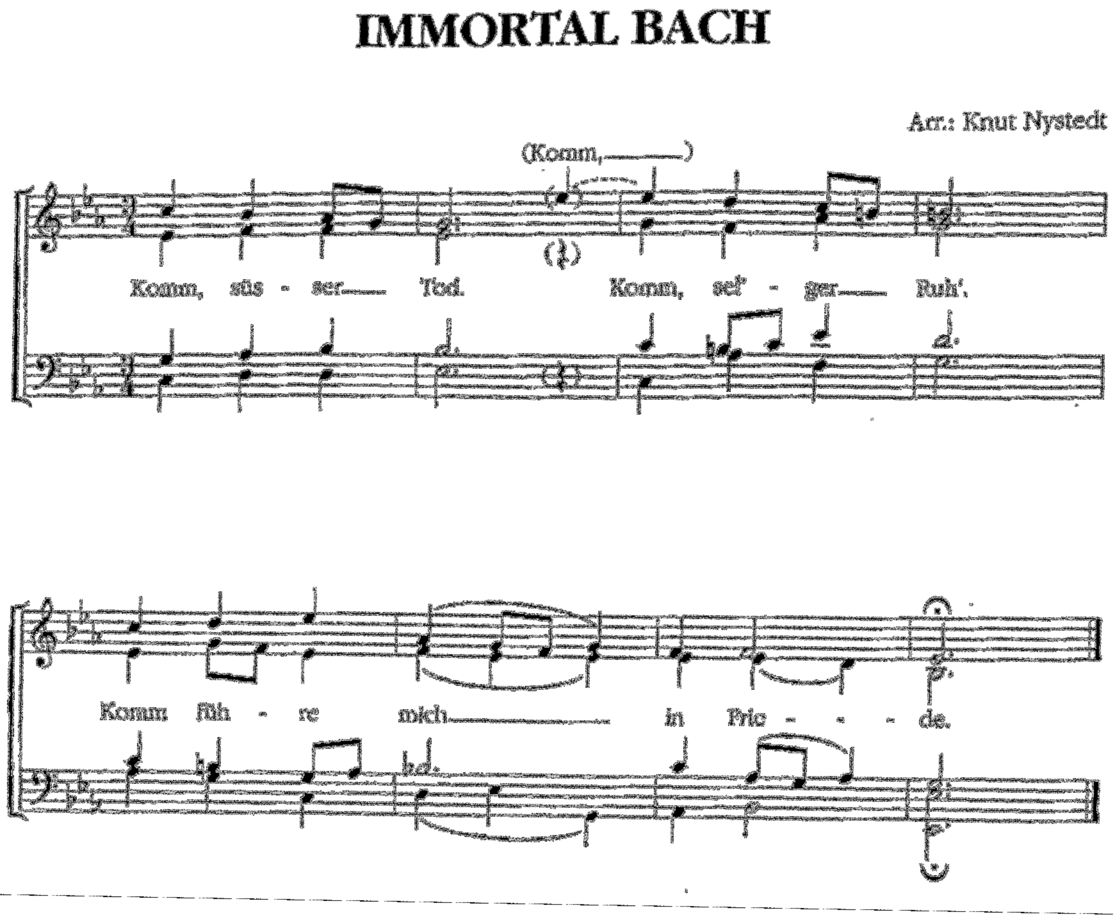
Weiss, R., Brown, Jr., W. S. & Jack Morris. 2001. Singer's Formant in Sopranos: Fact or Fiction? *Journal of Voice* 15: 4 pp. 457–468.

APPENDIX 1

Piece no. 1: *Great Minds against Themselves Conspire* – Henry Purcell (1659-1695)

APPENDIX 2

Piece no. 2: *Immortal Bach – Knut Nystedt* (1915-)



APPENDIX 3

Track-listing sheet for auditing session

Venue: Recording studio, Dept. of Music, SU **Date:** 21-05-2013 **Time:** 13:00

Indicate, by drawing an ' X ' next to the corresponding track number on the track-listing sheet, which *one* track would represent the choice of *exemplary resonance* for each grouping. This choice is subject to consensus between all auditors involved.

