LESSAC'S TONAL ACTION IN WOMEN'S VOICES AND THE "ACTOR'S FORMANT": A COMPARATIVE STUDY.

Marth Munro

Thesis submitted for the degree Philosophaie Doctor in General Linguistics and
Literary Theory at the Potchefstroom University for Christian Higher Education

Promoter: Prof. D.P Wissing
Assistant Promoter: Prof. A.M. Laukkanen
Assistant Promoter: Prof. K. Verdolini

2002

ABSTRACT

The purpose of this study is to investigate the Tonal NRG (previously known as the Tonal Action) of the Lessac Approach as a voice-building tool for the female voice in theatre. It provides an overview of existing scholarly writings on the pedagogical, physiological and acoustical qualities of the Lessac Approach. It reflects on current literature about the characteristics of good voice quality and especially the actor's formant. The empirical research on voice building in this study demonstrates that a randomised pre/post test/control group quasi-experimental design was used and is fourfold in dimension: it deals with a control group, a test group with 14 contact hours, a test group with 28 contact hours and a test group who had an extensive six-week workshop. Other variables that are reflected on are language, teacher and training methodology specificity. Investigative procedures include a questionnaire, various means of acoustic analysis and a perception panel. Where applicable, inferential statistics were done on the data. Results of the investigation are compared with existing, parallel research outputs.

This study indicates very strongly that the Tonal NRG of the Lessac Approach influences the female voice positively irrespective of language, teacher and training methodology specific parameters.

OPSOMMING

Die doel van hierdie studie is om 'n ondersoek te loods aangaande die "Tonal NRG" van die "Lessac Approach" as 'n stembousisteem vir die vrouestem in teater. Die studie verskaf 'n oorsig van bestaande vakkundige geskrifte rakende die pedagogiese, fisiologiese en akoestiese kwaliteite van die "Lessac Approach." Dit ondersoek die standpunte gestel in bestaande literatuur aangaande goeie stemkwaliteit en veral die begrip "actor's formant." Die empiriese navorsing rakende stembou in hierdie studie demonstreer die gebruik van 'n willekeurige, voor-/natoets-/kontrolegroep, kwasi-eksperimentele ontwerp. Vier groepe is gebruik, naamlik 'n kontrolegroep, 'n toetsgroep met 14 kontakure, 'n toetsgroep met 28 kontakure en 'n toetsgroep wat 'n intensiewe werkswinkel van ses weke deurloop het. Verdere veranderlikes waaraan aandag gegee word is taal-, onderwyser- en onderrig-metodologieverskille. Die ondersoeksinstrumente sluit in 'n vraelys, verskillende akoestiese ontledingswyses en 'n persepsiepaneel. Waar moontlik is inferensiële statistiese ontledingsmetodes gebruik. Resultate van die ondersoek is vergelyk met die uitkomste van bestaande parallelle navorsingsuitsette.

Daar is sterk aanduidings in die studie dat die "Tonal NRG" van die "Lessac Approach" positief inwerk op die bou van die vroulike stem nieteenstaande talige, onderwyser- en onderrig-metodologiese verskille.

Acknowledgements

- My promoter Prof. D.P Wissing for making this study a rich, learning experience, and for being prepared to venture into fields unknown and territories yet to be explored.
- Prof. Anne-Maria Laukkanen from the University of Tampere, Finland, for guidance in Chapter 3 specifically, and about the acoustic analysis of voice in general.
- Prof. Katherine Verdolini, from the University of Pittsburgh, USA, for guidance in Chapter 2 specifically and about the physiology of voice in general.

- All the wonderful female participants in this study – may your voices be heard far and wide.
- Emily Groenewald, for practical guidance about analysis, and for all the stability a researcher requires.
- Izak Smit and Joanne Clark from the Department of Statistical Services, Pretoria Technikon for significant help and guidance with the dreaded numbers.
- Rina White from the library, Arts Faculty, Pretoria Technikon who were always willing to "do a search."
- Prof. Munro for guidance with academic writing style and proof reading.
- Jan Louw for initial help with the acoustical analysis and interpretation.
- Charmaine Morris from Pretoria Technikon for computer assistance
- Ulrike Janke for technical support.
- Elsa, Emma and Hestelle for administration.

- Allan and Andrew for support and patience.
- My parents for support, encouragement, and child-care.
- Our Eternal Father.

And, from whom this study originated, and for the visionary that he is:

- Arthur Lessac for sharing the Lessac Approach with us, and his involvement in this study.

Dedication:

To my family

and the Lessac community.

Table of contents

Abstract ii
Opsomming ii
Acknowledgements iii
Dedication iv

Table of contents v
Illustrations viii

Chapter 1. Introduction 1

1.1. Contextualization 1
1.2. Stating the problem 7
1.3. Aims and objectives 8
1.4. Central thematic argument 8
1.5. Method of investigation 9
  1.5.1. An analysis of the literature 9
  1.5.2. Empirical research 9
    1.5.2.1. Design 9
    1.5.2.2. Method 9
    1.5.2.3. Instrumentation 10
    1.5.2.4. Participants 11
    1.5.2.5. Ethical considerations 12
    1.5.2.6. Data collection , procedure and analysis 12

Chapter 2. Lessac's Tonal NRG 14

2.1. Introduction and overview 14
  2.1.1. Arthur Lessac: the person behind the Lessac Approach 14
    2.1.1.1. The Lessac Approach: an overview 15
    2.1.2.1. The three NRG's 20
      2.1.2.1.1. Consonant NRG 20
      2.1.2.1.2. Tonal NRG 21
      2.1.2.1.3. Structural NRG 32
  2.1.3. The Tonal NRG and existing literature in Voice Science 35
    2.1.3.1. Critique of the Lessac Approach as a teaching tool 35
    2.1.3.2. Critique on the physiological input of the Lessac Approach 38
    2.1.3.3. Critique on the acoustical outputs of the Lessac Approach 40

Chapter 3. The acoustic characteristics of the "projected" voice 45

3.1. Introduction and overview 45
3.2. The "Singer's Formant" 48

3.3. The “Actor's Formant” 51

Chapter 4. Empirical research 57

4.1. Problem statement 57
4.2. Empirical research aim 58
4.3. Section one: American Test Group 58
    4.3.1. Aim 58
        4.3.1.1. Sub-aim 58
    4.3.2. Participants 59
    4.3.3. Ethical considerations 59
    4.3.4. Training period 59
    4.3.5. Training process 59
    4.3.6. Data collection 59
        4.3.6.1. Recordings 59
        4.3.6.2. Tasks expected from participants 60
    4.3.7. Data analysis and processing 61
    4.3.8. Results and discussion 61
4.4. Section two: Control Group 67
    4.4.1. Aim 67
        4.4.1.1. Sub-aim 67
    4.4.2. Participants 67
    4.4.3. Ethical considerations 68
    4.4.4. Training period and process 68
    4.4.5. Data collection 68
        4.4.5.1. Recordings 68
        4.4.5.2. Tasks expected from participants 69
    4.4.6. Data analysis and processing 69
    4.4.7. Results and discussion 78
4.5. Section three: Test Group S 81
    4.5.1. Aim 81
        4.5.1.1. Sub-aims 81
    4.5.2. Participants 81
    4.5.3. Ethical considerations 82
    4.5.4. Training period 82
    4.5.5. Training process 82
    4.5.6. Data collection 82
        4.5.6.1. Recordings 82
        4.5.6.2. Tasks expected from participants 83
    4.5.7. Perceptual evaluation: design 84
        4.5.7.1. Data analysis and processing of the perceptual evaluation 85
        4.5.7.2. Perceptual evaluation results 87
    4.5.8. Acoustic investigation 93
        4.5.8.1. Design, data analysis and processing of the acoustic investigation 93

4.5.8.2. Acoustic investigation results 95
4.5.9. Discussion 114
4.6. Section 4: Test Group V 117
4.6.1. Aim 117
4.6.1.1. Sub-aims 117
4.6.2. Participants 117
4.6.3. Ethical considerations 118
4.6.4. Training period 118
4.6.5. Training process 118
4.6.6. Data collection 118
4.6.6.1. Recordings 118
4.6.6.2. Tasks expected from participants 119
4.6.7. Perceptual evaluation: design 120
4.6.7.1. Data analysis and processing of the perceptual evaluation 121
4.6.7.2. Perceptual evaluation results 123
4.6.8. Acoustic investigation 129
4.6.8.1. Design, data analysis and processing of the acoustic investigation 129
4.6.8.2. Acoustic investigation results 131

Chapter 5. Summary, Conclusions and Shortfalls 154

5.1. Restating the field of investigation 154
5.2 Structuring the discussion of the findings 155
5.3. Findings 156
5.3.1. Acoustical findings 156
5.3.2. Acoustic findings in correlation to existing literature 165
5.3.3. Perceptual findings and existing literature 168
5.3.4. Physiological findings and existing literature 169
5.4. Conclusion 170
5.5. Shortfalls of this study 172

Bibliography 174

Addendum 181

Illustrations

Chapter 2

Figure 1: Lessac's Leitmotif 19

Chapter 4

Section one

Figure 1: An example of the pre- and post- recordings of the Y-buzz of one participant 62
Figure 2: An example of the pre- and post- recordings of the +Y-buzz of one participant 63
Figure 3: An example of the pre- and post- recordings of the Call of one participant 64
Figure 4: An example of the pre- and post- recordings of the English Phrases and Call words of one participant 65

Section two

Figure 5: The LTAS of the first and second recordings of the Y-buzz of all four participants in the Control Group 70
Figure 6: The LTAS of the first and second recordings of the +Y-buzz of all four participants in the Control Group 70
Figure 7: The LTAS of the first and second recordings of the Call of all four participants in the Control Group 71
Figure 8: The LTAS of the first and second recordings of the English Phrases and Call words of all four participants in the Control Group 71
Table 1: An excerpt of the LTAS of participant CA 72
Figure 9: An LTAS representation of the 1st and 2nd recording of the Y-buzz of participant CA 73
Table 2: An excerpt of the approximation of the highest amplitude for the frequency window of 100 Hz of the Y-buzz of participant CA 73
Table 3: An excerpt of the subtraction process: dB of 2nd recording minus dB of 1st recording, minus difference between dB of F0 1st and F0 2nd recording for all 100 Hz windows. Participant CA: Y-buzz 74
Figure 10: The difference in relation to F0 for all the 100 Hz frequency windows, including F0. Participant CA: Y-buzz 75
Figure 11: The difference in relation to F0 for all the 100 Hz frequency windows, including F0. Y-buzz average of all participants in Control group 76
Figure 12: The difference in relation to F0 for all the 100 Hz frequency windows, including F0. +Y-buzz average of all participants in Control group 76

Figure 13: The difference in relation to F0 for all the 100 Hz frequency windows, including F0. Call average of all participants in Control group 77
Figure 14: The difference in relation to F0 for all the 100 Hz frequency windows, including F0. English phrases and Call words average of all participants in Control group 77

Section three

Table 1: T-test for independence, Test Group S 87
Table 2: Group S, Y-buzz: Table of phase by score 87
Table 3a: Cross tabulation in preparation for chi-square of the Lessac explorations 89
Table 3b: The chi-square results of the Lessac explorations 89
Table 4a: Cross tabulation in preparation for chi-square of the Lessac applications 90
Table 4b: The chi-square results of the applications 90
Table 5a: Group S, Calls African languages: Table of phase by score 91
Table 5b: Group S, Calls, Afrikaans: Table of phase by score 92
Table 5c: Group S, Calls, English: Table of phase by score 92
Table 6: Cross tabulation table: Group S, participant AA, all sound modes 93
Figure 1: The average LTAS of the pre- and post- training of the Y-buzz of Test Group S 96
Figure 2: The average LTAS of the pre- and post- training of the +Y-buzz of Test Group S 96
Figure 3: The average LTAS of the pre- and post- training of the Calls of Test Group S 97
Figure 4: The average LTAS of the pre- and post- training of the English phrases and Call words of Test Group S 97
Figure 5: The average LTAS of the pre- and post- training of the First language texts of Test Group S 98
Table 7: An extract example of the processing of the data. Test Group S, participant AA, +Y-buzz 99
Figure 6: Bar graph of mean/median differences of the Y-buzz. 100
Figure 7: Bar graph of mean/median differences of the +Y-buzz. 101
Figure 8: Bar graph of mean/median differences of the Calls. 101
Figure 9: Bar graph of mean/median differences of the English phrases and Call words. 102
Figure 10: Bar graph of mean/median differences of the first language texts. 102
Figure 11: The LTAS and Bar Graph of the Test Group S: Y-buzz 103
Figure 12: The average LTAS of the pre- and post-recordings of the Calls of Test Group S, First language African 107
Figure 13: The average LTAS of the pre- and post-recordings of the first language texts, Test Group S, First language Afrikaans 109
Figure 14: A bar graph demonstrating the percentage increase of the

energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group S: + Y-buzz 111
Figure 15: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group S: Calls 112
Figure 16: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group S: English phrases and Call words 113
Figure 17: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group S: First language texts 114

Section four

Table 1: T-test for independence, Test Group V 123
Table 2: Group V, Y-buzz: Table of phase by score 124
Table 3a: Cross tabulation in preparation for chi-square of the Lessac explorations 125
Table 3b: The chi-square results of the Lessac explorations 126
Table 4a: Cross tabulation in preparation for chi-square of the Lessac applications 126
Table 4b: The chi-square results of the applications 126
Table 5a: Group V, +Y-buzz, African languages: Table of phase by score 127
Table 5b: Group V, +Y-buzz, Afrikaans: Table of phase by score 128
Table 5c: Group V, +Y-buzz, English: Table of phase by score 128
Table 6: Cross tabulation table: Group V, participant AA, all sound modes 128
Figure 1: The average LTAS of the pre- and post- training of the Y-buzz of Test Group V 132
Figure 2: The average LTAS of the pre- and post- training of the +Y-buzz of Test Group V 132
Figure 3: The average LTAS of the pre- and post- training of the Calls of Test Group V 133
Figure 4: The average LTAS of the pre- and post- training of the English phrases and Call words of Test Group V 133
Figure 5: The average LTAS of the pre- and post- training of the First language texts of Test Group V 134
Table 7: An extract example of the processing of the data. Test Group V, participant AA, +Y-buzz 135
Figure 6: Bar graph of mean/median differences of the Y-buzz. 136
Figure 7: Bar graph of mean/median differences of the +Y-buzz. 136
Figure 8: Bar graph of mean/median differences of the Calls. 137
Figure 9: Bar graph of mean/median differences of the English phrases and Call words. 138
Figure 10: Bar graph of mean/median differences of the first language

texts. 138
Figure 11: The LTAS and Bar Graph of the Test Group V: Y-buzz 139
Figure 12: The LTAS of the Calls of the Test Group V, First language African 144
Figure 13: The average LTAS of the +Y-buzz, Test Group V, first language English 145
Figure 14: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: Y-buzz 147
Figure 15: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: +Y-buzz 148
Figure 16: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: Calls 149
Figure 17: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: English phrases and Call words 150
Figure 18: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: First language texts 151

Chapter 5

Table 1: Y-buzz mode: A comparative table providing results that demonstrate tendencies 158
Table 2: +Y-buzz mode: A comparative table providing results that demonstrate tendencies 159
Table 3: Call mode: A comparative table providing results that demonstrate tendencies 160
Table 4: English phrases and Call words mode: A comparative table providing results that demonstrate tendencies 161
Table 5: First language texts mode: A comparative table providing results that demonstrate tendencies 162
Table 6: Most consistent changes observed in all training groups for all modes. 163

Chapter 1

Introduction

1.1. Contextualization

The fundamental task of an actor/actress in the theatre is to communicate. To do this the artist uses every means at his/her disposal to complete the task successfully. In essence, this task would employ the auditory, the visual and the kinesthetic dimensions of performance. The actor operates in and through time, and in and through space, and follows a set series of sign-systems in the process of encoding the goals of the performance.

One measure of the success of the process lies in the ability of the actor/actress to make the spoken sound (both simply as sound and also as an encoded sign) auditively available to an audience. Such an audience is, inevitably, at some distance from the actor/actress. Consequently, every actor/actress strives towards audibility/vocal projection (Munro, Leino & Wissing, 1996:26) on stage. The ultimate aim is thus the projection of the vocal sound by the actor/actress, without the loss of flexibility in the voice (Acker, 1987:77; Linklater ,1976: 2 & 3). Such a loss of flexibility would impede the clarity of the communicative act. One of the "tools" that the actor/actress requires, therefore, is a flexible vocal instrument that has the ability to project the created (and encoded) sound over some distance.

A number of divergent yet interrelated factors contribute to the performer's ability to play several different characters in several different genres (Hanson, 1997:1). These may include aspects such as:

- Body/voice integration, where the actor's entire physical and vocal instrument is trained to function optimally and holistically;
- Voice-building, where the vocal apparatus becomes "stronger and more pliable;"

- Speech-building and pronunciation, where articulation for the encoding practice for meaning is enhanced;
- Connecting voice and emotions, where the encoded message is layered with the basic theatrical demands of "emotional living:"
- The use of voice in violence on stage, where the inherent conflictual nature of drama as an art form is enhanced and "healthily integrated" into the performer's instrument;
- Text explorations, where subtleties and nuances of the text are vocally explored and discovered;
- Characterization through voice and body, as the specific theatrical demands of character presentation are carried through into the encoded instrument (body and voice);
- Accent acquisition and accent reduction,[1] where regionalities for character are encoded in idiolects and dialects;
- Coping with different performing spaces with different acoustics, in which actors learn to adapt the vocal demands to the spaces in which they perform;
- And, specifically in the South African case, orientations to the different language rhythms, sound placements, and other vocal demands (such as the Xhosa "click").

Taking all of these into account, the training of the actor's/actress's voice is a multi-faceted task. It is expected of the voice and speech teacher (and therefore artist) to focus on several different aspects that will eventually contribute to the performer's ability to execute with "sufficient intensity and loudness to be adequately heard and understood" (Hanson, 1997:1). This "intensity and loudness" requires a significant amount of voice building to take place.

---

[1] The author is aware of the political and educational controversies that surround the problems of accent reduction. The comprehension, across different first language groups, of the meaning of an utterance that is presented in different accents is an avenue of research that needs to be explored and exploited in the drama departments.

"Voice-building" might tentatively be defined as building the voice to deal with the rigours of performance. Laukkanen (1995:11) states that the muscular functions of voice production have to be strengthened and improved in such a way that it will result in an "improved voice and/or phonatory quality." In much the same way as an athlete might not only learn the skills of pole-vaulting, for example, but he or she would have also to build and condition the body to carry out the skills required in the vaulting process. The two fundamental criteria for the process of pole-vaulting, therefore, can be seen as the preparation of the body to perform the tasks at hand optimally, effectively and without waste or loss of effort, and the maintenance of the body as a healthy unit to perform the required task repeatedly. In similar fashion, the voice of the actor is required to work optimally in the process of performance, but is also required to be able to repeat the task on a regular basis. For this to occur, the voice needs to be built and strengthened (Barton & Dal Vera, 1995:79).

However, it must be remembered that this "built voice" needs also to satisfy the demands of quality. Good voice quality (sustainable over periods of time and repeatable) for theatre specifically may be "defined as a result of an optimal use of the vocal organ in order to establish the maximum possible acoustic output by minimal muscular effort" (Laukkanen, 1995:18), to which will be added the aesthetic dimension, which is culturally bound. This is thus the main goal in voice building.[2]

There are several recognized theatre voice-training systems, which claim to provide the actor/actress with such "good voice tools" (Berry, 1973; Lessac, 1967; Linklater, 1976; Martin, 1991; Mulholland, 1984). Most of these systems claim to be effective *in praxis*. Nevertheless, very few of these systems have been tested *scientifically,*[3] or in a scholarly way.

---

[2] At this point in the thesis I would simply define "most efficient acoustic output" as "projection". This is, of course, a very rough definition and the concept will be attended to in much deeper and clearer detail in Chapter 3.

[3] I wish to make a term clear here. The bulk of literature uses the term "scientific" to explain its procedures. I wish to use the term "scholarly" as an alternative to, if not an improvement on, the term "scientific." The reason for this is that the word "scientific" unfortunately is strongly associated with the "logical-positivist" research paradigm and method with it, whereas I find that the word "scholarly" anticipates, precedes and extends this term to include all aspects where scholars have applied their respective minds to particular

Very little was previously scientifically known about voice-building. In this light the voice systems should be acknowledged for their attempts to address voice-building. Most teachers, relying on a good ear, opted for the use of metaphors to explain, or imitation to guide, their students to better voice production. Typical metaphors are: "Putting it in the mask, putting more metal into the voice, allowing the voice to come from the centre of the body, experiencing a fountain of breath," etc. The dangers of these metaphorical teachings, especially in a multi-lingual and multi-cultural country are obvious. Although voice and speech teachers always have good intentions in their teachings some actors may actually hurt their voices in their attempts to do what they understood the metaphor to be, or in imitating a voice quality which is completely different from their own.

Only anecdotal evidence of this exists, yet the potential dangers are very real. This project will assist in the development of a system that is rooted in the scholarly, and that will assist in circumventing the potential problems that the metaphor brings to teaching voice.

During the past thirty years, the science of voice started to blossom, especially the science of the singer's voice. Here Vennard (1967) and Rose (1978) need to be mentioned. Yet, even in their time the development of science technology did not assist their thinking enough to really make an impact on the teaching of voice. The science of the singing voice came to full fruition only with scholars like Sundberg (1987), Sataloff (1991), Titze (1994b), Richard Miller (1986), Nair (1999) and Donald Miller (2000), to name but a few. These scholars have concentrated on investigating the physiological and acoustical functioning of the voice, with specific applications to the pedagogical practice.

Put into perspective, the healthy use of the voice may be conceived of as a continuum, with the concepts of Western classical singing (vocal manipulation) at the one end of the

---

problems. "Scholarly" would therefore include "scientific" completely, but not vice versa. The word scientific will be used only when referring to a "logical positivist" research paradigm.

continuum, and conversational speaking at the other. Given this concept, the "actor's speaking voice" may rest somewhere between the two "extremes."[4] Nevertheless, the study will indicate the fundamental differences and similarities between the singing voice and the speaking voice. Scholars of the performer's speaking voice include Acker, Laukkanen, Leino, Nawka, Raphael and Scherer. Reference will be made to them in Chapter 2.

One of the founding scholars on the development of the speaking voice, with specific reference to the actor's voice, is Arthur Lessac. Lessac (1997a: 9) claims that part of his system is the building of a good voice. Here he refers specifically to the Tonal NRG[5] part of the work and states that it will enhance audibility (1997a:139; 1967:20).[6] To the author's knowledge the Lessac Approach is the only theatre voice system that has been adopted by speech therapists. It is nowadays a fully-fledged part of speech therapy and it is called "Lessac Madsen Resonant Voice Therapy" (Verdolini-Marston, et al., 1995).[7]

Lessac acknowledges that he begins his voice work at the physical end of voice and speech.

> If we think about the human body as a superb instrument, we can proceed from the premise that careful observation of how the body *wants* to function - how it would function in the absence of adverse conditioning - is a good guide to the production of fine tone and excellent sounds. When the body produces excellent tones, the voice is not throaty, nasal, of forced; it is produced and resonated effortlessly. It has **stentorian**, resonant qualities and projection, full pitch range, and rich, warm, colorful timbre. (Lessac, 1997a: 9. accents Lessac's own)

It is, though, of great importance to Lessac that one does not divorce the voice-building work from the other aspects of the system. He takes care to stress that the "human voice is more than an extension of the pure mechanics of voice and speech" (ibid). According to Lessac (1997b: 17) "Voice and speech training is body training and body training is

---

[4] This study will be investigating the healthy voice, and will not attempt to address areas that are the domain of pathology.

[5] NRG is a specific abbreviation in the Lessac Approach and refers to a) "Neurological Regenerative Growth" and b) to energy.

[6] The Tonal NRG will become of fundamental importance in the following chapters. At this point it is necessary to define this basically as "tonal energy" that uses the focus of optimal vowel sounds to enhance vocal health and the carrying power of the voice.

[7] This therapy will be referred to in greater detail in Chapter 2 of this work.

language/communication training, as well as bio-neuro-physic 'heightened sensitivity' training." This concept also reflects very strongly in the work of other well-known voice systems like those of Berry (1997:25), Linklater (1997:4) and Rodenburg (1997:38). However, the specific aim here is to investigate whether the Tonal NRG of the Lessac Approach will enhance the voice quality of the female actor's voice.

With particular reference to the acting voice, scholars like Laukkanen (1995), Leino (1993) and Nawka, Anders, Cebulla & Zurakowski (1997) are leading the way to a better understanding of the acoustic science of voice building for the theatre. Similarly, previous research done by Munro et al. (1996:25-36) strongly indicates that Lessac's Tonal "Action" (or Tonal NRG) may indeed enhance the formant patterns -- particularly the relationship and amplitude of the third, fourth and fifth formants, which represent the projection capabilities, or audibility, of the produced voice. The tests were conducted specifically on the male voice to correlate with the findings of Leino (1993:206-210). This research thus argued that the Tonal NRG of the Lessac Approach does improve audibility/projection in the trained male voice.

The female voice and actress have to all intents and purposes been marginalized in scholarship of the speaking voice for acting. The result has been that, whereas attempts have been made to account for the development of vocal projection in male training, female training is contingent upon an extrapolation of this scholarship. Consequently she is forced to rely on systems that seem to help towards audibility of the male voice, without a scientific exploration of these systems, relative to the audibility of female voices.

Thus it becomes necessary to investigate whether the Tonal NRG of the Lessac Approach, as a system developed to train actors in the enhancement of audibility and projection of the voice, may be used effectively in the training of the female voice. The way of measuring such effectiveness is by creating a graphic representation, using LTAS (Long-Term Amplitude Spectrum), of the formants of the sound created. Leino has done some research on what the formant patterns of an acceptable, "good," audible, female actor's voice would graphically resemble or approximate.[8] The bulk of the Leino analysis is presented in the form of LTAS representations of vocal formants. Leino's results show marked clustering around the third, fourth and fifth formants in voices that are deemed to have been projected well.

Leino's work provides the primary material for a comparison between the formant patterns of a Lessac trained female voice and the accepted formant patterns for a "good" female voice.

The training of the human voice, however, is problematic, and is constrained by, amongst other things, the following:

- The extent to which the method of training, and the "successful" outcomes are dependent upon particular cultural or national parameters. In this case the cultural determinants of the society in which a particular sound arises might have stipulated certain "aesthetic" values to certain sound patterns.
- The extent to which the perceived "natural ability" of a voice to project is actually contained in the particular language of the user. It is conceivable that certain languages appear to be more "projectable" than others (e.g. Italian and English).
- The extent to which the vagaries or idiosyncrasies of a single teacher might determine the projected outcomes of the training. In this case one is concerned about the teacher's "subjective" influence.[9] In correlation with this one also has to acknowledge the fact that different training methodologies (i.e. intensive workshop, short workshop, modules or a year long period situation) will influence any system of voice building.

---

[8] Leino presented a paper on this work at PEVOC 2001 and kindly agreed to make his findings available to me.

[9] Although there are no specific references here, it is a well-known fact among theatre practitioners that professional voice users "cluster" themselves around a specific voice teacher, and consider that voice teacher to be nearer the "correct" form of voice teaching and usage than other teachers. This occurs irrespective of whether the voice system of that specific teacher has been tested scientifically or not.

1.2. Stating the problem

Can the Tonal NRG of the Lessac Approach enhance the projection in the female actor's voice, irrespective of

- language or
- teacher and training methodology specificity,

so that the formant patterns of the female voice, analysed by means of an LTAS (Long-Term Average Spectrum) method and then graphically represented, will demonstrate such patterns as are usually seen in effective vocal projection in general?

1.3. Aims and objectives

The general aim of this study is to determine whether, by training the female voice in the Tonal NRG of the Lessac Approach so that the voice appears to project better, such a voice might indicate, following an LTAS analysis and the graphical representation of such an analysis, the specific formant enhancement that is associated with better vocal projection. This study further aims to ascertain whether such a formant enhancement might occur irrespective of

- language or
- teacher and training methodology specificity.

1.4. Central thematic argument

It is the central thematic argument of this study that the female voice, trained to produce the Tonal NRG of the Lessac Approach, will be enhanced in terms of audibility/projection, by such training. An LTAS analysis of such a trained voice, graphically represented, will indicate that the formant patterns of the female actor's voice will correlate with accepted patterns associated with better audibility/projection. The Tonal NRG of the Lessac Approach may show itself to be bound by neither

- language, or
- teacher and training methodology

1.5. Method of investigation

1.5.1. An analysis of the literature

In Chapter 2 the Lessac Approach, as described in The Use and Training of the Human Voice (1997a; 1967) will be discussed and analysed. Research which has already been undertaken in this and related fields (Acker, 1987; Hanson, 1997; Raphael & Scherer, 1987; Titze, 1994b; Verdolini, 1995,1998;) will be examined and integrated into the findings. For a scholarly explanation of the Tonal NRG, existing research concerning the science of voice will be incorporated (such as: Leino, 1993, 1995; McKinney, 1982; Miller, 1986; Sundberg, 1987).

In Chapter 3 the work of Timo Leino (1993, 1995, 2001), and peer researchers (Laukkanen, 1995; Nawka, et al., 1997; Wedin, et al., 1978), investigating the voice quality of male and female voices will be presented and analysed.

Chapter 4 will reflect on the empirical research of this study.

1.5.2. Empirical research

1.5.2.1. Design

A randomised pre/post test/control group quasi-experimental design will be used (Cohen & Manion, 1997; Mouton, 2001).

1.5.2.2. Method

Any research undertaken in this field of investigation needs to consider the following:

- The training needs to be done by a number of different Lessac teachers and in a number of training styles, so as to "disqualify" potential idiosyncratic differences.

- The base language (first language) of every group in training should be different, so as to combat perceptions of the "natural projectibility" of certain languages and with the different language groups.

By doing this, possible problems can be counteracted and may indicate that the Tonal NRG of the Lessac Approach is not language bound, nor indeed gender and teacher dominated, and also does not rely on any specific training methodology.

Consequently, the following methods will be employed:

1.5.2.3. Instrumentation

A **questionnaire** has been distributed during the Lessac Summit that took place 12-14 June 1998 in Swarthmore, Pennsylvania, USA to all the Certified Lessac Voice and Movement Teachers present. The purpose of the questionnaire was to ascertain the teachers' views on the Tonal NRG as an effective tool to overcome the lack of projection, or to enhance existing projection in the speaking voice in general, and in particular the female actor's voice. In this questionnaire data have been gathered concerning the "subjective" experience of Lessac teachers of the Tonal NRG. The interpretation of the answers to this questionnaire will reflect the perception of the Tonal NRG of the teachers in practice. This will be dealt with in Chapter 2.

The **LTAS analysis system** will be used to analyse the formant patterns of the pre- and post recordings of the test groups, as well as of the control group. These LTAS analyses will then be turned into graphic representations by using Microsoft Excel. The graphic representations of the pre- and post-recordings of each group will be compared. The outcomes of each group will be compared with the other outcomes. These comparisons will be dealt with in Chapter 4. In the case of Chapter 4, Sections 3 and 4, the numerical results of the LTAS will be drawn in to SAS, version 8 and various applicable statistical measurements will be done to ascertain the difference between the pre- and post-training recordings.

Chapter 4, Sections 3 and 4 will also have as part of the study the use of a perception panel. This is important seeing that it is eventually the "ear of the audience" that will "approve of, or prefer" the sound quality of the female actor's voice.

1.5.2.4. Participants

Different female groups[10] in two different countries will be trained in the Tonal NRG of the Lessac Approach. Two different countries have been identified, each using different languages as base language, each containing a certified Lessac teacher or teachers and each having a different culture or cultures:

- In the United States of America, Arthur Lessac himself supervised the training procedure that was in the form of an intensive six-week workshop, being taught by four certified Lessac teachers. American English was used as the base language. More definite information about the participants of this test group will be provided in Chapter 4, Section 1.
- At the time of planning this study in South African Marth Munro would teach three months as part of a yearlong course structure, focusing on the Tonal NRG. This class may have anything between four and eleven South African languages as first languages. The language of instruction will be English.[11] The actual contact hours will be 28. More definite information about the participants of this test group will be provided in Chapter 4, Section 3.
- At the time of planning this study in South Africa Marth Munro will teach a module at a tertiary institution on Tonal NRG. Afrikaans will be use as the base language. Classes will be conducted in English and Afrikaans. The actual contact hours will be 14. More definite information about the participants of this test group will be provided in Chapter 4, Section 4.

---

[10] In each group there should be no less than five participants. This number may increase, depending on the enrolment in each workshop/class/module.

[11] Although these three languages (American English, South African English and Afrikaans) may be argued to have similar, Germanic, language bases, there are no other trained and qualified Lessac teachers outside of these languages, except for Saschiko Nakagawa in Japan. The cost factor involved in connecting her to this study would be prohibitive. The researcher is of the opinion that the ideal situation would have been to access some of the other South African languages not included in the study, but the fact that there are no trained teachers in these languages excludes them.

- A group of South African women who will not undergo any kind of training will be used as a control group.[12] More definite information about the participants of this test group will be provided in Chapter 4, Section 2.

1.5.2.5. Ethical considerations

The women participating in the various sections of this research project all volunteered to be part of this study, after being assured that their identity will be protected. As such no direct reference will be made to the place, tertiary institution or person used in these various sections (Cohen & Manion, 1997:373).

1.5.2.6. Data collection, procedure and analysis

The following overall procedure will be followed:

- **Before** each training period in each country, the untrained voices will be recorded. This will be referred to as pre-training recordings.
- The control group will also have **before** recordings done. This will be referred to as first recordings.
- The recordings will be standardized in terms of the sounds uttered and the equipment used for the recordings. The typical Tonal NRG sounds that occur in the Lessac Approach and which will be used for these recordings are the Tonal NRG explorations, Y-buzz, +Y-buzz, Calls and the application thereof in the English Phrases and Call words. For the South African test groups, first language texts have been added.
- The recordings will then be sent to a central point, and converted for LTAS analysis. Further relevant analytical proceedings will be done as needed – these will be specifically discussed in the various sections as applicable.
- **During** the actual training, the Tonal NRG part of the Lessac Approach will be focused on the Y-buzz, +Y-buzz and Calls.

---

[12] I am aware of the fact that there are no control groups for each of the experimental groups. Optimally this is desirable. However, two factors militate against it. Firstly the cost of such control groups makes the study prohibitive. Secondly, this cost is even greater because, should the findings be positive, I would be ethically bound to train the control group, specifically in an educational setting, so that they would be "paid" for their contribution to the research and not have any potential impediment of quality of education.

- **Following** each **completed** training course, the trained voices will be recorded, producing the same set of sounds as used during the pre-training recording, and on the same equipment. These recordings will be referred to as post-training recordings.
- The findings of the before training (pre-) and the after training (post-) of each person; of each language group and of each group in training will be aggregated and compared.
- The control group will also undergo the **after** recordings. These recordings will take place on the same day as the first recordings but after a period of voice rest. This will be referred to as their second recordings. These will be compared to the first recordings of this group.
- Finally, the findings will be compared to existing literature specifically referring to the findings of Timo Leino and the like.

The data collection, procedure and analysis of each of the empirical sections will be discussed in detail when applicable.

By way of foreshadowing what is to come, the dissertation's framework is as follows:
Chapter 2 will provide an overview of the Lessac Approach. The results from a questionnaire done as part of this study will be reported. A reflection on existing research about the Approach will take place as part of the contextualization of the Lessac Approach.
Chapter 3 will reflect on existing literature on voice acoustics specifically as it relates to the concept of projection. The singer's and actor's formant will be addressed and discussed.
Chapter 4 will present the empirical research of this study. This chapter has four separate sections, which each address an aspect of the overall research design as discussed above.
Chapter 5 will be comparative of nature, referring to the four separate sections of Chapter 4, as well as the existing literature discussed in Chapters 2 and 3. Certain conclusive recommendations to future research possibilities will be made.

Chapter 2

Lessac's Tonal NRG

2.1. Introduction and Overview

This chapter will initially provide a short background to the work of Arthur Lessac and then proceed to a strategic summary of the approach[1] he developed. The chapter will further develop into an overview of relevant current literature. This overview will specifically focus on the research undertaken investigating the Lessac Approach where this Approach is used as a means towards optimal voice usage.

2.1.1. Arthur Lessac: the person behind the Lessac Approach

Arthur Lessac's name is synonymous with theatre voice work in the United States (Hanson, 1997:27) and further afield. In theatre voice training programmes there are currently three key names that are considered cardinal and innovative contributors to the fields of voice and speech for theatrical performance: Cicely Berry (Britain: Royal Shakespeare Company), Kristin Linklater (from Britain, now working in the U.S.A) and Arthur Lessac. In the next "generation" in the theatre voice training league one encounters the systems of Patsy Rodenburg (Britain: National Theatre Company), Roy Hart (France), Catherine Fitzmaurice (U.S.A), Edith Skinner, Jo Estill and others.

Lessac trained originally at the Eastman School of music, Rochester, New York. He furthermore studied anatomy and physiology of the human body and voice (Raphael, 1997:209) and did a clinical internship at St. Vincent's Hospital, New York (personal conversation, May 2001). Lessac retired as Professor Emeritus from the State University New York at Binghamton (Hampton & Acker, 1997). Through the years Lessac developed his own American voice and movement approach focussing on behaviour modification

---

[1] Lessac himself prefers not to refer to his work as a "method" or a "system", seeing that the usage of these words to describe the work may restrict the freedom and holistic nature of the work (personal conversation, May 2001).

(Raphael, 1997:209) as a teaching tool, and eventually started training teachers in this, now fully-fledged, kinaesthetically based mind-body system (Tabish, 1995:138). At the time of writing, although already 93 years of age, Lessac is still actively involved in teaching workshops of his approach during the American Summer months. During May 2001 he actively taught in South Africa. He is the author of two books: The Use and Training of the Human Voice: A bio-dynamic approach to Vocal Life (originally published in 1960, updated third edition published in 1997); and Body Wisdom (originally published in 1978, second edition 1990). This study will concern itself only with the work and philosophy discussed in The Use and Training of the Human Voice. [2]

2.1.1.1. The Lessac Approach: an overview

Lessac is adamant that his work is seen as a holistically integrated approach to voice, speech and movement training that will eventually promote health and healing (Lessac, 1997b: 19).[3] He suggests that body integration, or, as he calls it, "Body Wisdom encompasses all of Verbal Life and its verbal dynamics, as well as all the non-verbal communicating behaviours" (Lessac, in Hampton & Acker, 1997:17). What he means by this is that the body and the voice are parts of an entire communicating synergy that is not really distinguishable. This viewpoint echoes and parallels many of the other theatre voice training systems (for example Berry, 1973; Linklater, 1976) and with the current requirements of theatre training in general (Tabish, 1995).

With this in mind Lessac spends time in the beginning of his voice book discussing four concepts, which he deems cardinal to the "vocal life" of the actor. These concepts are "Body esthetics (Lessac's term), Inner harmonic sensing, Organic instructions to the body, and the 'familiar event' principle" (Lessac, 1997a: 4-8). According to Lessac, these concepts should be interlinked and interwoven and should not be used separately.

---

[2] Although Lessac should be applauded for his attempt to provide a completely integrated body/voice approach with the addendum of the Body Wisdom book, this author has reservations about the use of the concept of "c-curve" or "parenthesis curve" when referring to the optimally integrated spine. It is a pity that an approach that can potentially be so powerful, may disempower itself on such a cardinal concept, that has been so well researched that any doubt of the spine as an "elongated s" has be disregarded long ago. Nevertheless, this body of knowledge are beyond the range and scope of this study.

Nevertheless, for the purposes of this introduction, it is useful to describe them in separate categories.

Body "esthetics" is, according to Lessac, "anything that promotes sensitivity and induces awareness of sensation" (1997:4). This concept of "sensory awareness" plays a pivotal role in the Approach. Although Hanson (1997:209) categorises the Lessac Approach as an external approach, this turn of phrase has as specific goal the enhancement of the internal involvement of the performer (Chabora, 1994:114) and voice producer. Crucially, in this concept of "sensory awareness," the trainee will be guided to be aware of and to use the different senses to assist in the monitoring of the verbal, and also the auditory sense.