Piece # 2: *Immortal Bach* (Nystedt)

Group description	Take #	Track #	Mark choice with X
M3B1	1	1	
	2	2	
	3	3	x
F1A1	1	4	x
	2	5	
	3	6	
F2A2	1	7	
	2	8	
	3	9	x
F4S2	2	10	
	3	11	x
	1	12	
M2T2	1	13	x
	2	14	
	3	15	
F3S1	1	16	
	3	17	
	2	18	x
M4B2	1	19	x
	3	20	
	2	21	
M1T1	3	22	
	2	23	
	1	24	x
Ensemble	1	25	
	2	26	x
	3	27	

Piece # 1: *Great Minds* (Purcell)

Group description	Take #	Track #	Mark choice with X
M1T1	3	28	
	2	29	x
	1	30	
F1A1	2	31	x
	1	32	
	3	33	
M4B2	2	34	x
	3	35	
	1	36	
F3S1	2	37	
	1	38	
	3	39	x
F2A2	2	40	x
	1	41	
	3	42	
M3B1	3	43	
	2	44	
	1	45	x
F4S2	2	46	
	3	47	x
	1	48	
M2T2	1	49	x
	2	50	
	3	51	
Ensemble	2	52	
	3	53	
	1	54	x

APPENDIX 4

Bash script used to invoke FFmpeg for downsampling to 48 kHz and summing to mono:

```
find ./ -type f -name "*.wav" | while read i;
do ffmpeg -i "$i" -ac 1 -ar 48000 "${i%.wav}_48-M.wav";
done
```

Bash script used to invoke SoX for normalizing:

```
find ./ -type f -name "*.wav" | while read i;
do sox -v 1.8 "$i" "${i%.wav}-NORM.wav";
done
```