Inner harmonic sensing (1997:5; 1998:n.p) is a phrase that, yet again, is supposed to induce a holistic approach to voice production (Tabish, 1995:164). "In self-use, inner harmonic sensing provides us with extended and expanded vistas for heightened sensitivity, perception, awareness, response, subtext, synergistic activity, and research" (Ibid, 1998). In essence, Lessac suggests that the trainee is capable of "harmonising" the results of the "sensory awareness" into a fluid and meaningful dynamic and whole, or at least to cluster, or pattern events harmonically and through that "generate newly revealed dynamics, essences and intelligences (Lessac, 1998:n.p).

Organic Instruction is the acknowledgement of the organic flow of movement in the body where the focus will be on the sensorial and the proprioceptive as a self-teaching tool (Tabish, 1995:154). Lessac (1998:n.p) states that the organic instruction in the human body is related to " the search and use of holistic unifiers to reduce complexity". As has been suggested, for Lessac vocal work is part of movement (primarily in terms of the vocal apparatus, but also, significantly, in terms of the movement and expressive potential of the entire body). Lessac uses organic instruction as a self-use and self-monitoring tool, using the concepts of sensory awareness, and the inner harmonic sensing as the guiding mechanism and controller.

---

[3] As will be demonstrated, Voice pertains to the production of vocal sound, and Speech refers to the shaping of the sound into words that carry potential meaning. Lessac intertwines the two concepts.

The Familiar Event is a concept known to all actors. Stanislavski, the great theorist on acting, states the following: "we retain our memories of them (the experienced events*)*, but only outstanding characteristics that impressed us and not their details. Out of these impressions one large, condensed, deeper and broader sensation memory of related experience is formed" (Stanislavki, 1980: 173, parenthesis added).[4] The Familiar Event is thus something familiar that is "thought of" to aid with the task at hand. According to Lessac (1998:n.p) it leads to "body activities.... performed with a sense of creative ease and pleasure due to one's acquired skills, innate talents or natural instinct." In acting, the method is used to conjure up an event experienced by the actor that contains parallel experiences to those demanded by the character for the scene in the play. The actor is thus able to use the "familiar event" as a catalyst for the new, created, moment. In the Lessac Approach this concept is applied to create an ease in voice production. It acts as a guiding tool or (as Lessac states it) the familiar event will provide the voice producer with "a kinesensic image resulting primarily from association with the initial familiar event (that) will become part of body physical memory and then constitutes itself as internal organic motivation for new and future unique events" (Lessac, 1997:7). In other words, by creating (under tutelage) effective vocal sensory awareness moments, that eventually become "automatic," the performer is able to draw on these "familiar events" in times of creating new work.

The cornerstone of the learning programme is the concept of "sensory awareness", and it is in the attempt by Lessac to describe the anatomy and physiology of the voice that is connected to "sensory awareness" that the matter becomes problematic (Hanson: 1997:111,112,169), especially in comparison to the work of scientists like Kent (1997). Verdolini (1998:n.p) argues, though, that although his language is different from the "operationalized, quantitative mode of expression that is increasingly common within the speech-language pathology community", it is a refreshing and imaginative reading that goes beyond the literal. "Sensory awareness" can be seen as a perceptual process, and therefore, in Lessac's defence it should be mentioned that his aim is not to give an accurate anatomical and physiological description but to provide the theatre student with a

---

[4] Indeed, the entire section of pages 172-178 deals with the Sense Memory and the Familiar Event.

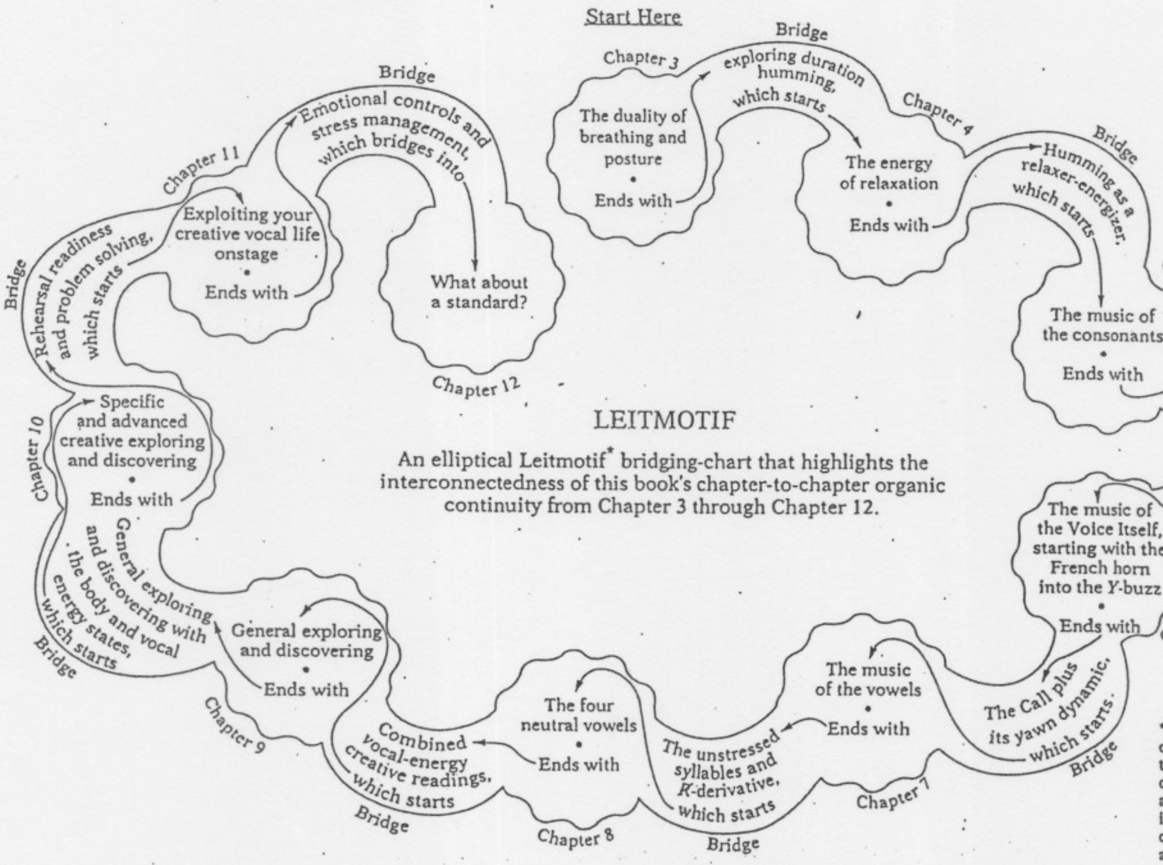
somatosensory understanding of the organic process of voice production in the human body. These somatosensory experiences provided as guidelines to the student contribute to the effectiveness of this approach, as Verdolini (2002) posits that this may lead to heightened motor learning as it implements schemas of recall and recognition (Schmidt, 1991). It is the view of this author, that this should go hand in hand with an introductory overview of the anatomical and physiological functioning on voice production as scientific grounding, as enough evidence in literature exists to clarify the Lessac Approach.[5] Although a biomechanical understanding of voice production does not lead to optimal voicing (Verdolini, 2002), this does reflect on Verdolini's idea that for successful and optimal voice training, three parameters have to be taken into account. These are a) Physiology (biomechanics) b) Learning and c) Compliance (Verdolini, private e-mail about the LMRVT, 2002). Verdolini, however, refers here to the three parameters at play in voice therapy and does not necessarily insist on the "patient" having the knowledge of these three. This researcher argues that in a tertiary institute teaching situation, it is necessary for the students to not only improve their own voice quality and experience good voice production, but also understand these parameters in order to eventually train themselves as well as possible future students. This will support the analytical and empowering approach of any institute of higher learning.

The Lessac Approach is organised in such a way that there is an organic continuity to the work, although Lessac acknowledges that one can actually start working with this approach from any place in the approach (Hanson, 1997:111). In a personal conversation (May 2001) Lessac stressed the fact that the starting point will be determined by the needs of the student/s. The way the work is structured in the latest edition of the voice book, is portrayed here in Figure 1 in Lessac's Leitmotif (1997a:1).

---

[5] The exact make-up of such a course will depend on the demands for this from the student and the administrative structure of the course.

Figure 1: Lessac's Leitmotif.



*Leitmotif: a motive, or theme associated throughout a music-drama or poem with a particular situation, idea or concept; a dominant motive of action.

(Lessac: 1997a:1)

After focussing on the integration of body and breath (1997:17), the Lessac Approach proceeds to work on three different aspects of voice and speech training. These three aspects are: Consonant NRG (a Lessac acronym for "Neurological Regenerative Growth" and an abbreviation for "energy", substituted for the word "action" in the previous editions of the book), Tonal NRG and Structural NRG.[6] In the previous editions the three NRG's are presented in a different order: Structural, Tonal and Consonant (Lessac, 1960, 1967). Lessac, in consultation with his certified teachers, has decided to address, in the 1997 edition, the consonants first, as it is, according to him, a more organic and gentle way of working seeing that the consonant musical energy is primarily humming, which is, in itself a familiar event.

---

[6] It needs to be mentioned that within the complete Lessac Approach, there are three body NRG States that form the cornerstones of the "kinesensic" training (Lessac, 1990; 1997; 1998).

2.1.2.1. The three NRG's

2.1.2.1.1. Consonant NRG

In line with the holistic and "aesthetic" orientation of the Lessac Approach, the approach metaphorises the human body as a musical apparatus (Lessac, 1997:61). Consonants are explored as musical instruments and each consonant is compared to an instrument[7] in the classical orchestra (Ibid, 1997:70). In the 1997 edition no formal/academic descriptions of the consonants are given. In the previous (1960, 1967) editions there were, what Lessac called, "conventional classification"(s) of each consonant. The "playing" of different consonants is guided by either the sensorial awareness of the vibrations of the voiced consonants, or the sensorial awareness of the noise pattern of the voiceless consonants (Hanson, 1997:171). The student is thus using his/her "inner harmonic sensing" (Lessac, 1997:66) to "play" (Ibid, 68) the various techniques of consonant forming such as obstruction, impedance, interruption and friction (Ibid, 67). This approach opposed most traditional articulation exercises, as the emphasis is here on "optimal, gentle esthetic qualitative experiencing" (1998:n.p) of the consonants as "relaxers/energizers" that lead to optimal articulation, but stretches beyond it into a "therapeutic vital configuration" (Ibid). The idea of playing a melody or rhythmic pattern with a specific instrument/ consonant, provides a valuable tool for the performer to relate to consonants as carriers of subtext and meaning, and not just to use consonants as "necessary evils" to shape vowels into meaning.[8]

---

[7] Including percussion instruments.

[8] Examples of musical instruments that, within this Approach, correlate directly to the consonants are: Violin = N; Viola = M; Cello = V; Tympani drumbeat = D; Snare drum = T. For a complete correlation see Lessac, 1997.

Lessac states that consonants are the "anatomical 'spine' of words" (Ibid, 67) and thus stresses that consonants are necessary to provide intelligibility to speech as well as, in combination with vowels, melody and rhythm to language (Hanson, 1997:115). The Consonant NRG of the Lessac Approach is a unique, albeit experimental, way to approach the forming and using of consonants in speech. The Consonant NRG for nasals as well as the [j] can be an excellent tool, as a familiar event (recall schema), to introduce the bone conduction experienced in the Tonal NRG. Verdolini (personal e-mail, 2002) notes that the "Lessac Madsen Resonant Voice Therapy (LMRVT)" follows this approach, as a way to introduce the bone conduction experience to people with vocal pathology as it is easy to experience the sensation of the bone conduction in this manner. This again, reflects the organic processing within the Lessac Approach. This study does, however, not concern itself directly with the Consonant NRG part of the Lessac work.

2.1.2.1.2. Tonal NRG

In the Tonal NRG the focus is on the sensory awareness of the vocal vibrations on the bony parts of the face (Lessac 1997a: 115). These bony parts include "the hard palate at the upper gum ridge, including the teeth, the nasal bone, including the cheekbones and ... the forehead reaching into the cranium" (Lessac, 1997a: 124). Bone conduction is thus within this approach of utmost importance as a self-guiding tool and is used as a familiar event (Lessac, 1997a:123). Although respected scholars note that vocal production happens primarily through the use of an air resonator, and that bone conduction happens as a secondary effect, Lessac chooses to not spend time acknowledging or discussing the air resonator. The only reference to air conduction is when he mentions that the "resonated vocal sounds... reach the outside world through air conduction" (Ibid)[9]. The argument behind this deliberate choice seems to be the fact that the shaping of the air resonator is primarily done by involuntary muscle usage. Focussing on the actual muscle function will tie in with a biomechanical approach to learning, which can actually impede motor learning responses (Verdolini, 2002). Should the voice user experience the vocal vibrations in the bony parts of the face, and thus actually focus on an "external" (or in this case "removed") movement effect, enhanced learning will take place (Wulf & Prinz, 2001). In any event, the

---

[9] The use of the "structural NRG" focuses primarily on the shaping of the oral cavity through muscle memory and as such the air resonator is addressed, albeit, again, not named as such.

assumption can be made that the air resonator is optimally shaped for the specific sound, taking into account both its pitch and vowel quality.

Bone conduction, as a primary sensory awareness tool within the Lessac Approach, is a scientific fact in the sense that the facial bones do vibrate in response to the acoustic waves. Verdolini (private e-mail, 2002) argues that "acoustic waves are alternating regions of dense versus rarefied molecules, i.e. differential propagating pressure regions. Such pressure pushes and releases bony tissue at rapid rates." As hinted above, then, the emphasis on the bony vibrations in voice production acts as an external focus point and the subconscious process of voicing takes place without interference. This is supported by the position of Wulf and Prinz (2001), for example, who argue that optimal motor learning takes place when the attention is placed on the effects of the movement rather on the biomechanics of the movement. This enhances both immediate performance as well as learning.

In this respect, again, it should be noted that the Lessac Approach is not the most accurate in explaining or drawing from the anatomy and physiology of voice production. Hanson (1997:111), as well as Verdolini (1998:n.p) state that it is a perceptual and metaphorical system. Terms are created from Lessac's own vocabulary to describe the sensorial awareness in order to emphasize the holistic approach. These new terms and this "esoteric" approach may be confusing if the student's inclination is to stay within the scientific facts or biomechanical understanding. On the other hand, this approach can be seen as a process where the integration of "contemporary self-use systems (in performance) and the current neuropsychological research" (Chabora, 1994:114) allow the actor/performer to engage kinaesthetically (Tabish, 1995:10) in order to allow optimal body/voice integration and development. Put another way, through a process of sensory and proprioceptive awareness, the actor/performer engages, by means of organic instruction, in self-monitoring and self-development in the dynamic and creative process. This leads to enhanced motor learning.

The Tonal NRG progresses from the French horn consonant (Lessac, 1997:121), which is the [j] sound through the Y-buzz [i], the +Y-buzz, which is the diphthong [eɪ], to the Call, which starts off with the [oʊ] sound preceded with either a [j] or a [l]. Specifically, Chapter 6 in the 1997 edition starts with a reminder about the sensation of the [j] consonant.

The tuning fork experiment, where a tuning fork is set into vibration and the base is then put onto the front upper teeth, is done to stimulate the sense memory of the voice user in order to create a proprioceptive familiar event of bone conduction. According to Lessac this also puts the proprioceptive and tactile sensory awareness as the primary source, seeing that the voice user is "hearing what you feel rather than attempting to feel what you hear" (Lessac, 1997a:122). This is Lessac's attempt to overcome the fact that the voice user never has a clear auditory "picture" of what he/she sounds like to the outside world. Bone conduction will always influence the voice producer's perception of his/her own voice, as this is an effect of optimal voice production.

After creating this familiar event, the approach proceeds to "playing" an elongated [j] sound in the lower pitches of the speaking voice range while focussing on the vibratory sensation (buzz) on the forward upper gum-ridge, hard palate and nasal bone. This exploration will "establish the vibratory foundation for bone-conducted tone in your singing and speaking voice" (Ibid: 123).

One should proceed to link the [j] consonant[10] with the Y-buzz vowel, which is the [i]. This can be done with words like "peace, piece, easy." One then links the [i] vowel with different consonants whilst keeping the sensation of the buzz. Examples of this would be "evening, breeze, free, easy, weepy."

Within this approach, the concept of Tonal NRG is built on the concept of bone conduction as a method of inner perception in order to create a recognizable and re-creatable sensorial event that the voice user can rely on. Neither the vibratory action of the vocal folds, nor the shaping of the air resonator are addressed, but just briefly mentioned, seeing that Lessac claims that these parts of the process of voice production function "involuntarily" and "without conscious sensation" (Ibid, 124). This approach thus implies

---

[10] For example in words like yonder, yes.

that healthy vocal fold functioning and optimal shaping of the air resonator, for the specific pitch and vowel at hand will happen "effortlessly" when the "buzz" is felt optimally.

To assist in the feeling of the buzz, and by implication adjusting the resonator length, a forward facial posture (Ibid, 125) is advised. This researcher prefers the use of the term "forward facial orientation", as the danger exists with the use of the term "posture" to create a "held position" (Thurman & Welch, 1997:143), whereas Lessac stresses the moveability of the musculature and thus refers to an e-centric muscle usage. To create sensory awareness and a familiar event for this forward facial orientation the "Shhh" sound [ʃ:], as if telling a child to be quiet, is used. This is then combined with the French horn and Y-buzz sounds. It should be noted that the jaw should always be relaxed by allowing a flexible space between the teeth, while still feeling the sensation. This space between the teeth should be allowed to function as a "variable space cushion" (Ibid, 126).

The position of the tongue (Ibid, 126) needs to be observed, but not forced. The tip of the tongue will be gently making contact with the front lower teeth, whilst the sides of the dorsum of the tongue touch the upper molars. Lessac (e-mail conversation, 2000) insists that the tongue "should never be withdrawn towards the back of the throat. The farthest back position for the tongue is while sustaining the American sound [ɽ], which doesn't go very far back, ..." In a personal conversation (May 2001) Lessac noted that he will address the "placement" of the tongue only if it is a problem and then only to guide the voice user to an orientation of the tongue towards this position. Nothing must ever be fixed of forced. "If the facial posture is being used in a natural way, and the anti-gravitational flow of the bone conducted tone is alive and vital, the tongue will always be where it should be, out of harm's way" (e-mail conversation, 2000).

With all the above as guidance, as a means to holistically explore the sensory awareness of the sound, the Y-buzz can be developed as a voice building tool. In developing the Y-buzz the Lessac Approach focuses on:

- working in the lower third of the vocal range;

- keeping all organs of voice production free and moveable at all times, without any pressure or unnecessary constrictions, seeing that free shape and space adjustments need to be possible at all times;
- combining the forward facial orientation with the sensory awareness of the sound vibrating into the frontal bony parts of the face;
- body integration;
- overcoming and side-stepping nasality by focussing on, and feeding the Y-buzz sensation which will guide the voice user to an involuntary lifting of the soft palate;
- optimal energy, without tension but with intent;
- integration between breath and sound in such a manner that the voice user should eventually not be aware of breath as a separate entity and therefore create optimal vocal fold action and output through the use of sensory awareness;
- changing and refining the sound while sustaining it so that the sensory awareness and "kinesensic training" will always be there as a guidance tool (Ibid, 127).

Lessac emphasises the concept of the Y-buzz as a synergistic event where the forward facial orientation in combination with the vibratory sensation will act as a jaw and muscle relaxant (Ibid, 128). A taut jaw will have a direct, negative effect on the sound quality produced, and on the ability to create the buzz.

The range of the Y-buzz can be adjusted by using progressively lower pitches. Throughout the exercise the vibratory sensation is required to act as a sensorial guidance and as a familiar event. The reduction of the opening between the lips is observed when the vibratory feeling is kept when moving to the lower pitches. An increased opening between the lips is observed when the pitch is higher and the frequency thus increased. Again this underlines the holistic approach of this system. The change in the resonator is led by the sensory awareness (Ibid, 129), as well as the interrelationship between the sensory awareness, pitch, and the modulating shape of the resonator, leading to the creation of a "familiar event," or series of "familiar events."

The approach proceeds with the use of words and sentences that contain Y-buzz-like sounds.[11] The aim is to link the vowel qualities with consonants without losing the "buzz" as a guideline for optimal, comfortable and healthy voice production.

The next step in the development of the Tonal NRG will be the +Y-buzz (1997a:131-133). As mentioned before, this is a diphthong ([eɪ]). During this exploration the stress, unlike in the natural English speech pattern, is placed on the second part of the diphthong, which is, of course, that part which has become the original familiar event created by the vibratory sensation of the Y-buzz – again linking the new exploration with the familiar. All the characteristics of the Y-buzz remain, including now a slight pulsating rhythm where the voice user proceeds from the "yeee-yeee" [ji:-ji:] into a "yey-yey" [jeɪ-jeɪ]. There is thus a slight change in the shaping of the frontal part of the air resonator, as there is a difference in the vowel that is produced (from [i] to [eɪ]). The voice user thus starts from and returns to the familiar Y-buzz as a familiar event, passing through the smallest possible kinesensic change that delivers a different vowel. In this manner the sensation of the + Y-buzz is becoming, through the kinesensic training, a familiar event in itself. After being able to recognise and control this sensation the voice user will proceed to words and sentences containing the +Y-buzz sound.[12] Initially it may be necessary to elongate the second part of the diphthong (i.e. the Y-buzz), in the explorations, but as soon as possible, the voice user must proceed to using the natural rhythm of speech whilst maintaining the "buzz sensation" (Ibid, 131-133).

The logical organic development of this leads, according to Lessac, into "Carry over opportunities" (Ibid, 133) where this, now familiar, sensation -- which is supposed to provide the voice user with a "rich, vibrant, gently ringing" (Ibid, 134) tone - will be carried over into different vowels.[13] The "Y-buzz tonal current" (and the "current" might be described as an "ongoing action") is an exploration where the voice user will start producing a Y-buzz, then proceed into a word containing the Y-buzz vowel, then into a

---

[11] Words like "thieves, kingly, tincture, conceived meagre" are used. "Please don't leave me; I need you" and "We need steam heat immediately" are two examples of sentences being used (Lessac, 1997a:130).

[12] Words like "play, nation, shapely, ungainly" are used. " Every baby needs to play each day" and "They came very late, they also stayed very late." are examples of sentences used (Lessac, 1997a: 132).

[13] The potential "subjectivity" of these terms is surprisingly constant across many researchers and practitioners attempting to describe the tone. (Nair, 1999:7)

word containing another vowel, then again uttering a word using a Y-buzz vowel.[14] In this way, the voice user can constantly monitor the sensation of the vowels using the familiar event of the sensory awareness. The Lessac Approach starts this "Carry over" with vowels which are traditionally seen as high front vowels, which have a small lip opening and a high, forward tongue placement (Nair, 1999:95). Gradually vowels are introduced where the lip opening increases, and the point of tongue constriction moves back and where the tongue itself is lower in the mouth (Ibid, 96). These vowels appear in the following order in the system: [u] as in cool; [oʊ] as in go; [ɔ:] as in called; [ɒ] as in odd; [ɑ:] as in calm; [æ] as in the American pronunciation of asked or cat. Where there is a diphthong combination where the vowel is followed by an [i] or an [ʊ] sound, it also becomes integrated in this exploration (Lessac, 1997:134). Once again, the voice user is using the principles of sensory awareness, familiar events and kinesensic training in order to shape the resonator to carry the concept of voice building over into vowel formation. This work is then carried over into readings where the conversational speech mode is explored (Ibid, 136).

The Y-buzz and + Y-buzz training focuses the voice building around the lower third of the speaking voice range. For development in the higher range of the speaking voice (as well as the singing voice) the Lessac Approach uses the Call (Ibid, 136-155). As is the case with the entire approach, the Y-buzz (as established familiar event) is used as the basis for this work. Following the Y-buzz and then the + Y-buzz the voice user is guided to focus on the forward facial orientation and the sensation of the buzz through the frontal bony structures of the face. From here on the voice user is guided into the Call where a "fuller forward facial posture (sic), ... a definite yawn feel, and, therefore a wider space between the teeth" (Ibid, 137) are used. In effect what Lessac suggests is a change in the shape of the resonator. The Call must be noted to emerge out of the Y-buzz, but it is different from the Y-buzz as the Y-buzz has a "reduced facial form and shape and minimal space between the teeth" (Ibid). (Again, this may be seen to be related to the shape of the air resonator.) Lessac stresses the concept of moveability and comfort of sensation during the Call action. In this specific exploration it is also the sensation of the vibration on the hard palate and in the frontal bony structures of the face that is of primary importance. The

---

[14] Example: "Keep cool Mimi" (Lessac, 1997:134)

actual placement of the vowel is influenced and dependent on the pitch of the Call. "(T)he subtler qualities of vowel articulation are secondary" (Ibid). Vowel modification is thus allowed to happen (Ibid, 147) in order to experience the optimal vibratory sensation on each pitch.

The Call starts with the [oʊ] diphthong proceeding from the [h] and [l] sounds. The first "word" that is thus used is "h'LLO".

In addressing the "control of the Call" Lessac underlines the idea that the Tonal NRG is the control of a vibratory current of sound in a state of constant movement, radiation, and transmission propagating in the hard palate and teeth, nasal bone, cheekbones, sinuses, forehead and cranium" (Lessac, 1997a:139). The concentrated tone of the Call will be used in situations where a "big voice" is needed, for instance when addressing a large group of people. This concentrated tone will not be used in everyday conversation but a diluted version of it will rather be used.[15] If the Call is experienced as an inner kinesensic experience, the adaptation of the usage of the well-focused voice will happen organically. In such a case it will be possible to move easily from one pitch to another.

It is possible to use short or sustained Calls. It should be encouraged to use both as both these Calls add to building the voice. Lessac states that whilst doing the Call the voice user should:

- be in a large space so that distance can be experienced;
- be able to sense the vibration on the hard palate;
- experience a pleasant yawn-like sensation that should be "pleasant, flexible, energy-giving" (Ibid, 141);
- experience no heaviness or throatiness;
- allow no force or breathiness into the Call action.

---

[15] Dilute tones or sounds with a "diffused focus" are, according to Lessac (1997a:152) used about eighty to ninety percent of the time during normal conversational speech and about twenty percent less in performance. A concentrated tone (which will project optimally) will become diluted when the space in the oral cavity is altered to be too big for the specific pitch of the tone. Lessac (Ibid, 182) defines the diluted tone as referring to

Lessac reminds the reader that when the Call is reinforced by focusing on the "forward facial vocal yawn" sensation, it proves to have "many virtues" (Ibid, 141). These include:

- a therapeutic effect. Lessac claims that the body will be stimulated to optimal breathing;
- flexibility and a resilience of facial muscles will take place;
- through the awareness of bone conduction of sound while yawning, an awareness of the hard palate will be created. This will lead to an organic raising of the soft palate, that will lead to the domino effect of creating space in the pharynx which will translate into the shaping of the air resonator;
- every Call should thus have a relaxing as well as an energizing effect on the body as it should include a full body and voice yawning which act as a synergistic "relaxer/energizer" (Ibid,142).

In relation to the pitch of the Call, the sensation of the vibration on the hard palate will differ in size. This phenomenon is known, in Lessac terminology, as the size of the "focus pocket." The pocket will be smaller, and closer to the gum-ridge, for lower pitches and bigger, filling the whole hard palate, for higher pitches. The size of this pocket should correlate with the size of the opening of the mouth. This area of sensation on the hard palate is directly connected with the space used, three dimensionally, in the oral cavity. The bigger the "pocket," the more space in the mouth is needed. The result (as with the Y-buzz), in the release of the jaw, is to allow a bigger lip opening. This is thus the shaping of the air resonator whilst focussing on the sensation of the bone conduction. It should be remembered that the actual change in size is minute. While producing the Call, the voice user should actively associate him/herself with the sensation, as it is still the kinesensic and proprioceptive awareness that allows the involuntary actions (through organic instructions) of the generator, vibrator and resonator to take place.

The approach proceeds from short spontaneous Calls into the sustaining of Calls. After producing the sustained h'LLO Call successfully, according to the voice user's own sensorial judgements, other vowels and diphthongs are used for Calling. Words that are

---

"a non-concentrated vocal focus with the variable use of facial posture. Anytime the Call or the Y-buzz spreads or lightens in texture, it becomes dilute tone."

used in the system are "aWAY, unTIL, unEARTH" (Ibid, 146). Body integration is always of importance -- movements that will act as muscle relaxants should be consciously employed as aids to relieve stress generally, but specifically in the voice-production areas.

Lessac introduces a "Yodel Call" where the voice user does a sustained h'LLO Call but while sustaining pitch the tongue moves toward the position for the Y-buzz and back to the position for the [oʊ]. Should this be done more rapidly a yodel effect will be obtained (1997a:146).[16]

Other variations of the Call are:

- the siren/sliding Call (Ibid,147) that happens as a glide up and down the voice user's Call range. Gradually this range will be extended, being guided by the sensory awareness. Depending on the pitch used for these sliding Calls, the voice user will go through different vowels as the mouth openings change;
- a Call on the American consonant "R" [ɾ] into the [oʊ] (Ibid: 148). In this case the tongue is isolated for free movement while the facial form, shape and size will stay stable in the optimal position for the [oʊ] diphthong and specific pitch;
- Call phrases where in one sentence different vowels occur. In these phrases the length of the resonator will be the same for each pitch but the shape of the lip opening as well the tongue position will change according to the vowel. [17]

There are several variations of the basic Call sentences. In these variations the focus is on combining the sensorial, kinesensic and synergistic effect of the quality of tone with the freedom of movement and the artistry of the performer (Ibid, 151-159).[18]

It is very important to note that, according to Lessac, the Call is "the common denominator in speech and singing" (Ibid, 158). Titze (1994b:248) supports this when he refers to the Call as highly resonant. It has to be clarified here that this is a specific reference to the

---

[16] Please note that Lessac is referring to the rapid changes of vowels in the diphthong and not to the abrupt register shift that is understood by the "yodel effect" (where the voice user rapidly change between the chest and falsetto registers) in voice science.

[17] Examples are: "You may go"; "Ahoy there"; "The dirty old grey boats will float again Thursday!" (Ibid: 150/1).

[18] The reader is referred to "Improvisation in communication through the Tonal NRG" (1997:151) and to "Role Calls" (Ibid: 153).

German style of classical singing rather than the Italian. This concept becomes vital in the next chapter, as the study engages with the singer's and actor's formant.

The essence of the direct influence of the Tonal NRG of the Lessac Approach on the resonator, and indirectly on the vibrator, should be separated from the methodology where the holistic approach and behaviour modification are essential. The guidelines for shaping of the vocal tract as reflected in the Tonal NRG are the following:

In general:

- sensory awareness of the vibratory sensation in the bony frontal parts of the face with specific focus on the upper gum-ridge and hard palate (1997a:123);
- the forward orientation of the lips (1997a:125) – in a more traditional vocabulary this may read as protrusion, although the difference here may be the use of the facial muscles in an e-centric manner;

Specifically:

For the Y-buzz

- the forward high tongue position (1997a:126);
- the specific vowel ([i]);

For the + Y-buzz

- the return to the Y-buzz vowel (1997a:131);

For the Call

- the agility of the tongue (1997a:146) in order to articulate the different vowels;
- the space between the teeth that leads to a released jaw action (1997a:144);
- the yawn sensation (1997a:137) that should be leading with the sensation of the "buzz" at all time.

2.1.2.1.3. Structural NRG

"Structural NRG refers to the mold, shape and size of the human voice and speech instrument ..." (Lessac, 1997a: 160). Lessac uses a seemingly strange and unscientific description for the shaping of the front part of the air resonator as he here refers to this oral cavity as the "sound box" (Ibid).[19] The structural NRG focuses on the muscle actions for shaping the oral cavity for clear speech through organic instruction and muscle memory. This way the articulators responsible for the shaping of the vowels are used in a synergistic manner. True to the holistic approach, Lessac focuses on the "kinaesthetic action" as well as the "kinesensic application"[20] (Ibid, 160/161) of the movement patterns that have to take place in the oral cavity, cheeks, jaw, and lips.

A controversial aid, which is sometimes used in the Lessac Approach, is the use of a cork between the upper and lower side teeth. The possible logic behind this is that it would provide the student with a familiar event (recognition and recall schema) of the inverted megaphone. Once the familiar event is established the cork should be discarded. The cork is lightly held in position with the teeth. The lip opening for the [ɑ:] vowel (as in "father") is produced. The mouth opening must then gradually be reduced while the space between the teeth stays the same. The yawn sensation should again be employed to facilitate the space in the mouth. The facial structure that is induced by this experiment is known, within the approach, as the "reverse megaphone" which is synonymous with the forward facial orientation. Yet again, it can be noted here, that the whole approach is intended in a holistic manner and can thus not be divided into separate sections. Although there may be a focus on one specific aspect of the work, the other aspects will always be implied. The total will thus always be greater than the sum of the parts.

---

[19] As has been pointed out above, it should be acknowledged that this is not strictly scientific use. It may be to the advantage of the performer that metaphors are used in teaching, but they may create confusion, as they are tangential to what is used in voice science. However, to discard the approach because of its deviation form the traditional vocabulary used by scientist will be short sighted. Using this approach and fusing it with the traditional terminology of the field of science of voice seems to be optimal and the responsibility of the voice teacher.

[20] Lessac has coined the term "kinesensics" to explain the neuro-physical and neuro-muscular sensorial process that happens within the human body (Lessac: 2000:191). The concept "kinesensic application" thus demonstrates the "perception of motion and sensation ...through the senses rather than the intellect" (Ibid).

Although there may be merit in the use of the cork – especially in a caring environment and with the organic approach as foundation, the author opposes such a practice and would rather guide voice users towards the next exploration in the Lessac Structural NRG work where two fingers are lightly placed vertically on the face just next to one corner of the mouth. The awareness of the space is created, but it is not fixed. There is, in the author's view, no exact distance which teeth should be apart. It depends entirely on the anatomical build of each voice user and can therefore only be used as a guide to make the voice user aware of optimal space in the oral cavity. The "reverse megaphone" must always be soft and pliable, never forced. If this exploration is taken too far it may result into tension in the tongue, jaw and around the larynx. Lessac underlines that there must be no force or pushing, when he states that the voice user must focus on the behavioural image of a forward diagonal elasticity which will cause a freeing and softening effect in the whole area (Ibid, 164). In this way the voice user will be able to have a good healthy voice quality while producing "natural, proper vowels" (Ibid, 164).

The vowels that are used as "Structural Vowels" (Ibid, 164-167) correlate with the size and shape of the lip openings. The vowels are given numbers and start from the smallest lip opening to the largest. Should the sound be a diphthong then two numbers are allocated to the sound. It must be remembered that this is in this sense a truly American system and that, although the concepts are generic and can thus be universal to all voice and speech development, adaptations may have to be made for application in other languages.[21] The choice of vowels here is based on American English. The vowels/diphthongs and their numbers are:

#1 [u:] ooze, crude
#21 [oʊ] ode, boat
#3 [ɔ:] law, all
#4 [ɒ] odd, yonder
#5 [ɑ] father
#51 [ɑʊ] ounce, down
#6 [æ] add (and in American: class, ask)

---

[21] Again the responsibility here will lie with the voice teacher. The motivation behind providing the different lip openings with a number is again to provide a familiar event through muscle memory.

#3y [ɔi] oil, toys
#6y [ɑ:i] isle, kind
#3r [ɚ] early, bird

In line with the kinesensic training approach, it should be stressed here that the focus is on proprioceptive and sensorial feedback. The students should be guided towards an integrated muscle memory. It is thus again, a holistic orientation towards pronunciation, rather than a use of the auditory aspect alone to make judgements about vowel qualities. The emphasis for the voice user is on forming and feeling the three-dimensional shape of each vowel.

There are several different ways in which the Structural NRG can be explored. An exploration that enhances and develops "flexible lip movement while maintaining the full inverted-megaphone" (Ibid, 168) is "woo... woe...war... wah... wow." This combines the optimal shaping of the oral cavity with the possibility of healthy voice usage, as it employs the initial constriction due to the inverted megaphone and forward orientation at the lips that relates to the lip position of the Y-buzz. Important to note is that this Structural NRG, through the sensing of the space and the shaping of the oral cavity, addresses the release of the jaw without emphasising it.[22] This is thus a gentle approach to allow the jaw through muscle engagement, to release downward. The jaw is not pushed or forced. Proceeding explorations include word lists and sentences where different structural vowels are combined. Eventually these structural vowels are linked through the "carry-over exploration" and reading selections with the Tonal NRG and the Consonant NRG.

After guiding the voice user through the three NRG's, the system proceeds to investigate the use of neutral vowels that influence the rhythm of speech directly. Although this is an important concept in the use of language as it feeds into the natural rhythm of language, this section will not be further discussed, because it is the section, within this approach, which is the most American (seeing that it deals with neutral pronunciation of vowels within American English). This will differ from language to language, even from country to country

---

[22] Employing the concept of organic instructions.

where both countries use English as a spoken language. This borders on the field of accents and dialects, which is not the focus of this study.[23]

Lessac's voice text proceeds further by focussing on the exploration of creative vocal life and is thus the application of the "so-called technical phase"[24] of the approach in the creative – moving from building the craft to exploring the art. It should be noted that the phase of "building the craft" should always happen as a holistic intrinsic approach and not as a series of technical exercises. The voice teacher should always take into account the psycho-physical nature of this work and never attempt to force the voice and speech into something purely technical.

2.1.3. The Tonal NRG and existing literature in Voice Science

Seeing that the primary focus of this study is the Tonal NRG as a possible way of voice-building, existing scholarly literature about the Lessac work will be reviewed, especially where it specifically provides scientific evidence towards the viability of the Tonal NRG as a voice-building system. Existing literature seems to comment on the Lessac Approach in three different areas: that of the approach as a teaching tool, an analysis of the physiological results of the approach, and considerations of the acoustical outputs of the approach.

2.1.3.1. Critique of the Lessac Approach as a teaching tool

Hanson (1997: 111) warns that the Lessac system can be confusing, because Lessac does not use scientific explanations for his explorations. She reacts negatively to the Lessac Approach when she states that the Lessac Approach is the most complex of all voice systems known to her. She claims that a highly trained teacher's input is necessary in order for the voice user to know that his/her voice production is optimal. She furthermore observes that the approach is perceptually and metaphorically based (Ibid) without a direct correlation and control between the student's subjective experience and the intended

---

[23] It should be noted that "neutral vowels" as a generic concept will exist in any language. Again the Lessac Approach provides the framework that can be developed and applied in any language.

[24] Although this may seem to be "technical" is should be noted that the approach is always holistic and engages the sensorial, if not the psychological, in each voice user during every experiment or exploration.

experience. This critique may be valid to a point but this researcher disagrees with Hanson and posits that this approach is first and foremost embedded in the organic and sensorial awareness of the voice user,[25] which is critical and central to motor learning (see Verdolini, 2002; Wulf & Prinz, 2001).

Tabish (1995), reacts positively to the Lessac Approach as a teaching tool in actor training, arguing that the Lessac Approach is a "kinaesthetically based mind-body system" (1995:138) which not only develops the performer's voice but also encompasses a "sense of well-being and a holistic sense of self" (Ibid). Chabora (1994:114) hails Lessac as an innovator in the application of neuropsychological research to actor training. She admits that Lessac himself acknowledges that the pedagogy of the Lessac Approach is still "pre-scientific" (1994:119), but claims that scientific support for this teaching approach is manifested in areas like "bio-feedback, hemispheric research and systems research" (1994:120). Chabora suggests that the Lessac Approach reflects the benefit of neuroscience and neuro-psychology for actor training.