APPENDIX 5

MATLAB code for all spectrographic tools employed in this study:

```
function varargout = VocalAnalysis(varargin)
% VOCALANALYSIS M-file for
VocalAnalysis.fig
%      VOCALANALYSIS, by itself, creates a new
VOCALANALYSIS or raises the existing
%      singleton*.
%
%      H = VOCALANALYSIS returns the handle
to a new VOCALANALYSIS or the handle to
%      the existing singleton*.
%
%
VOCALANALYSIS('CALLBACK',hObject,event
Data,handles,...) calls the local
%      function named CALLBACK in
VOCALANALYSIS.M with the given input
arguments.
%
%      VOCALANALYSIS('Property','Value',...)
creates a new VOCALANALYSIS or raises the
%      existing singleton*.  Starting from the left,
property value pairs are
%      applied to the GUI before
VocalAnalysis_OpeningFcn gets called.  An
%      unrecognized property name or invalid value
makes property application
%      stop.  All inputs are passed to
VocalAnalysis_OpeningFcn via varargin.
%
%      *See GUI Options on GUIDE's Tools menu.
Choose "GUI allows only one
%      instance to run (singleton)".
%
% See also: GUIDE, GUIDATA, GUIHANDLES

% Edit the above text to modify the response to
help VocalAnalysis

% Last Modified by GUIDE v2.5 22-Apr-2013
17:57:36

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                   'gui_Singleton',  gui_Singleton, ...
                   'gui_OpeningFcn',
@VocalAnalysis_OpeningFcn, ...
                   'gui_OutputFcn',
@VocalAnalysis_OutputFcn, ...
                   'gui_LayoutFcn',  [] , ...
                   'gui_Callback',   []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State,
varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT

% --- Executes just before VocalAnalysis is made
visible.
function VocalAnalysis_OpeningFcn(hObject,
eventdata, handles, varargin)
% This function has no output args, see OutputFcn.
% hObject    handle to figure
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)
% varargin   command line arguments to
VocalAnalysis (see VARARGIN)
% Choose default command line output for
VocalAnalysis
handles.output = hObject;
global WaveParam
% Update handles structure
guidata(hObject, handles);
% Initialize the waveform figure
if strcmp(get(hObject,'Visible'),'off')
end
%set(findobj('Tag','axesWaveform'),'Units','pixels')
;
%set(findobj('Tag','axesWaveform'),'Position',[10
20 600 300]);
%set(findobj('Tag','axesWaveform'),'XLim',[0 4]);

% UIWAIT makes VocalAnalysis wait for user
response (see UIRESUME)
% uiwait(handles.figureMain);
```

```
% --- Outputs from this function are returned to the
command line.
function varargout =
VocalAnalysis_OutputFcn(hObject, eventdata,
handles)
% varargout  cell array for returning output args
(see VARARGOUT);
% hObject    handle to figure
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)

% Get default command line output from handles
structure
varargout{1} = handles.output;



% ---------------------------------------------------------
---------
function FileMenu_Callback(hObject, eventdata,
handles)
% hObject    handle to FileMenu (see GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)


% ---------------------------------------------------------
---------
function OpenFileItem_Callback(hObject,
eventdata, handles)
% hObject    handle to OpenFileItem (see GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)
global WaveData WaveParam FreqParam
% Define the frequency analysis parameters
%LowBandLimits =
struc('LowerLimit',0,'UpperLimit',2000);
%HighBandLimits =
struc('LowerLimit',2000,'UpperLimit',4000);
%FreqParam
=struc('LowBand',LowBandLimits,'HighBand',Hig
hBandLimits,'Window','Hanning');
% Define the filepath
FileFilter = '*.wav';
[FileName,FilePath] = uigetfile(FileFilter,'Select
the wave file');
File = [FilePath FileName];
% Import the data