Raphael (1997:209) underlines the idea that the Lessac Approach uses behaviour modification as a teaching tool and that it is a holistic approach to voice training as it induces an "ongoing state of habitual awareness." Raphael admits, though, that not all voice educators react positively to the Lessac Approach because of the "untraditional terminology," and until the voice user has heightened and refined his/her own sensory awareness, the Lessac Approach may seem to advocate forced or uncomfortable speech patterns. Should the Lessac Approach, according to Raphael, not be taught with great care and integrity, it may lead the voice user to produce a pretentious and self-conscious way of speaking (1997:210). In this sense Hanson (1997:111) is echoed. However, this researcher would suggest that it is true for any other voice system, and the responsibility lies in the hands of the teacher to make sure that s/he is clear in his/her instructions and explanations. It seems that the shortfall with this approach lies in the fact that as a holistic approach, no written material can actually justify and explain the body of the work in such a way that the teacher will not need classes from somebody who is thoroughly trained in this approach. This argument needs, however, to be counteracted by the idea that it is the

---

[25] The sensation of vibration on the hard palate can serve as an example here.

current Western ideology that does not allow the individual to be so "in touch with the inner-self" that the recognition of the optimal neuro-muscular or neuro-physical usage of the body will take place immediately by the individual.[26]

A questionnaire[27] answered by Certified Lessac Teachers, people on the verge of being certified as teachers and a speech pathologist doing research on the Lessac Approach, provided further information about the Lessac Approach as a teaching tool. This adds a valuable subjective scholarly evaluation of the Lessac Approach. These questions were responded to from their subjective viewpoints as experts within the Approach. For the purpose of this study only the questions directly relevant to this study will be reflected on.[28]

- Of the respondents, 67% use the Lessac Approach in combination with other voice systems, while 33% use only the Lessac Approach.
- All the respondents report the Lessac Approach to be an effective voice-building tool.
- All respondents agree that the Tonal NRG enhances projection of the male, as well as the female voice.
- Although only 67% of the respondents teach in a multi-lingual classroom situation, they all agree that the Approach (including the principle of the Tonal NRG) is, according to their perceptions, language bound.
- 83% indicate that they assume that the Tonal NRG affects the aerodynamic-myoelastic balance in a positive way. 8% comment that this should be looked at experimentally.
- 75% posit that the use of the Tonal NRG probably leads to a free relatively low laryngeal position that is not forced, during voice production.
- Although 17% indicate that they did not understand the question about the effect of the Tonal NRG on voice onsets, 83% agree that the Tonal NRG would lead to a "smooth, healthy, simultaneous" onset.

---

[26] This ideological framing necessitates body integration systems like Alexander, Feldenkrais etc.

[27] Set-up and distributed by this researcher and filled in at the Lessac Swarthmore convention, Swarthmore, Pennsylvania, USA, 1998. Participants had been involved in the Lessac Approach for 7-28 years. All 12 were teaching the Lessac Approach at the time of the questionnaire, either in a tertiary setting, private studios and/or workshops. One person taught the Approach in a clinical situation.

[28] The complete questionnaire as well as the interpretation thereof is available from this researcher.

- In reaction to a question about the advantage/benefit of the Tonal NRG as a teaching tool, 58% mention the sensory awareness; 33% the easy access to resonance; 33% indicate that it leads to the use of a balanced voice producing mechanism; 17% posit minimum vocal fold impact as the greatest benefit, whilst 17% indicate an ease in voice production.
- Although only 8% are working in a clinical situation, 92% indicate that they thought the Lessac Approach could be used in voice therapy.

2.1.3.2. Critique on the physiological input of the Lessac Approach

As noted earlier, the Lessac Approach is used in voice therapy and one of its permutations in that context is, within that discipline, known as Lessac–Madsen[29] Resonant Voice Therapy" (Verdolini, 1998:34). The "Resonant Voice Therapy" system, developed by Verdolini (Ibid), is primarily built on the Tonal NRG of the Lessac Approach (Berry,Verdolini, Chan & Titze, 2001; Peterson, Verdolini-Marston, Barkmeier & Hoffman, 1994; Verdolini, Druker, Palmer & Samawi, 1998:147; Verdolini-Marston, Burke, Lessac, Glaze & Caldwell, 1995:77). Verdolini and her fellow researchers did groundbreaking work with in-depth physiological studies of the "resonant voice."[30] Verdolini et al. (1994:147) found an intermediate level of laryngeal adduction for the production of resonant voice, in comparison to pressed voice that utilises hyper-adduction, and breathy voice that utilises hypo-adduction (see also Peterson et al., 1994). In a comparative study done by Peterson, Verdolini et al. (1994) they indicated that resonant voice has:

- (again), an intermediate adduction rating, similar to the normal voice (1994:341),
- relatively low contact stress (Verdolini et al., 1998; Berry et al., 2001),
- a closed quotient proportion (of EGG) that is similar to that of the normal voice (Ibid: 340).

---

[29] Madsen was one of Verdolini's teachers and his work, together with that of Lessac had a great influence on shaping her teachings.

[30] It is, for the sake of this project, important to note that Verdolini equates resonant voice with the fundamental principles of the "Tonal NRG" (Verdolini, Druker et al., 1994:147; Peterson, Verdolini et al., 1994:335; Verdolini-Marston, Burke et al, 1995:75; Verdolini, Druker et al., 1998:316).

Verdolini-Marston, et al. (1995) did a comparative study between two voice therapy systems: confidential voice[31] and resonant voice therapy. They argued that the resonant voice is produced with a relatively "complete anteroposterior vocal fold closure during phonation" which correlates with the use of a well-projected voice (Ibid, 75). In a more recent study, Verdolini et al. (1998) further investigated the laryngeal adduction in resonant voice. They underline their previous findings that in resonant voice usage (and therefore the Tonal NRG of the Lessac Approach) a barely abducted and adducted laryngeal configuration is taking place (Verdolini et al., 1998:325) again comparing favourably with normal voice. Further research that has been done in this field (for example Berry et al., 2001) supports the notion that voice production where the laryngeal configuration maximizes voice output whilst keeping vocal fold impact stress small and minimizing subglottal pressure required for phonation. It is clear that in most aspects of voice production, resonant voice is similar to normal voice. The question then arises why the use of the resonant voice, especially for the performer, is so important? Although the laryngeal configuration is similar for resonant voice and normal voice, the difference, for one, lie within the vocal tract changes. Verdolini, Druker et al. provide further information, although within the discipline of speech therapy:

> Theoretically, Ev-max[32] involves an optimal tradeoff between the voice output intensity (maximized) and intraglottal impact stress (minimized), which, computationally, should be produced with barely abducted vocal folds. The barely adducted laryngeal configuration used for resonant voice appears to include the range needed for maximum vocal economy (Verdolini, Druker et al., 1998:325).

And:

> "The barely ab/adducted laryngeal configuration used for resonant voice may have further benefits, beyond the potential for producing Ev-max. Within the range of glottic configurations, the subglottic pressures required for vocal fold oscilation are smaller than for any other configuration, assuming constant F0 and constant tissue viscosity conditions" (Ibid, 326).

And:

> "...it is likely that resonant voice is often produced with subtle supraglottic adjustments .... to enhance the glottal spectrum and thus increase oral vibratory sensations. Such adjustments are in fact frequently part of resonant voice training, and can boost oral output by a maximum of about 6 dB with no change in subglottic pressure or glottal resistance. Thus, resonant voice

---

[31] "...an easy, quiet, breathy voice, as if speaking confidentially to someone at close range" (Verdolini, 1998:38).

[32] Ev-max: vocal economy.

> training might accomplish output advantages not present in normal voice without additional glottal loading" (Ibid, 326).

This research is an ongoing process and although no specific empirical data has yet been collected on the supraglottic adjustments in resonant voice in direct relation to the Tonal NRG, the Berry et al. (2001) article definitely indicates that available data reflect positively on the use of resonant voice and that further studies will investigate this phenomenon further. From this research done, so far, by Verdolini and partners, the essence of the physiological functioning of the resonant voice therapy, which equals the Tonal NRG of the Lessac Approach, can be extracted as *a physiologically efficient approach to voice production, as it requires relatively small impact stress between vocal folds (thus putatively protecting them from injury), whilst it creates a maximum output with sufficient harmonics for projection capabilities.*

2.1.3.3. Critique on the acoustical outputs of the Lessac Approach

Acker (1987) did research on vocal tract adjustments for the projected voice and described the tonal quality of the Lessac Approach as "ring phonation" (1987:78). This term, according to Acker, describes the production of vowels produced at considerable volume, with the forward facial orientation and a lowered released jaw, where the pitch of the sound is related to the lip opening. Acker's study indicated that with the same amount of subjective input (effort), the "ring phonation" was perceptually louder and proved acoustically to have an increased amplitude if measured against what she perceived as constricted voice use. The "ring phonation" also showed an amplitude increase in the overtones clustered around 1.3 kHz, 2kHz and ~3.6kHz[33] (1987:79).

Raphael and Scherer (1987) used the Call[34] of the Tonal NRG as the basis for a research project. Their reasoning for using the Call was that although it may not be the best or only way to train actors' voices, it has a "rather specific, replicable quality (that) can be taught, and (that) has had obvious usefulness for projection on stage" (1987:83). They also suspected that the Call combines laryngeal function and "voluntary shaping of the vocal tract" (ibid) in a way that will positively contribute to "enhanced vocal resonance and

---

[33] Also read as 1300 Hz, 2000 Hz and more or less 3600Hz.

[34] The Call forms the third building block in Lessac's Tonal NRG.

projection." In the female voice samples of the Call, they found that F1 was "tuned more sharply in Call mode than in speech mode" (Ibid, 84/5) as well as a "midfrequency enhancement ... at ~2,000 – 2,500 Hz" and a "drop in energy right beyond F3 (at ~3,000Hz)" (Ibid, 85). This information is of particular importance to this specific study, as current findings will be related to these outcomes of Raphael and Scherer. In the male voice samples their findings compared favourably with their findings in the female voice samples.

They speculate that the optimal space between the teeth, the tip of the tongue "resting gently against the back of the bottom front teeth" (Ibid, 86), forward orientation of the facial muscles and the adjusted lip-opening as well as the idea of an open and relaxed throat, advocated in the Lessac Approach, relate to the acoustical findings – especially to the frequency lowering and amplitude enhancement of the first formant centre frequency (Ibid, 86) as the vocal constriction moves anteriorly. They postulate further that the lowering of the jaw to adjust to different higher pitches could assist in tuning the first formant "in the presence of lip protrusion or larynx lowering, since jaw lowering tends to raise the centre frequency of the first formant" (Ibid). Another aspect that needs further investigation is, according to them, the vocal tract wall that may have an increased stiffness, due to the "yawn sensation" (Lessac, 1997a:142) that may lead to enhancement of the amplitude of the formants.

Titze (1994b:248) reacts positively to the Tonal NRG of the Lessac Approach, when he states that the usage of the Call concept leads to a speaking voice with good resonance. In a later article (1999:41) he again mentions the attribution of Lessac's Tonal NRG to vocal pedagogy when he ponders the idea that the usage of vowels like the [i] and [u] elicit a "lighter, yet more resonant, vocal mechanism." Laukkanen (1995:77) relates Lessac's concepts of the forward facial orientation and the reverse megaphone to phonation into a tube, where the reasoning is that better voice quality will emerge when phonating with a narrowed lip opening or with "only a narrow pathway between the tip of the tongue and the alveolar ridge." This phenomenon is due to vocal tract impedance where the supraglottal pressure is increased, which happens because of the anterior constriction. This has a domino effect on the "sub- and supraglottal mechanoreceptor-based transglottal pressure sensing control system" (Laukkanen, 1995:77), which in turn influences laryngeal functioning with decreased adduction. Laukkanen (Ibid) speculates that this may lead to an optimal laryngeal setting for a maximum output, which ties in with Verdolini's thoughts on this (Verdolini et al., 1998:326) and possibly explains the vibratory sensation on the hard palate.

Laukkanen (personal communication, 1999) explains that this "buzzing sensation" in the oral cavity happens because of the phenomena of the standing wave. In the resonating pipe where the one end is closed and the other open, it happens that nodes or particle vibration maxima of the standing wave of each formant manifests at the end of the pipe where the pipe is open, thus in the case of the vocal tract, nodes of each standing wave occur at the lip opening (Kent & Read, 1992:26). This may create a vibratory sensation in the frontal bony parts of the face, especially the gum ridge and hard palate. Titze supports this when he postulates that these sensations may be related to the "localization of pressure maxima" (1994b:167) of the standing wave in the vocal tract. He mentions that the [i] vowel will have high pressures in the palatal region. Furthermore he argues that concepts like the sensation of the tone on the hard palate may also be related to the acoustic pressure maxima (Ibid), otherwise put as the particle vibration maxima.

Titze (1994b:167) underlines the fact that singers and speakers sense certain vowels in specific locations in the vocal tract, due to acoustic pressures in the front of the mouth that provide "ample vibrational feedback" (1994b:248) and that it is referred to as the focus of a vowel when placed correctly. McKinney (1982:31) as well as Vennard (1967:95) warned that the awareness that voice users may have of the vibrations in the bone, which can be called bone conduction, will not have an effect on the "radiating" quality of the voice. However, it is generally accepted amongst professional voice users that these "buzzing" sensations can be use as a guiding tool.

According to existing literature on voice science the forward facial posture/orientation will have a lowering effect on all the formant frequencies due to the lengthening of the vocal tract as resonating tube (Kent & Read, 1992:16,24; Titze, 1994b:165; Sundberg, 1987:22; see also Raphael & Scherer, 1987). The decreasing or increasing of the formant frequencies in a non-uniform pipe (Kent & Read, 1992:17), such as the vocal tract in different vowel configurations, are measured against the formant frequencies of a uniform resonating pipe. The average male vocal tract (17,5 cm) has, when producing a mid-central or neutral vowel, the formant frequencies of 500, 1500, 2500, 3500 and 4500 (1992:14,15). In this specific study it will be taken into account that the average woman's vocal tract is shorter and thus the formant frequencies of the mid-central, neutral vowel will be slightly higher. The forward and high position of the tongue assists in lowering the first formant frequency as F1 frequency "varies inversely with tongue height" (Kent & Read, 1992:22). Sundberg (1987:22/3) advises that the shape of the tongue influences the second formant (F2) frequency. The second formant frequency will increase if the tongue is forward in the mouth due to the influence that tongue advancement has in the anterior-posterior dimensions in the vocal tract (Ibid). This pattern correlates exactly with the vowel that is used in the Y-buzz. The [i] vowel has a typical decreased F1 frequency and an increased F2 frequency (Titze, 1994:166). It can be reasoned that the return to and stress on the Y-buzz, from and after the first part of the diphthong [əi] of the + Y-buzz, will influence the formant patterns of the first part of the + Y buzz diphthong so that a lower first formant (F1) frequency will be present.

The agility of the tongue as one of the main articulators is of cardinal importance to the voice user in theatre, as messages will not be conveyed as clearly as possible should the vowels and consonants not be articulated as clearly as possible. The other main articulators are the lips and the jaw (Perkins & Kent, 1986:119). The method of focussing on the space between the teeth as an ever-changing volume instead of the focussing on the jaw may be a good teaching tool, as the jaw will then be allowed to open in a released way without having more pressure put onto it. It often happens that voice users force the jaw down and that too much pressure is then put on the jaw-hinge. The sensation of the yawn will lead to an "open throat," which will translate as a vocal tract with as little constriction as possible where only the necessary local constriction for the vowel produced will be used. The lowering of the jaw, in itself results in a raise of the first formant (Sundberg, 1987:22). Titze (1994b:165) mentions that the attempt to get the yawn sensation will lower the larynx, which will lengthen the vocal tract (in the same way as the protruding lips will). The formant frequencies will thus again be lowered. In Chapter 3 specific references will be made to the direct correlation between the main articulators and the specific formants that they are affecting.

Munro, Leino and Wissing (1996) did some preliminary studies on the effect of the Y-buzz as a teaching tool for voice projection. A favourable comparison between the "actor's formant" (Leino, 1993, 1995) and the LTAS analysis of the Y-buzz-trained men's voices was made (1996:34). The results of this investigation will be dealt with in Chapter 3.

Given the above overview of current literature in voice science dealing with the Lessac Approach, it can be concluded that although some researchers and educators have reservations regarding aspects of the Lessac Approach, the majority who are writing and doing research about the approach, and specifically about the Tonal NRG, react positively to it in one or other way. The "Lessac trained" teachers using this approach, either as the only pedagogical tool or, as part of their pedagogical approach in teaching, reacted favourably towards the Lessac Approach as a teaching tool. It is thus necessary to deduce further from existing literature why the Tonal NRG may have a positive influence on voice building and voice production. To achieve this goal it is necessary to borrow from a scientific investigative field that is already further developed than that of the performer's speaking voice, namely, the acoustic characteristics of the singing voice. These acoustic properties of voice will be the central thrust of the following chapter.

Chapter 3

The acoustic characteristics of the "projected" voice.

3.1. Introduction and Overview

One of the most important characteristics of the successful performer's voice is the ability to project. Projection, or more specifically "carrying power", is an acoustic characteristic of the performer's voice (Leino, 1993:209). Sundberg (1988:12) states that the acoustic characteristic of voice is determined by two factors. The first is the vibratory process of the vocal folds, also known as the voice source, and the second is the shaping of the vocal tract in order to optimally resonate each sound that is produced with the vocal folds. The vocal tract thus acts as a filter, or a "frequency selective transmission system" as described by Kent and Read (1992:13).

Each resonator, depending on its length and shape, has the ability/tendency to enhance certain frequencies that pass through it, more than other frequencies (Kent & Read, 1992:17). These favourable frequencies are called the "resonance frequencies" and, in the case of the vocal tract, are referred to as "formant frequencies" (Nair, 1999:44; Sundberg, 1988:12). Should a sound that contains these frequencies by way of harmonics or overtones pass through the vocal tract, these frequencies will be enhanced with a greater amplitude than the other harmonics or overtones that make up the complex voice signals. Should these enhanced frequencies happen in groups, they are referred to as clusters. The vocal tract is highly moveable and adaptable and thus constantly adjusts shape and length (Nair, 1999:42) in order to optimalise the sound produced in the voice source. The moveable structures in the vocal tract are referred to as articulators and are the mandible, tongue, lips, soft palate, larynx and walls of the pharynx (Miller, 1986: 256-268; Sundberg, 1988:12).

The fact that these articulators are so moveable is of extreme importance to humans, as all voice and speech formations happen because of this moveability. All vowels and consonants are formed by moving the articulators in specific ways required in order for the human ear to discriminate between the different sounds. Perkins and Kent (1986:6) posit that one of the design features of speech is the distinct differences between sounds so that the change from one sound to another in a word will be sufficient to "change meaning". The resonator's malleability further provides potential formants.[1] These formants are different for different sounds and so each vowel, for instance, has its own formants or formant patterns (Miller, 1986:259). These formant patterns will be somewhat influenced by the pitch of the sound, but not to the extent that the vowel is not recognizable to the ear – unless the pitch of the sound is extremely high, which is usually the case for the soprano, and sometimes tenor, singing voice in the upper range.

There are generally five formants, below 5000 Hz, that are of importance for the analysis of the human voice (Leino, 1993:207; Stone, 1999:161). These formants each have an acoustical as well as a physiological "characterisation" – this means that physiological change to the voice production system will directly impact on the acoustic patterning of the sound. With the utterance of a single vowel it is easy to determine these formants. When an LTAS of the voice (singing or speaking) over some time, using various words is made, it is impossible to pinpoint the various formants. Therefore, references are made to the clusters around certain frequencies.

The fundamental frequency (F0) determines the vocal pitch. F0 is a direct result of voice source function. The first and second formants are F1 and F2. These are primarily responsible for vowel recognition (www.ncvs.org/vpt/tutorial/filter.html). For example, when auditively recognising a vowel as an [i], F1 would be lower in frequency than the expected F1 in a uniform pipe and the F2 higher in frequency on the spectral analysis, whereas when the auditive recognition will be an [a], F1 will be higher than the expected F1 in a uniform pipe and F2 lower in frequency on the spectral analysis (Nair, 1999).[2]

F1 is influenced by the mandible as an articulator (Sundberg, 1988:12). An increasing mouth opening raises the frequency of F1. A decreasing mouth opening and protruding

---

[1] As will be demonstrated later, each formant has two characteristics: a formant frequency and a bandwidth (Kent & Read, 1992: 18,19).

[2] Laukannen (personal e-mail 2002) posits that F3 may sometimes also be influenced by language.

lips lower the frequency of F1. The manipulation of F1 with the adjustable mouth opening, leads in the soprano voice to the phenomenon that F1 is clustered around F0 as F0 becomes higher than what F1 would normally be.[3] This aids in the projection of the high female singing voice and happens due to the fact that the soprano singer releases/drops the mandible as low as possible to create a bigger mouth opening (Sundberg, 1987:132; Miller, 2000). This may also explain the phenomenon that F1 moves closer to F0 where the lower speaking or singing range is used, as the smaller mouth opening and longer vocal tract will lead to the lowering of F1 beyond its normal frequency.[4] The vertical position (or height) of the tongue in relation to the space between the teeth also has an influence on the frequency of F1 (Kent & Read, 1992:24).