if ~isequal(File, 0)
   % open(File);
   WaveData = importfile(File);
end
% Define the wave data channel parameters
WaveParam.Channels.Number =
length(WaveData.data(1,:));
WaveParam.Channels.ChannelSelect = 1;
WaveParam.ChannelList = [];
for Index = 1:WaveParam.Channels.Number
   WaveParam.ChannelList =
[WaveParam.ChannelList;Index];
end
% Update the wavedata channel select popup menu
set(findobj('Tag','popupmenuChannelSelect'),'Strin
g',WaveParam.ChannelList)
% Define the wave data timeline parameters
WaveParam.TimeLine.Start = 1;
WaveParam.TimeLine.End =
length(WaveData.data(:,1));
WaveParam.TimeWin = WaveParam.TimeLine;
WaveParam.TimeLine.TimeVec =
linspace(0,(WaveParam.TimeLine.End-
1)/WaveData.fs,WaveParam.TimeLine.End);
% Display the waveData
%set(gca,'XLim',[WaveParam.TimeLine.Start
WaveParam.TimeLine.End]);
%set(gca,'YLimMode','auto');
line(WaveParam.TimeLine.TimeVec,WaveData.da
ta(:,WaveParam.Channels.ChannelSelect)');
set(gca,'FontSize',10);
title('Wave File
Signal','FontSize',10,'FontWeight','bold');
xlabel(gca,'Time
[s]','FontSize',9,'FontWeight','bold');
ylabel(gca,'Amplitude','FontSize',9,'FontWeight','bo
ld');
% Define the wave data time window parameters
Init(hObject);
% Initialize the time window start cursor
hLineWinStart =
line(WaveParam.TimeWin.StartCur.XData,
WaveParam.TimeWin.StartCur.YData);
set(hLineWinStart,'Tag','lineWinStart')
% Initialize the time window start slider
hSlider = findobj('Tag','sliderWinStart');
set(hSlider,'Min',WaveParam.TimeLine.TimeVec(
WaveParam.TimeWin.Start));
set(hSlider,'Max',WaveParam.TimeLine.TimeVec(
WaveParam.TimeWin.End));
set(hSlider,'Value',WaveParam.TimeLine.TimeVec
(WaveParam.TimeWin.Start));
% Initialize the time window end cursor
hLineWinEnd =
line(WaveParam.TimeWin.EndCur.XData,
WaveParam.TimeWin.EndCur.YData);
set(hLineWinEnd,'Tag','lineWinEnd');
% Initialize the time window start slider
hSlider = findobj('Tag','sliderWinEnd');
set(hSlider,'Min',WaveParam.TimeLine.TimeVec(
WaveParam.TimeWin.Start));
set(hSlider,'Max',WaveParam.TimeLine.TimeVec(
WaveParam.TimeWin.End));
set(hSlider,'Value',WaveParam.TimeLine.TimeVec
(WaveParam.TimeWin.End));
```

```
% ---------------------------------------------------------
---------
function PrintMenuItem_Callback(hObject,
eventdata, handles)
% hObject    handle to PrintMenuItem (see GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)
printdlg(handles.figureMain)

% ---------------------------------------------------------
---------
function CloseMenuItem_Callback(hObject,
eventdata, handles)
% hObject    handle to CloseMenuItem (see
GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)
selection = questdlg(['Close '
get(handles.figureMain,'Name') '?'],...
                     ['Close '
get(handles.figureMain,'Name') '...'],...
                     'Yes','No','Yes');
if strcmp(selection,'No')
   return;
end

delete(handles.figureMain)


% --- Executes on selection change in
popupmenuWindow.
function popupmenuWindow_Callback(hObject,
eventdata, handles)
% hObject    handle to popupmenuWindow (see
GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)

% Hints: contents = get(hObject,'String') returns
popupmenuWindow contents as cell array
%        contents{get(hObject,'Value')} returns
selected item from popupmenuWindow


% --- Executes during object creation, after setting
all properties.
function popupmenuWindow_CreateFcn(hObject,
eventdata, handles)
% hObject    handle to popupmenuWindow (see
GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    empty - handles not created until after
all CreateFcns called

% Hint: popupmenu controls usually have a white
background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end
WindowOpt = {'Bartlett','Bartlett-
Hanning','Blackman','Blackman-
Harris','Bohman','Chebyshev','Flat