F2 depends on the influence that the shape of the tongue has on the oral cavity (Kiukaanniemi, Siponen & Mattilla, 1982:23), and specifically to the sagital orientation of the tongue in the mouth. The tongue divides the mouth sagitally into a front and back cavity. Kent & Read (1992:23) refer to this as the "tongue advancement." Nair (1999:94) advocates that the term "tongue advancement" has to be substituted in the singing voice studio by the term "longitudinal movement" as it implies movement of the tongue in either the sagitally forward or backward direction. Should the tongue move to the front of the mouth, then the front cavity will become smaller and F2 will rise in frequency. The same logic holds then when the tongue moves posteriorly: the front cavity increases in volume/size and the back cavity decreases. This lowers the frequency of F2.

F3 also influences vowel characteristics but is linked to F4 and F5 for voice quality. F4 and F5 are responsible for a part of voice quality and, in combination with F3, for the projection capabilities of a voice as will be seen in further discussion below. F3 is, according to Sundberg (1988:12), sensitive to the "position of the tip of the tongue or, when the tongue is retracted, to the shape of the cavity between the lower incisors and the tongue."

---

[3] This phenomenon is reliant upon a standardised, 17,5 cm uniform pipe as a resonator (the average length of the male vocal tract), where F1 is usually located at 500Hz (Kent and Read, 1992: 14-15).

[4] This obviously reflects in the research done by Raphael and Scherer (1997) where they found F1 to drop in frequency with the use of the Call, taking into consideration that the Lessac Approach advocates the use of the forward facial orientation.

The vocal tract shaping and lengthening that take place that are not related to vowels and consonants, are responsible for voice quality. It is the laryngeal tube that acts as "a separate resonator" to provide the resonance characteristics that lead to a well projected singing voice (Sundberg, 1974:842). It is these higher formants or formant frequencies that relate directly to voice quality that is of importance to this study.

3.2. The "Singer's Formant"

As early as 1934, Bartholomew (Nair, 1999:46; Leino, 1993:206) observed a specific cluster of formants in sonograms of the singer's voice. This cluster was around 2800 Hz. Vennard (1966) observed the same phenomenon. Coffin (1976) designed a whole teaching method focussing on "matching formants with the harmonics of the voice source" (Coffin, 1976; 1980).[5] Sundberg (1974) did an in-depth study of this phenomenon and coined it the "singer's formant." Sundberg explains that it is a cluster around 2,8kHz that "belongs to the acoustical characteristics of professional male singing in Western Opera and concert performances" (1974:838). This formant cluster is independent of vowel quality and pitch. Sundberg (1987:119) reasons that this "singer's formant" is "a clustering of the third, fourth and fifth formants" with a smaller frequency separation between each formant, that leads to an increased ability of the vocal tract to transfer sound. Due to this increased ability to transfer sound, the frequency range of these formants increases. This, together with optimal closing of the glottis in the vibratory act (Leino, 1993:209), will lead to an increase in amplitude of the higher frequencies, or, as Laukkanen (1995:14) explains, "the spectral slope becomes less steep together with SPL (Sound Pressure Level) increase as a consequence of the increased glottal closing speed." Sundberg (1988:13) suggests that the "singer's formant" phenomenon is also present in the female alto voice and that the frequency of this formant cluster differs slightly, depending on voice type (Stone, Cleveland & Sundberg, 1999:161). Sundberg (1988:19) also indicates that there is a resemblance of the singer's formant in the professional soprano voice with "clearly higher level of partials in the 2-4kHz band" but that this level is not as clear as in the other voice types. This comes about because of the range of the soprano voice. Dmitriev and

Kiselev (1979:240) foreshadow Sundberg's view on the phenomena that the cluster is different for each voice type, and indicate that there is an observable formant cluster for the soprano voice. In their study they have found a favourable comparison between the formant clusters and the length of the buccopharyngeal tract used by each voice type. This research is supported by Leino (1994:45) who indicates the same tendency of different frequency clusters for the singer's formant due to different voice types. As a matter of fact, Leino stresses that this tendency should be used as a guideline and not be exclusive or inclusive.

Bartholomew suggested that this cluster around 2800 Hz was due to a constriction in the epilarynx (Nair, 1999:46). This correlates with Estill's (1986) research that suggests an aeryepiglottic sphincter constriction responsible for this acoustic characteristic. Sundberg (1974:839; 1987:121), however, supports Vennard's (1966) idea that F4 and F5 (which are not affected by the articulators that shape the front part of vocal tract for the vowels) are dependent on the laryngeal vestibule and the sinus piriformes. The size and shape of the "back part" of the vocal tract is thus responsible for the voice quality or the projectibility of the singer's voice. Titze (1998:27,28) enhances this view when he suggests pharyngeal widening as an important contribution to the tonal balance of the singer's formant. Laukkanen (1995:11) reminds, however, that "voice production at a laryngeal level" happens subconsciously and is not under voluntary control. Voice users use auditory and proprioceptive means to manipulate these structures. Despite these various ways of explaining the shaping processes of the resonator, what is of utmost importance to this study is that there is a specific F3, F4 and F5 formant cluster present in the well-projected classical singing voice.

An investigation into the formant patterns of professional Country Singers (Stone et al., 1999) indicated that the formant patterns of the professional country singers' singing compared favourably to the formant patterns of their normal speaking voices.[6] In some

---

[5] It has to be taken into account though, that the matching of the formants with the harmonics is not always in direct correlation to the "singer's formant" although it may contribute to the "singer's formant" as phenomenon in the well-projected voice.

[6] It has to be noted that the comparison was made between the formant patterns of their singing and normal speaking and not with the use of the speaking voice as an instrument for performance, thus the projected speaking voice.

cases the formant frequencies were a little higher in the singing than in the speaking voice, but that may be simply due to the fact that the fundamental frequency during singing was higher than during speaking (Ibid, 66). It may also be argued though, that the country singers used a higher laryngeal position during voice production. Research into the formant patterns of belting indicated that the higher harmonics are enhanced (Bestebeurtje & Schutte, 2000:202) to create the necessary loudness and brightness expected from the belting style. This was previously foreshadowed by the work of Estill (1980; 1984) who posited that "belting projects slightly higher than opera"[7] (1984:3), indicating that the acoustic energy of the belting sound is higher than 3kHz.[8] This correlates with a finding by Schutte and Miller (1993) who commented on F1 adjusted upward to "stay close to H2".[9] They further commented that this adjustment is thus also a way of formant tuning[10] albeit used differently to that of the way a classical singer would use the concept. The different formant patterns characteristic of different singing styles indicate that there is a different "configuration in the vocal tract" (Stone et al, 1999:161) for each different style of singing.

Nair (1999:51) states that "well-executed song is not speech." He explains the difference between singing and speaking, and although he is referring to "every day speech" (Ibid, 50) and not to the performer's speaking voice, most of the differences are valid. In singing the vowels are usually more lengthened, pitch range usage usually wider (see also McKinney, 1982:169), and the use of rhythm determined by notation is very specific and controlled. The difference in volume between singing and conversational speech also has to be noted – the singing voice is much louder than the conversational speaking voice (Bloothooft & Plomp, 1986). This is reflected in the very clear formants that can be observed in the well-executed singing voice (McKinney, 1982:170). Titze (1993:248) mentions that the performer's speaking voice differs acoustically from the classical singing voice, as far as formants are concerned, seeing that no prescribed F0 frequency (pitch) is

---

[7] Indicating towards a higher peak frequency of the clusters typically observed in belting as opposed to the "singer's formant" of classical singing.

[8] Although Estill does not refer to the term "singer's formant" it is clear that she is referring to the acoustic energy cluster that is commonly referred to as the singer's formant.

[9] H2 is an abbreviation for the second harmonic.

[10] "...singers tune their two lowest formant frequencies to harmonic partials in order to increase the audibility of the voice" (Carlson & Sundberg, 1992:256).

expected from the professional actor's voice. It is thus necessary to study the professional actor's speaking voice as an independent subject.

3.3. The "Actor's Formant"

Fant (1970) observed earlier that F3 and F4 seemed to be closer in trained than in untrained speaking voices (and even closer for the trained singing voice). Leino (1993:206) comments that very few studies on the "spectral characteristics of good professional speaking voice" have been done and that this is necessary, as it may contribute to the outcomes for voice training in theatre training. This paves the way for much needed research in this field.

Leino proceeds with several studies on the quality of the professional actor's speaking voice. His subjects all have Finnish as a mother-tongue. It is of importance for this specific study to reflect on the results of two of the projects executed by Leino. The first research project reported about is a "Long-Term Average Spectrum (LTAS) study on speaking voice quality in male actors" (Leino, 1993). Recordings of 48 male professional actors' voices where played to several experts in the fields of theatre, speech therapy and voice research. These expert listeners perceptually graded the recordings into categories of good, fairly good, rather poor and poor voice quality. This perceptual analysis was done subjectively as no definition for good voice quality was provided. Acoustical analyses, by means of LTAS, were done of the voices in these four different groups and comparisons were made amongst individual voices in each group and amongst the different groups (1993:207). This reflects two of the three ways that voice can be analysed: acoustic and perceptual. The third one is the physiological (Laukkanen, 1995:13; Miller & Schutte, 1999:206). Results indicated that the voices perceptually defined as "poor" displayed a clearly steeper spectral slope than the voices defined as "good." Of greater importance, however, was the "most notable" peak near the region of 3500 Hz. Although all four different groups displayed a tendency toward this peak, the group that was perceptually defined as "good quality" displayed a higher amplitude than the other voice groups in this frequency cluster. It has to be mentioned, however, that this peak should be evaluated in relation to the other peaks (1993:208) as it was proven that in some cases the 3-4 kHz peak was too strong and in these cases the voices were perceptually placed in the "poor" voice quality group. The question arises whether this was due to the use of "pressed voice," as this will be in line with other literature (Gauffin & Sundberg, 1989; Sundberg & Gauffin, 1978). A further characteristic which differentiates this group was not only the clearer relative amplitude difference, but also the fact that the valleys around these peaks became deeper, which indicates that this results from a formant clustering, possibly F4 and F5 (Leino, 1993:209). An amplitude difference between peak and valleys of more that 10dB was reported. Although there were deviations where an actor's voice was perceptually rated as "good", but where the acoustical analysis did not reflect this 3500 Hz cluster, as a whole it seems that there is a strong indication that what is classified as a good actor's voice, will be reflected in a LTAS analysis with this frequency cluster between 3-4kHz with increased amplitude and definite valleys around this phenomena.

Leino coined this phenomenon the "actor's formant"[11] and relates it to the previously mentioned "singer's formant." The difference between these two phenomena seems to be a higher amplitude level in the singer's formant with a lower frequency, while the actor's formant has a slightly lower amplitude level but a higher frequency (Ibid, 208).

In "On the effects of vocal training on the speaking voice quality of male actors" (1995) Leino and Kärkkäinen report on a project where seven male students had, for eight months, extra training, above their normal voice training as part of their course work, with a real time analyser to provide visual feedback on their voices. The aim here was to verify the use of the visual aid in combination with "vocal exercises (that) consisted of nasal-vowel syllable strings produced aiming at a clear, bright, well-projecting voice quality" (1995:496). The students could monitor the spectrum of the voice exercises in real time, and could thus adjust the resonator in such a way that the peak around 3.5kHz would be present during all these exercises. Before- and after- recordings were played, at random, for perceptual evaluation. The importance of this study lies in the fact that the experts, in the fields of theatre and speech, indicated that the recordings with the clearer peak at 3.5kHz, and with a less steeper slope, were the preferred voice quality. These

---

[11] Other literature, for example Oliveira Barrichelo et al (2001) may refer to this as the "speaker's ring."

characteristics still persisted two years after the initial training albeit the fact that they were weaker in amplitude (Ibid).

These projects are supported by independent research done by Nawka, Anders, et al. (1997) on the "speaker's formant in male voices." Their study was done with German speakers as subjects. Their findings correlate with Leino's research as they indicate a less steep spectral slope and a frequency cluster around the "centre frequency of 3,400 Hz" (1997:425) "with borders of 3,150 to 3,700 Hz" (Ibid., 426) for the five actor voices that they have analysed. They indicate furthermore the tendency towards this cluster in the five normal male voices used in their study. Again supporting Leino (1995), they proclaim this cluster as a "distinguishing resonatory effect of the vocal tract" (Nawka et al., 1997: 427) that occurs as a characteristic of F4. It is evident from their report on this research that this acoustical profile is directly related to the sonorous sound and leads to an increase of intensity of the voiced sound.[12]

Research done by Oliviera Barrichelo et al. (2001:348)[13] indicates that the speaking voice of classically trained (male, as well as female,) singers reflects the formant characteristic of the "speaker's ring."[14] It has to be noted, though, that previous independent research done by Lundy, Roy, Roy, Xue & Evams (2000) did not find this correlation – an explanation for this finding may possibly be that their subjects were all still singers in training and not professionals.

As previously mentioned (in Chapter 2) Acker's research (1987) on the performance speaking voice indicates an increase in spectral energy in some frequencies when referring to the tonal quality of the Lessac Approach. Raphael and Scherer's findings (1987) support this as they posit that the Lessac Call (as part of the Tonal NRG) leads to a brilliance in voice quality. They specifically noted F1 "being tuned more sharply in call mode…" (Ibid, 84).

---

[12] " The steepness of the spectral slope (i.e the rate of decline) becomes less pronounced when the sonority or the intensity of the voice increases" (Nawka et al, 1997:422).

[13] By means of an LTAS analysis of both singing and speaking voice samples.

[14] Also read as "actor's formant."

Linking to Leino's research, Munro et al (1996:27-29) did a pilot study analysing untrained male voices versus Lessac trained male voices, with a specific reference to Lessac's Y-buzz, a sub-division of the Lessac Tonal NRG. They reported a strong indication of a frequency cluster in the region 3-4kHz with a less steep spectral slope[15] in both the Y-buzz and prose reading of Lessac's own voice and that of a Lessac trained male voice (Ibid, 33). It is of interest here that Lessac himself is American, although he speaks several other languages. The other subjects trained in the Lessac Approach were all South African with either Afrikaans or English as their first language. The sample recordings for this research were all done in English. This language diversity suggests that the Tonal NRG may not be a language specific teaching tool.

Concluding from the aforementioned research, it thus seems that the professional actor's voice, when it projects well and has a sonorous quality (Nawka et al., 1997:422), has as its characteristics, a less steep spectral slope and an enhanced peak at the frequency range between 3 and 4 kHz which seem to be related to F4 and F5 (Leino, 1993:209). It has to be noted that this was the tendency whether the subjects had Finnish, German, Afrikaans, English and/or American English as a first language.

So far most of the findings reflect on the possibility of the actor's formant in the male voice. Very little is known about the characteristics of the female actor's voice. It is generally accepted that it is more difficult to investigate the acoustic properties of the female voice.[16] Titze (1989) ponders on the idea that the source-filter theory of speech production might have been different should early studies have been based on the female voice.

---

[15] Scholarship always describes the slope of the trained voices as "less steep". This description compares the slope of the untrained and trained voices. The untrained voice slopes downward from left to right.

[16] "The woman's pitch has a higher frequency than the man's pitch..." (Mendoza, Valencia, Munoz & Trujillo, 1996). This leads to the harmonics of the female voice being further apart.

Leino (PEVOC[17] poster, 2001) reported on recent research done on the female Finnish voice. 20 Finnish female actors each read one minute of text. The recordings were played to a perceptual panel, all with at least four years of training in Speech Communication and Vocology. They subjectively differentiated between good and poor voice quality. An LTAS analysis of these voices indicated that both groups had the tendency to peak at 4300Hz but that the voices considered as good had a stronger peak (thus a higher amplitude). Leino concluded that good Finnish female speaking voices are, like the male counterpart, characterized by an actor's formant, but that this phenomenon is found around 4,3kHz. This is thus approximately 800Hz higher that the male actor's formant. This is in line with the idea that the female vocal tract is shorter than that of the male.

Weiss (1993:4)[18] warns against setting up acoustical guidelines for preferred voice qualities as his research indicates that there are different acoustical formant patterns according to language or, more specifically, language groups. In his research he makes a differentiation between Anglophone and Francophone languages. According to his findings (Ibid, 5) voice training affects voices differently depending on the language group and this is thus reflected differently in the spectral analysis of these languages. Even where it was stated earlier in this study that F4 and F5 are not vowel dependent, but voice quality specific, Weiss warns otherwise (Ibid, 6). It will thus be of importance for this study to reflect on his recommendations when analysing the LTAS of the non-Anglophone subjects.

It can be concluded that there is uniformity on the idea that different voice types, as well as different voice styles, have different specific formant patterns reflecting the sound quality of specific voice types or styles. There seems to be some conflicting evidence about the influence of language on the formant characteristics of the well-projected speaking voice.[19] Although this study presents an imbalanced representation of different language groups it may provide some indication towards a possible solution.

---

[17] Pan European Voice Conference.

[18] In this article Weiss used an LTAS analysis as means of investigating the speaking voices of 23 subjects.

[19] Laver (in Kent, 1997) indicates that the evaluation of voice quality is very complex, as organic and phonetics aspects are always at play.

If the claims made by Leino and others on the "actor's formant" holds true, then the training of the voice to achieve these formants needs to be considered. The Tonal NRG of the Lessac Approach as a teaching methodology seems to develop a voice that perceptually answers to the demands of this type of training. The following chapter will investigated the efficacy of the Tonal NRG in the training of projection of the female voice (which in turn will reflect the actor's formant).

Chapter 4

Empirical research

4.1. Problem Statement

As has been mentioned in Chapter 1, one of the "tools" that the actor/actress requires is a flexible vocal instrument that has the ability to project over some distance. This leads to the voice and speech teacher having to focus (amongst other things) in the training of the acting student, on voice building. Voice building is defined here as the building of the voice to deal with the rigorous vocal demands of performance in such a way that the performer is capable of providing the needed intensity and loudness without losing flexibility to "envoice" the inner intent of the performer/character. It is in this training situation that the teacher mainly has his/her ear as guiding tool to evaluate the vocal production of the student (see Kent, 1997).

Lessac claims that through his approach, and specifically the Tonal NRG of this Approach, voice building will take place in a healthy, holistic way. In the questionnaire done as part of this study and referred to in Chapter 2, the trained Lessac teachers all agree with Lessac's statement and see this approach as an excellent pedagogical tool. Verdolini (see Chapter 2) contributes positively to the understanding of the physiological effect of the Lessac Approach on voice production. Acoustically, research has indicated that the Lessac Approach enhances projection in the male voice (Acker, 1987; Munro, et al., 1996; Scherer & Raphael, 1987).

Although the Lessac teachers indicated that they believe that the Tonal NRG enhances the projection of the female voice, very little objective research has been done on this topic.[1] The main purpose of this chapter is to determine whether the Tonal NRG enhances the voice quality and projection capabilities of the female voice. A further belief of most of the Lessac teachers[2] is that the principles of the Tonal NRG are not bound to any specific culture or language. Consequently, the second purpose of this chapter is to investigate the possibility of the principles of the Tonal NRG not being language bound. This may obviously then bode

---

[1] Only the work of Raphael and Scherer (1987) and Acker (1987) reflect acoustically on this.

[2] This is reflected in their responses to the previously mentioned questionnaire.

well for the use of this approach in the multi-lingual, multi-cultural tertiary educational system in South Africa.

4.2. Empirical Research Aim

The overall aim of this study is to investigate whether the Tonal NRG of the Lessac Approach will enhance the voice quality[3] of the female actor's voice irrespective of

- Language or
- teacher and training methodology specificity.

Three different training groups and one control group (discussed in different sections below) will be used to conduct this study.

As a first phase of this research project an independent study[4] was done. The outcomes of this study were presented as a poster at PEVOC,[5] Stockholm, 2001. This study will be discussed in Section One and will be used in this project, as a base study on which to model further research.

4.3. SECTION ONE: American Test Group

4.3.1. Aim

To investigate the effects of the Tonal NRG of the Lessac Approach (three primary explorations[6] as well as English phrases and Call words), within an intensive workshop situation, on the acoustic quality of the American female actor's voices.

4.3.1.1. Sub-aim

To compare and interpret the graphic presentations of the LTAS of the pre-training and post-training recordings, by means of a visual comparison.

---

[3] See 1.1 p.3 for a description of "voice quality".

[4] In collaboration with Anne-Maria Laukkanen, Head, Department of Vocology, Tampere University, Finland.

[5] Pan European Voice Conference (Munro & Laukkanen, 2001).

[6] Y-buzz, +Y-buzz and Calls. Once recorded these explorations as well as the English phrases and Call words will be referred to as modes.

4.3.2. Participants

Seven American female actors, between the ages of 21 and 45, who attended an intensive six-week workshop, taught by Arthur Lessac himself and three of the certified Lessac teachers. None of the participants had any reported voice or hearing problems.

4.3.3. Ethical considerations

These women all voluntarily agreed to participate in this research knowing that the voice samples will be referred to anonymously and that their identities will be protected.

4.3.4. Training period

A six-week intensive Lessac Workshop.

4.3.5. Training process

During the workshop all the different aspects[7] of the Lessac Approach were covered. A typical workshop programme includes daily body work, big group sessions where new concepts are introduced and practised, small group sessions, as well as one-on-one sessions, buddy sessions and time for self-study and practical explorations.

4.3.6. Data collection

The three main vocal explorations, and the English phrases were recorded from the participants before and after training.

4.3.6.1. Recordings

- Recordings were done in an isolation booth (studio in College at Fredonia, New York State University, USA).

---

[7] During this specific workshop Structural NRG was introduced first followed by Consonant NRG and Tonal NRG. During the body sessions optimal body integration and breathing was introduced from the perspective of the Lessac Approach.

- a DAT recorder was used.
- An Audio Technica 4050 microphone (cardioid, flat response, no bass roll-off, no pad) was used.
- The mouth-to-microphone distance was 40 cm.
- A Sound Level Meter (CAT 42-3019) was used to control voice intensity. The aim was to keep the sound level between 65 -70 dB.

4.3.6.2. Tasks expected from participants

Initially the participants were asked to do an "uh-uh" sound. This was used to determine the pitch given for the Y-buzz sound. The pitch decided on was then given to each participant[8] when asked to do the Y-buzz for both recordings.

Instruction given for samples[9] 1-3: "Please do the (different sounds inserted – named and demonstrated) as long and as loud as is comfortably possible while staying in the parameters 65-70dB on the SPL meter."

Instruction given for sample 4: "Please read the following (either English sentences and call words or first language texts) in a comfortable volume for performers in speaking voice."

Different sounds named and/or demonstrated:

1. Y-buzz on certain pitch as determined. The pitch that suited the voice (as decided on for each participant) was given on the keyboard.
2. +Y-buzz
3. Calls
4. Additionally, the following English phrases and Call words were recorded:

"Leave me alone, I don't need grief!"
"It may rain today."
"He complained all day about the pain."
"H'lo, dover, potato, watermelon"

These phrases and words are used to help the trainees to transfer the principles of the exercises to speech.

---

[8] A specific note was played on a keyboard.

[9] Throughout this study the word "modes" will be used when referring to the different kind of utterances recorded. Mode 1 is the Y-buzz, Mode 2 the +Y-buzz, Mode 3 the Calls and Mode 4 the English phrases and Call words.

4.3.7. Data analysis and processing

Long Term Average Spectra (LTAS) were made with a Signal Analyzer (Hewlett-Packard 3561A) at the University of Tampere, Finland.
Visual inspection and interpretations of the graphic representations of the LTAS were done.

4.3.8. Results and discussion

LTAS for the three main Tonal NRG explorations and the phrases recorded before and after training can be seen in the Figures 1 for Y-buzz, 2 for +Y-buzz, 3 for Calls and 4 for English phrases and Call words.

The most consistent changes found after training were:

- An increase in the dB level difference between F1 and F0 in all three the sound samples of the explorations/exercises and in the phrases after training. The increase in F1-F0 dB level difference suggests better formant tuning of Fl to a harmonic, mainly to the second one. This seemed to be achieved by either changing F0 or the frequency of Fl. These results support the findings reported by Raphael and Scherer (1987). Better tuning of F1 is likely to improve the energy transfer from the vocal tract.
- A relative strengthening of the spectral peaks at the range of 2-4 kHz compared to the strongest spectral component (Fl) in the Y-buzz sound samples. This suggests a further improvement in the projecting capacity of the voice, since the threshold of hearing is lowest in the range 2-5 kHz.

What follow are randomly chosen figures to present the LTAS profile visually of the different modes. Figure 1 provides an example of the pre- and post- recordings of the Y-buzz of one participant. In Figure 2 an example of the pre- and post training recordings of the +Y-buzz of one participant is presented. In Figure 3 an example of the pre- and post- recordings of the Call of one participant can be observed. Figure 4 provides an example of the pre- and post-recordings of the English Phrases and Call words of one participant. These different visual profiles indicate clearly the increase in F1-F0 dB level difference, as well as the relative strengthening of the spectral peaks at the range of 2-4 kHz.

Y-BUZZ

BEFORE (TOP)
AFTER (BELOW) TRAINING

SUBJECT C

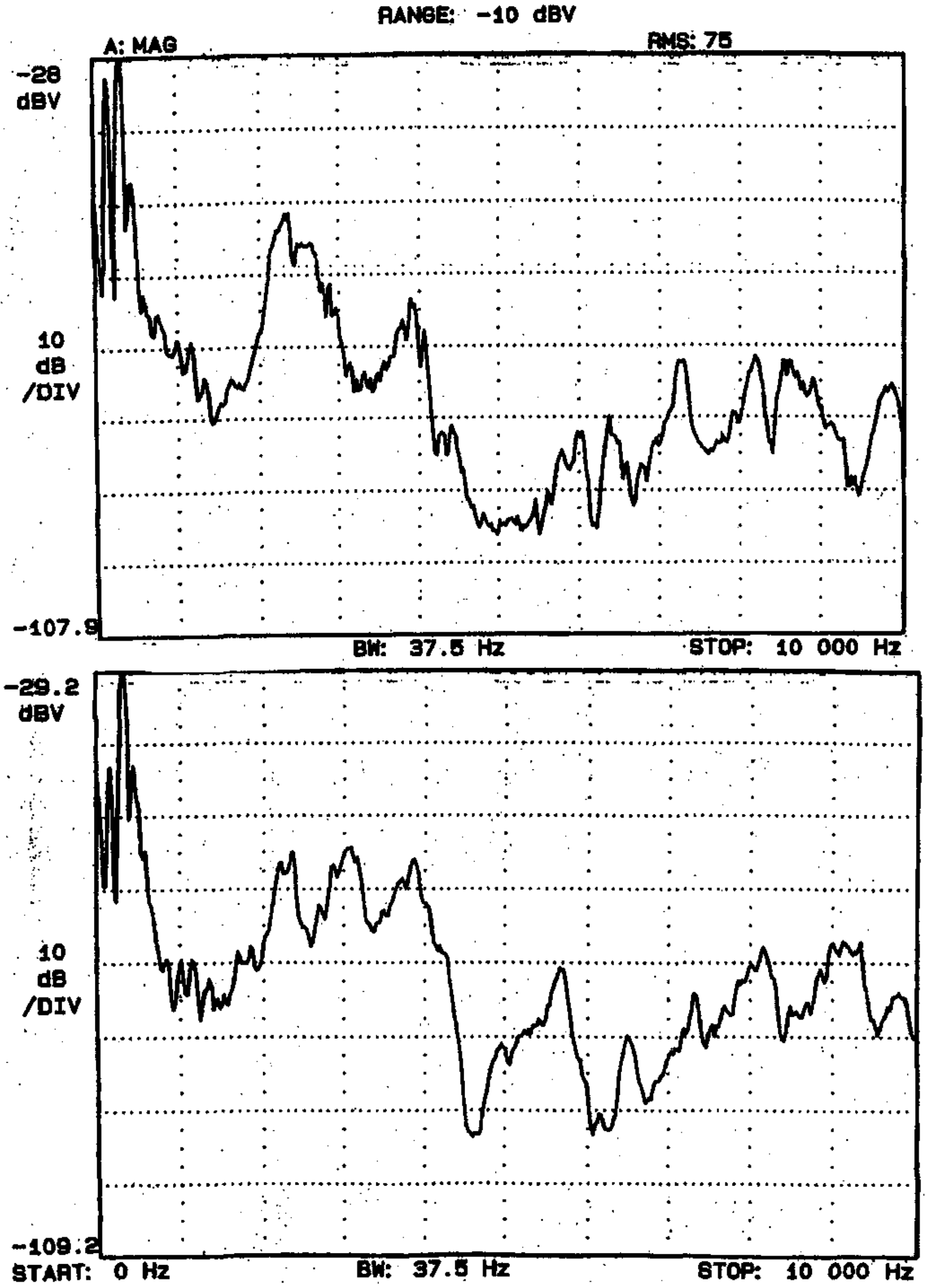


Figure 1: An example of the pre- and post-recordings of the Y-buzz of one participant.

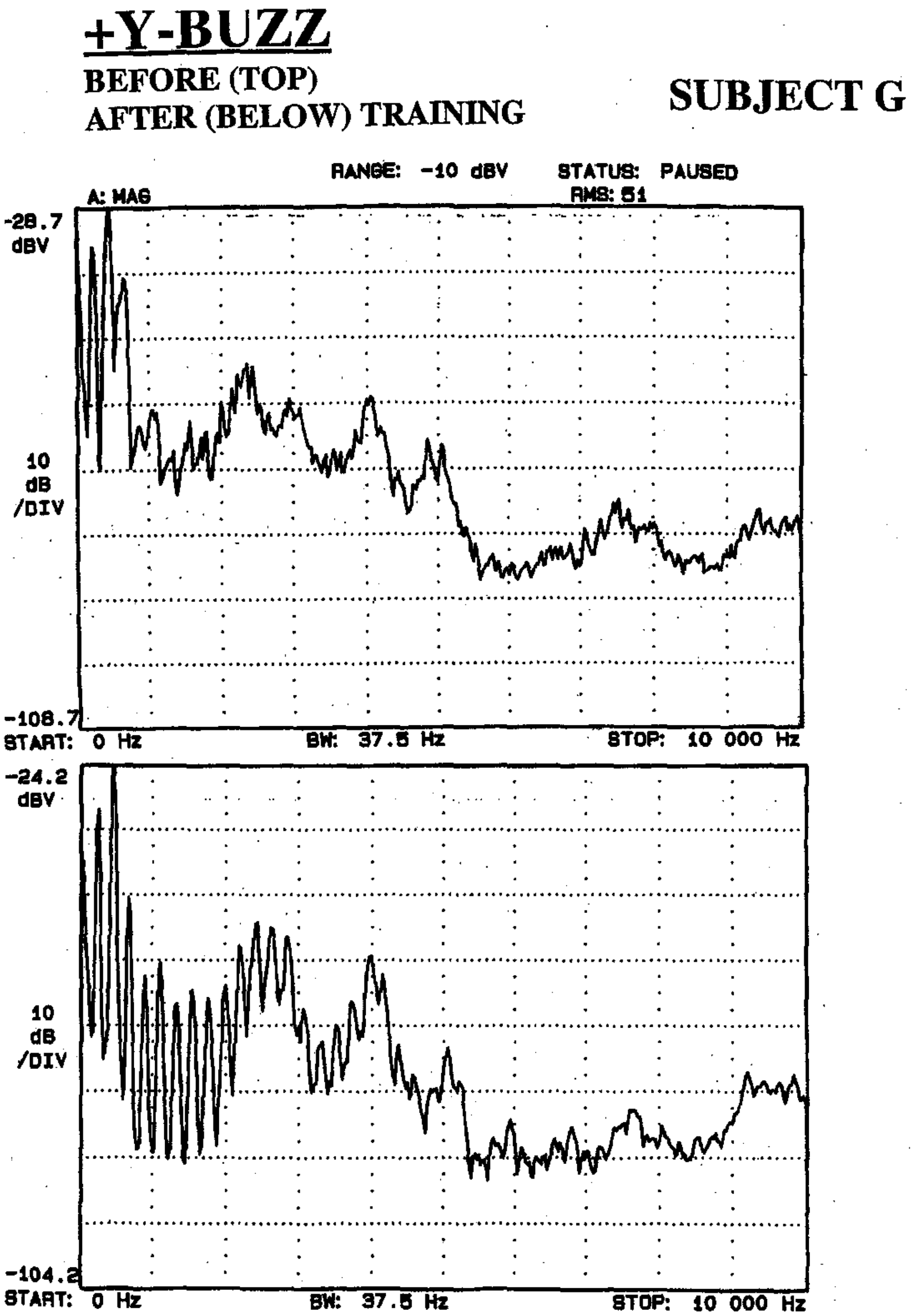


Figure 2: An example of the pre- and post-recordings of the +Y-buzz of one participant.

CALL

BEFORE (TOP)
AFTER (BELOW) TRAINING SUBJECT E

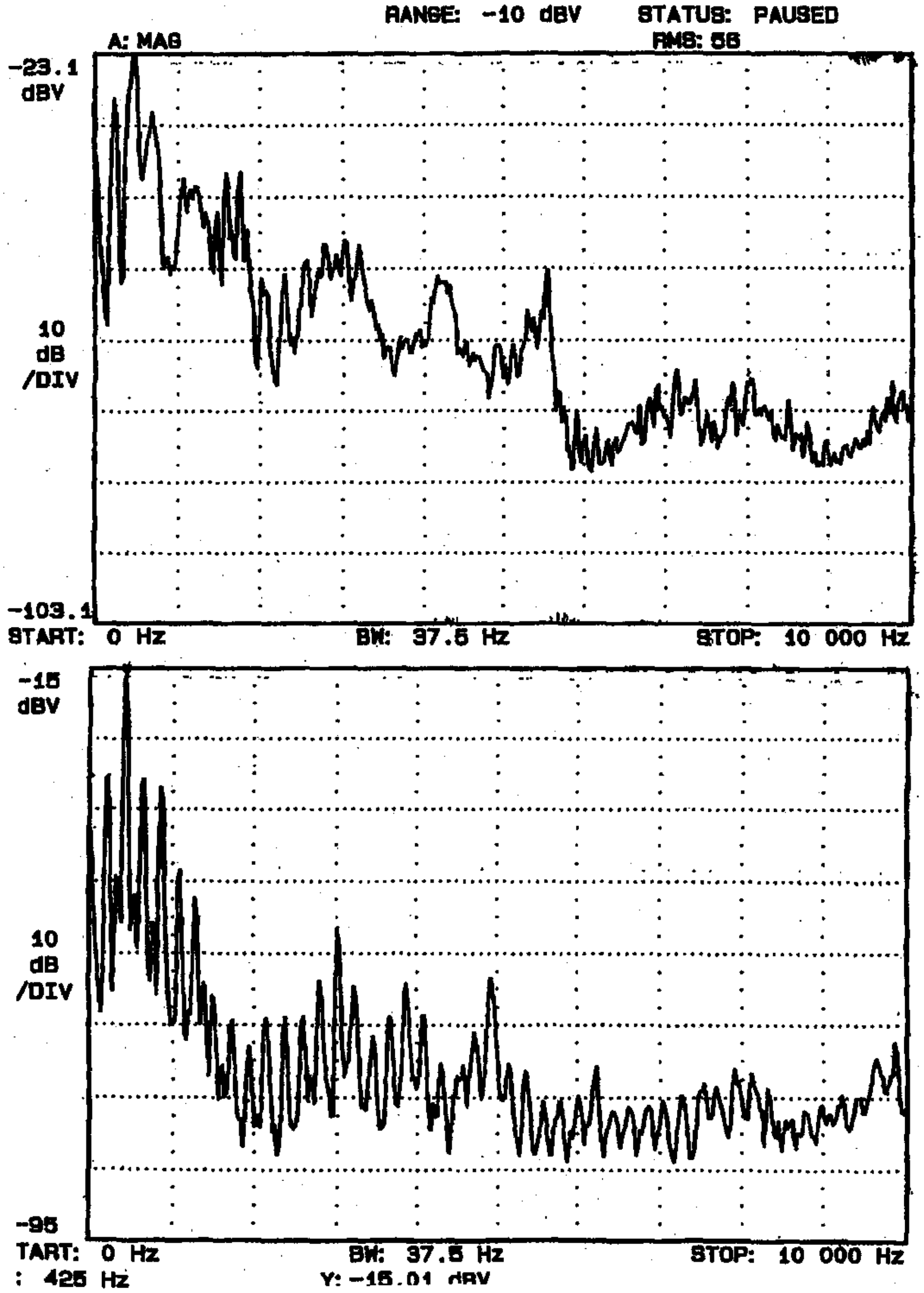


Figure 3: An example of the pre- and post-recordings of the Call of one participant.

PHRASES & WORDS

BEFORE (TOP)
AFTER (BELOW) TRAINING

SUBJECT E

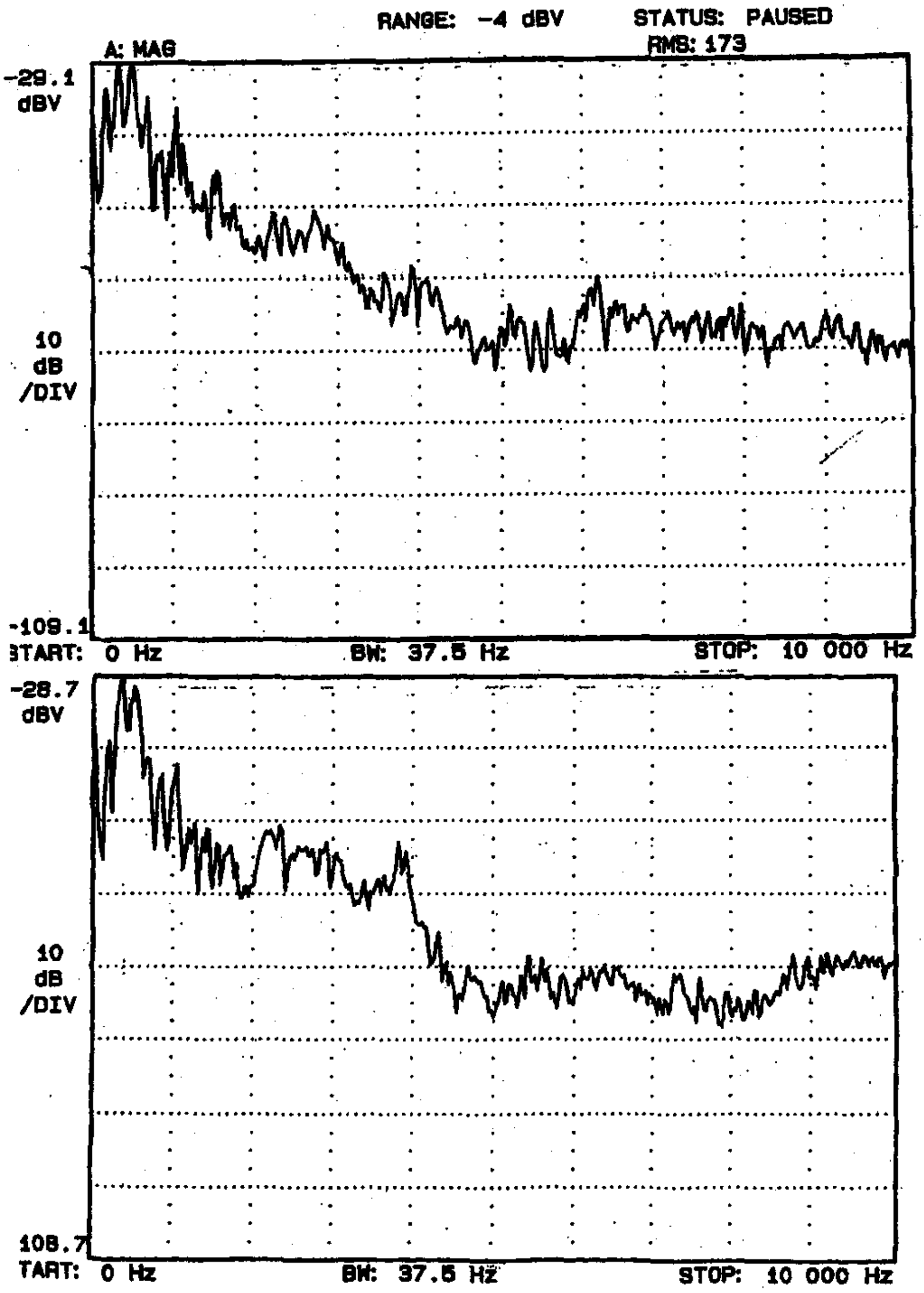


Figure 4: An example of the pre- and post-recordings of the English phrases and Call words of one participant.