Top','Gaussian','Hamming','Hann','Kaiser','Nutall','P
arzen','Rectangular','Triangular','Tukey'};
set(hObject, 'String', WindowOpt);


% --- Executes on button press in
pushbuttonAnalyze.
function pushbuttonAnalyze_Callback(hObject,
eventdata, handles)
% hObject    handle to pushbuttonAnalyze (see
GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)
global WaveData WaveParam
% Initialize the wave window
SlideWinOffset = 0;
WinWidth = WaveParam.SlideWin.Width;
WinStart =  WaveParam.TimeWin.Start;
WinEnd =  WinStart+WinWidth-1;
% Define the sampling frequency
Fs = WaveData.fs;
% Define the frequency increment
DeltaF = Fs/WinWidth;
% Define the frequency plot window
FreqVec = 0:Fs/WinWidth:Fs/2;
FreqWinStartIndex = 1;
FreqWinEndIndex = FreqWinStartIndex;
while (FreqVec(FreqWinEndIndex) <
WaveParam.FreqWin.Width)
   FreqWinEndIndex = FreqWinEndIndex + 1;
end
% Define the Low Band frequency window
FreqLBWinStartIndex = 1;
FreqLBWinEndIndex = FreqLBWinStartIndex;
while (FreqVec(FreqLBWinEndIndex) <
WaveParam.LFreqBand.End)
   FreqLBWinEndIndex = FreqLBWinEndIndex +
1;
end
%Define the High Band frequency window
FreqHBWinStartIndex = FreqLBWinEndIndex+1;
FreqHBWinEndIndex = FreqHBWinStartIndex;
```

```
while (FreqVec(FreqHBWinEndIndex) <
WaveParam.HFreqBand.End)
   FreqHBWinEndIndex = FreqHBWinEndIndex +
1;
end


% Create the PSD frequency plot figure
hFreqFig1 = figure('Name','Spectrum
Plot','NumberTitle','off','Visible','on');
% Create the PSD[db] frequency plot figure
%hFreqFig2 = figure('Name','Spectrum
Plot','NumberTitle','off','Visible','on');
% Define the spectrum algorithm
SpectralAlgorithIndex =
get(handles.popupmenuSpectrumAlgorithm,
'Value');
EnergyRatio = [];
SingingPowerRatio = [];
SpectralAlgorithIndex
switch SpectralAlgorithIndex
   case 1
      % FFT
      Index = 1;
      WaveParam.TimeWin.End
      while WinEnd <= WaveParam.TimeWin.End
         % Apply the window
         WaveDataWin =
WaveData.data(WinStart:WinEnd,WaveParam.Cha
nnels.ChannelSelect);
         WinFunc = hanning(WinWidth);
         WaveDataWin =
WinFunc'.*WaveDataWin';
         % Compute the FFT
         xdft = fft(WaveDataWin,WinWidth);
         % Extract first half of FFT
         xdft = xdft(1:ceil(WinWidth/2));
        % psdx = (1/(Fs*WinWidth)).*abs(xdft).^2;
        % psdx(2:end-1) = 2*psdx(2:end-1);
         % Normalize the FFT amplitudes using Fs
         psdx = xdft.*conj(xdft)/WinWidth;
         psdxDb = 10*log10(psdx);
         % Plot the power spectrum

plot(FreqVec(FreqWinStartIndex:FreqWinEndInde
x),psdxDb(FreqWinStartIndex:FreqWinEndIndex))
; grid on;
         hold on;
        % Calculate the Energy Ratio
         EnergyLB =
2*sum(psdx(FreqLBWinStartIndex:FreqLBWinEn
dIndex))*DeltaF;
         EnergyHB =
2*sum(psdx(FreqHBWinStartIndex:FreqHBWinEn
dIndex))*DeltaF;
         EnergyRatio = [EnergyRatio; Index,
EnergyLB/EnergyHB];
         % Calculate the Sing Power Ratio
         SingingPowerLB =
max(psdx(FreqLBWinStartIndex:FreqLBWinEndI
ndex));
         SingingPowerHB =
max(psdx(FreqHBWinStartIndex:FreqHBWinEndI
ndex));
         SingingPowerRatio =
[SingingPowerRatio;Index,
SingingPowerLB/SingingPowerHB];
         WinStart =
WinStart+WaveParam.SlideWin.Inc;
         WinEnd =  WinStart+WinWidth-1;
         Index = Index+1;
      end
      EnergyRatioAvg =mean(EnergyRatio(:,2));
      SingingPowerRatioAvg =
mean(SingingPowerRatio(:,2));
      FigureTitle = 'Spectrum using FFT';
   case 2
      % Periodogram
      [psdestx,Fxx] =
periodogram(WaveDataWin,hanning(NSamples),N
Samples,Fs);
      psdx = psdestx;
      FreqVec = Fxx;
      psdSquared = (psdx/Fs).*conj(psdx/Fs);
      EnergyLB = 2*sum(psdSquared( FreqVec>=0
& FreqVec<=2000,: ))*DeltaF;
      EnergyHB = 2*sum(psdSquared(
FreqVec>=2000 & FreqVec<=4000,: ))*DeltaF;
      FigureTitle = 'Spectrum using Periodogram';
      EnergyRatio = EnergyLB-EnergyHB;
      SingingPower = EnergyHB/EnergyLB;
end
% Calculate the Energies
title(FigureTitle);
xlabel('Frequency (Hz)');
ylabel('Power/Frequency (dB/Hz)');
% Create the Energy Ratio plot figure
hFreqFig2 = figure('Name','Energy Ratio
Plot','NumberTitle','off','Visible','on');
plot(EnergyRatio(:,1),EnergyRatio(:,2));
title('Energy Ratio');
xlabel('Window Index');
ylabel('Energy Ratio');
% Display the results
display('Energy Ratio:');
display(EnergyRatio(:,2));
display('Energy Ratio Average');
display(EnergyRatioAvg);
display('Singing Power Ratio:');
display(SingingPowerRatio(:,2));
display('Singing Power Ratio Average');
display(SingingPowerRatioAvg);



function [WaveData] = importfile(File)
%IMPORTFILE(FILETOREAD1)
```

```
% Imports data from the specified file
% FILETOREAD1: file to read
% Auto-generated by MATLAB on 02-Apr-2013
19:45:46
% Import the file
WaveData = importdata(File);
% Create new variables in the base workspace from
those fields.
vars = fieldnames(WaveData);
for i = 1:length(vars)
   assignin('base', vars{i}, WaveData.(vars{i}));
end



function editSlideWinWidth_Callback(hObject,
eventdata, handles)
% hObject    handle to editSlideWinWidth (see
GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)

% Hints: get(hObject,'String') returns contents of
editSlideWinWidth as text
%        str2double(get(hObject,'String')) returns
contents of editSlideWinWidth as a double


% --- Executes during object creation, after setting
all properties.
function editSlideWinWidth_CreateFcn(hObject,
eventdata, handles)
% hObject    handle to editSlideWinWidth (see
GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    empty - handles not created until after
all CreateFcns called

% Hint: edit controls usually have a white
background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
   set(hObject,'BackgroundColor','white');
end



function editSlideWinInc_Callback(hObject,
eventdata, handles)
% hObject    handle to editSlideWinInc (see
GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)
```

```
% Hints: get(hObject,'String') returns contents of
editSlideWinInc as text
%        str2double(get(hObject,'String')) returns
contents of editSlideWinInc as a double


% --- Executes during object creation, after setting
all properties.
function editSlideWinInc_CreateFcn(hObject,
eventdata, handles)
% hObject    handle to editSlideWinInc (see
GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    empty - handles not created until after
all CreateFcns called

% Hint: edit controls usually have a white
background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'),
get(0,'defaultUicontrolBackgroundColor'))
   set(hObject,'BackgroundColor','white');
end


% --- Executes on slider movement.
function sliderWinStart_Callback(hObject,
eventdata, handles)
% hObject    handle to sliderWinStart (see GCBO)
% eventdata  reserved - to be defined in a future
version of MATLAB
% handles    structure with handles and user data
(see GUIDATA)
% Hints: get(hObject,'Value') returns position of
slider
%        get(hObject,'Min') and get(hObject,'Max') to
determine range of slider
global WaveParam
% Get the time window start
XCur =  get(hObject,'Value');
% Update the time window start cursor
WaveParam.TimeWin.StartCur.XData = [XCur;
XCur];
set(findobj('Tag','lineWinStart'),'XData',WavePara
m.TimeWin.StartCur.XData);
% Update the time window start value
StartIndex = 1;
while (WaveParam.TimeLine.TimeVec(StartIndex)
< XCur)
   StartIndex = StartIndex + 1;
end
WaveParam.TimeWin.Start = StartIndex;

% --- Executes during object creation, after setting
all properties.
function sliderWinStart_CreateFcn(hObject,
eventdata, handles)
% hObject    handle to sliderWinStart (see GCBO)
```

% eventdata reserved - to be defined in a future version of MATLAB
% handles empty - handles not created until after all CreateFcns called
% Hint: slider controls usually have a light gray background.
if isequal(get(hObject,'BackgroundColor'), get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor',[.9 .9 .9]);
end

% --- Executes on slider movement.
function sliderWinEnd_Callback(hObject, eventdata, handles)
% hObject handle to sliderWinEnd (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)
% Hints: get(hObject,'Value') returns position of slider
% get(hObject,'Min') and get(hObject,'Max') to determine range of slider
global WaveParam
% Get the time window end
XCur = get(hObject,'Value');
% Update the time window end cursor
WaveParam.TimeWin.EndCur.XData = [XCur,XCur];
set(findobj('Tag','lineWinEnd'),'XData',WaveParam.TimeWin.EndCur.XData);
% Update the time window end value
EndIndex = 1;
while (WaveParam.TimeLine.TimeVec(EndIndex) < XCur)
    EndIndex = EndIndex + 1;
end
WaveParam.TimeWin.End = EndIndex;

% --- Executes during object creation, after setting all properties.
function sliderWinEnd_CreateFcn(hObject, eventdata, handles)
% hObject handle to sliderWinEnd (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles empty - handles not created until after all CreateFcns called
% Hint: slider controls usually have a light gray background.
if isequal(get(hObject,'BackgroundColor'), get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor',[.9 .9 .9]);
end

function Init(hObject)
global WaveData WaveParam
% Initialize the time window parameters
WaveParam.TimeWin = struct('Start',WaveParam.TimeLine.Start,'End',WaveParam.TimeLine.End);
% Initialize the time window start cursor parameters
XCur = WaveParam.TimeLine.TimeVec(WaveParam.TimeWin.Start);
WaveParam.TimeWin.StartCur = struct('XData',[XCur; XCur],'YData',[-1; 1]);
% Initialize the time window end cursor parameters
XCur = WaveParam.TimeLine.TimeVec(WaveParam.TimeWin.End);
WaveParam.TimeWin.EndCur = struct('XData',[XCur; XCur],'YData',[-1; 1]);
% Initialize the sliding window offset and width parameters in seconds
SlideWinOffsetSec = 0.0;
SlideWinWidthSec = 0.021;
SlideWinIncSec = 0.010;
WaveParam.SlideWinSec = struct('Offset',SlideWinOffsetSec,'Width',SlideWinWidthSec,'Inc',SlideWinIncSec);
% Calculate the sliding window offset and width index values
SlideWinWidth = ceil(SlideWinWidthSec*WaveData.fs);
SlideWinWidth = 1024 % Overide: Use 1024
SlideWinInc = ceil(SlideWinIncSec*WaveData.fs);
WaveParam.SlideWin = struct('Offset',WaveParam.TimeWin.Start,'Width',SlideWinWidth,'Inc',SlideWinInc);
set(findobj('Tag','editSlideWinWidth'),'String',WaveParam.SlideWin.Width);
set(findobj('Tag','editSlideWinInc'),'String',WaveParam.SlideWin.Inc);
% Initialize the frequency window parameters
WaveParam.FreqWin.Width = 6000;
% Initialize the Low Frequency Band and High Frequency Band parameters
WaveParam.LFreqBand = struct('Start',0.0,'End',2000);
WaveParam.HFreqBand = struct('Start',2000,'End',4000);

% --- Executes on selection change in popupmenuChannelSelect.
function popupmenuChannelSelect_Callback(hObject, eventdata, handles)
% hObject handle to popupmenuChannelSelect (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)

```
% Hints: contents = cellstr(get(hObject,'String')) returns popupmenuChannelSelect contents as cell array
%        contents{get(hObject,'Value')} returns selected item from popupmenuChannelSelect

% --- Executes during object creation, after setting all properties.
function popupmenuChannelSelect_CreateFcn(hObject, eventdata, handles)
% hObject    handle to popupmenuChannelSelect (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: listbox controls usually have a white background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'), get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end


% --- Executes on selection change in popupmenuSpectrumAlgorithm.
function popupmenuSpectrumAlgorithm_Callback(hObject, eventdata, handles)
% hObject    handle to popupmenuSpectrumAlgorithm (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: contents = cellstr(get(hObject,'String')) returns popupmenuSpectrumAlgorithm contents as cell array
%        contents{get(hObject,'Value')} returns selected item from popupmenuSpectrumAlgorithm


% --- Executes during object creation, after setting all properties.
function popupmenuSpectrumAlgorithm_CreateFcn(hObject, eventdata, handles)
% hObject    handle to popupmenuSpectrumAlgorithm (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns called

% Hint: popupmenu controls usually have a white background on Windows.
%       See ISPC and COMPUTER.
if ispc && isequal(get(hObject,'BackgroundColor'), get(0,'defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor','white');
end
```