In conclusion one has to take into account that phonation on closed vowels and with protruded lips increases the input impedance of the vocal tract. This may have beneficial effects on voice source and vocal fold vibration (Rothenberg, 1981 and 1988; Laukkanen, 1995). Results obtained by Verdolini et al. (1995; 1998) suggest that the "resonant voice" approach makes it possible to achieve sufficient carrying power in the voice economically, i.e. with less vocal effort and with less impact stress during voice production.

Evaluating Section One as pilot study in the research project, several questions arose which led to a possible refinement of the research design:

- The fact that there was no control group is problematic as it can be argued that possibly this was just a tendency towards "projection" seeing that the participant spoke louder during the post training recordings and this may lead to the increase of amplitude of the spectral peaks.
- The size of the test group is too small to generalise any of the findings statistically. A bigger test group will thus improve the quality of the study.
- The visual interpretations of the graphic representations of the LTAS may be one way of interpreting the difference in the pre- and post-training recordings but this can perhaps be supported by the use of statistics should the test group be bigger.
- There was no perceptual analysis of the pre- and post-recordings. This would contribute to the validity of the research as "the ear of the audience" is always of great importance, because this leads to aesthetic acceptability.

In the following sections attempts were made to overcome some of the shortfalls indicated about Section One above.

The necessity of a control group in the research design is eminent and supported in literature (Cohen & Manion, 1997:164-169). Yet it is immensely difficult to have a control group within the teaching set-up in any tertiary institute, for the following reasons:
All the students pay equal tuition fees and have the right to the same (and the best possible) training. Where previous research (Acker, 1987; Munro et al., 1996; Raphael & Scherer, 1987 Verdolini, 1998) have indicated that the Lessac Approach contributes positively to the

development of performance voices,[10] a dilemma exists in the fact that the lecturer is bound to offer all the students equal knowledge and skills. It is not possible to offer an initial control group the choice of the same training as that of the test group, should the investigation indicate the work done with the test group as positive, seeing that a specific syllabus within a specific time frame needs to be followed. The shortfall of a control group, due to these obstacles, is addressed in the best possible way in this study. The experiment of the control group is discussed in Section Two. In Sections Three and Four further research, investigating the use of the Tonal NRG as training mechanism, will be reported on. The control group is discussed first because, at the time of implementation, it was realised that there were still some foreseeable shortfalls in the design of the empirical study. These shortfalls – both in terms of design and analysis -- were adapted, once again, for Sections Three and Four.

4.4. SECTION TWO: Control Group

4.4.1. Aim

To investigate whether there is an energy/amplitude difference (within the spectrum F0 – 5000Hz) that enhances the carrying power of the female speaking voice when the intent is there to speak louder (as graphically demonstrated by the Long term Average Spectrum).

4.4.1.1. Sub-aim

To visually compare and interpret the graphic presentation of the LTAS of the first recording done when using conversational voice and of the second recording done when using a voice loud enough to speak on stage.

4.4.2. Participants

A call was put out for female volunteers who had no voice training, no experience of professional voice usage and no history of vocal health problems, who did not smoke. Five

---

[10] This suggests a contribution towards vocal health on the physiological level, and enhances the projection possibilities of the male voice as well as, when referring to Section One, possibly the female voice.

women volunteered, one was not used due to age.[11] The other four were all in the age group 20 – 26. Two of the participants have Afrikaans as a first language and two have an African language as a first language.[12]

4.4.3. Ethical Considerations

These women all voluntarily agreed to participate in this research being assured that the voice samples will be referred to anonymously and that their identities will be protected.

4.4.4. Training period and process

No training took place.

4.4.5. Data Collection

The three main vocal explorations/exercises and English phrases were recorded from the participants in a first ***and*** second recording. Except for the instructions on loudness as explained in 4.4.5.2. the modes recorded were duplicated. The second recordings were conducted on the same day as the first recordings, with a period of voice rest in between.

4.4.5.1. Recordings

- Recordings were done in a sound-treated room at a tertiary institution.
- Recordings were done onto a DAT (Sony ZA5ES Super bit mapping) recorder. Rec. level: 5; sampling rate of recording: 44.1kHz; Input: analogue microphone.
- Microphone used: Shure SM48. Dynamic LOZ- Unidirectional.
- Microphone-to-mouth distance: 40cm.

---

[11] This woman was 62 and the researcher was concerned that the aging of the voice will play a role in the acoustic profile of the voice (see Sataloff, 1991:146-150).

[12] This language difference was, however, not used, as the number of participants of each of these groups was too small to make any relevant contribution.

4.4.5.2. Tasks expected from participants

Initially the participants were asked to do an "uh-uh" sound. This was used to determine the pitch given for the Y-buzz sound. The pitch decided on was then given to each participant[13] when asked to do the Y-buzz for both recordings.

The instruction given for the first recordings of modes 1-3 was: "Please do the (different sounds inserted – named and demonstrated) as loud as what you would do it in conversational speaking voice." For the second recordings of modes 1-3 the participants were asked to do the different sounds as loud as what think they would have to do it when they are on stage.

The instruction given for the first recording of mode 4 was: "Please read the following English Sentences and call words in a comfortable conversational volume." For the second recording of sample 4 the mode were asked to read the English sentences and call words as loud as what is comfortably possible.

Sound modes 1-4 named and demonstrated were:

1. Y-buzz on certain pitch as determined. The determined pitch was given on keyboard.
2. +Y-buzz
3. Calls
4. English phrases and Call words:

"Leave me alone, I don't need grief!"
"It may rain today."
"He complained all day about the pain."
"H'lo, dover, potato, watermelon"

4.4.6. Data analysis and processing

Using KAY Elemetrics' Multi-Speech analysing functions, a Long Term Average Spectrum was done for each of the four sound samples of each participant. Settings used: FFT size: 512 (thus samples/frame: 257); pre-emphasis: 0.000; Window weighting: Hanning; Smoothing level: None. The numerical data of the acoustical profiles of the first and second recordings was imported into MS Excel where graphic presentations of the LTAS were made.

---

[13] A specific note was played on a keyboard.

In Figures 5,6,7 and 8 the average LTAS of the first and second recordings of each sound samples is graphically represented:

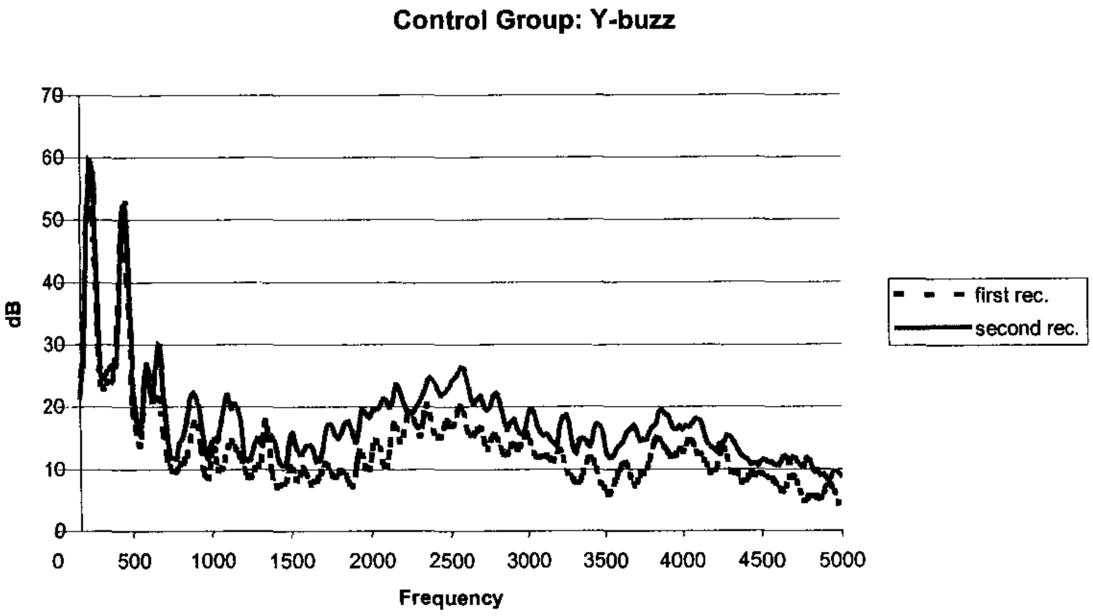


Figure 5: The LTAS of the first and second recordings of the Y-buzz of all four participants in the Control Group.

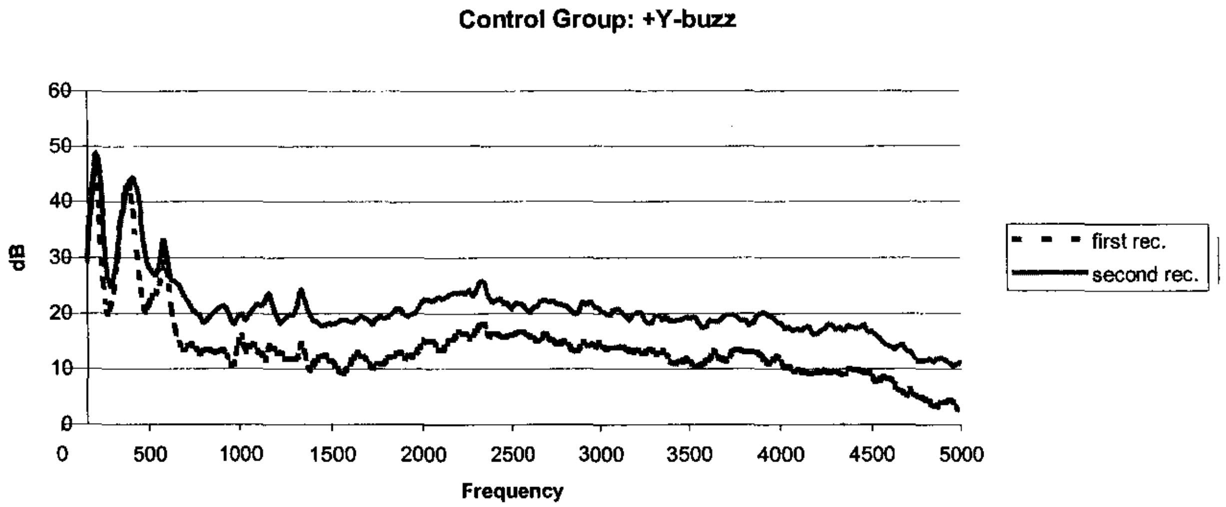


Figure 6: The LTAS of the first and second recordings of the +Y-buzz of all four participants in the Control Group.

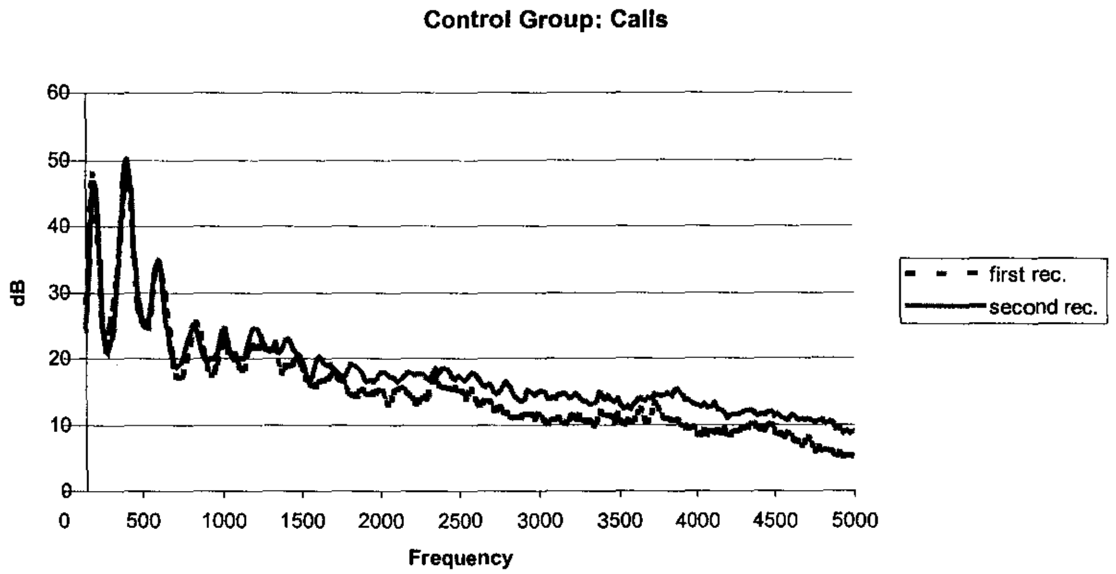


Figure 7: The LTAS of the first and second recordings of the Call of all four participants in the Control Group.

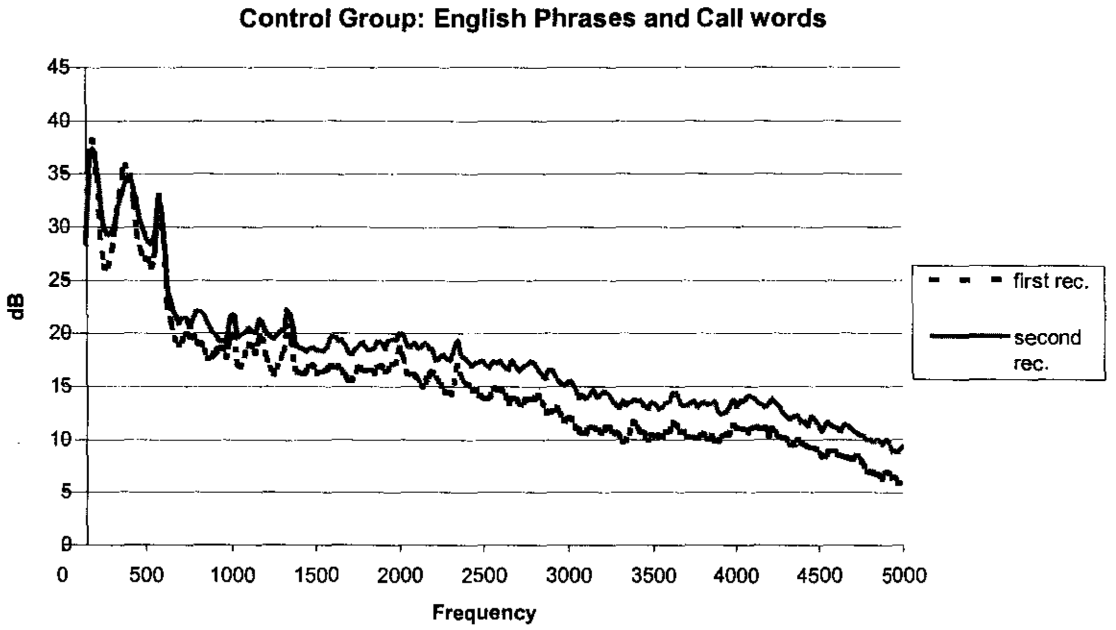


Figure 8: The LTAS of the first and second recordings of the English Phrases and call words of all 4 participants in the Control Group.

To show the method of processing the data that is followed, a step-by-step demonstration will be given. This will be done by means of using excerpts of numerical data of one participant who was selected at random. Table 1 indicates an excerpt (F0 to 495.26 Hz) of the numerical results of the LTAS of the Y-buzz of participant CA. The LTAS was done up to 5000Hz.

Table 1: An excerpt of the LTAS of participant CA.

Freq.	1st Rec.dB	2nd Rec.dB
193.8	54.33	52.64
215.33	53.84	59.82
236.87	38.7	54.93
258.4	23.91	27.65
279.93	21.91	22
301.46	23.69	23.92
323	23.76	25.21
344.53	24.38	24.66
366.06	31.03	22.6
387.6	52.7	24.18
409.13	57.76	45.1
430.66	50.4	53.67
452.2	25.63	49.8
473.73	17.01	26.56
495.26	14.42	16.75

Figure 9 posits a graphic representation of the numerical results (as seen in part in Table 1) of the LTAS of the first and second recordings of participant CA's Y-buzz.

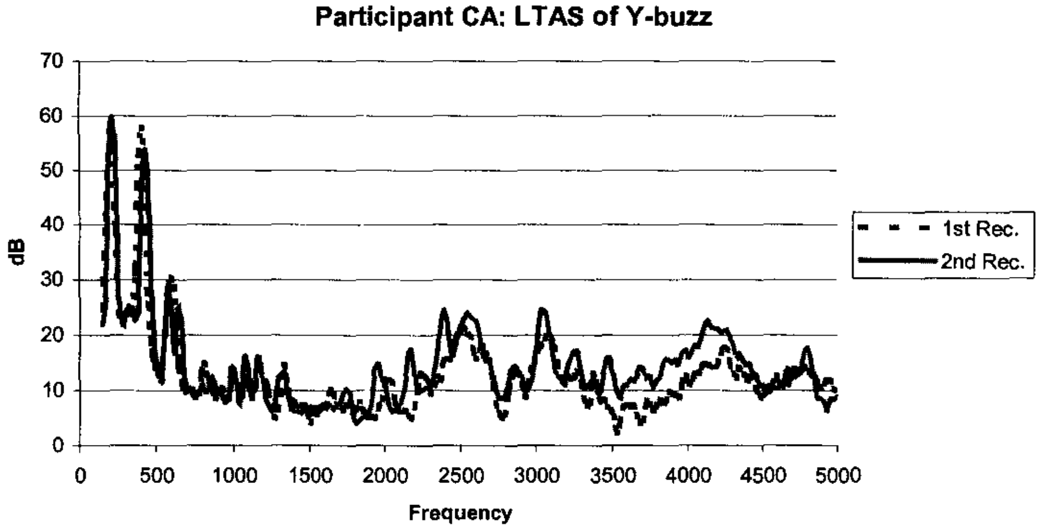


Figure 9: An LTAS representation of the 1st and 2nd recording of the Y-buzz of participant CA.

Following the LTAS an approximation of the numerical results was done where the highest dB within a window of each 100Hz was chosen. It was decided to use the window of 100 Hz seeing that the F0 was in all cases higher than 100 Hz and this system/approach would thus include the strength (dB) of each harmonic. Table 2 indicates an excerpt of the chosen highest dB's of each 100 Hz window of the numerical data presented in Table 1 (F0-5000Hz).

Table 2: An excerpt of the approximation of the highest amplitude for the frequency window of 100 Hz of the Y-buzz of participant CA.

Freq.	1st Rec.dB	2nd Rec.dB
F0	54.33	59.82
200-300	53.84	59.82
300-400	31.03	25.21
400-500	57.76	53.67

Following this approximation the amplitude of each 100 Hz window of the first recording was subtracted from the amplitude of each 100 Hz window of the second recording,[14] providing the amplitude difference between each frequency window of the 1st and 2nd recordings.[15] In an attempt to standardise these differences and interpret them in relation to the "loudness"[16] difference of F0, the difference between the amplitude of the F0 of the 2nd recording, and the amplitude of the F0 of the 1st recording, was subtracted from the differences of each of the 100 Hz windows. Table 3 indicates an excerpt of this process (F0-500Hz). It is thus clear what the energy increase in each 100Hz window is, above the difference observed in F0 dB. By using this method a clear indication of a less steep or steeper slope will be provided.

Table 3: An excerpt of the subtraction process: dB of 2nd recording minus dB of 1st recording, minus difference between dB of F0 1st recording and F0 2nd recording for all 100Hz windows. Participant CA: Y-buzz.

Freq.	1st Rec.dB	2nd Rec.dB	Difference	F0 diff.	Diff. in relation to F0
F0	54.33	59.82	5.49	5.49	5.45
200-300	53.84	59.82	5.98	5.49	0.49
300-400	31.03	25.21	-5.82	5.49	-11.31
400-500	57.76	53.67	-4.09	5.49	-9.58

A graphic presentation of the difference in relation to F0 was made for all the 100Hz frequency windows including F0. Figure 10 provides an example, again of the Y-buzz of participant CA, in order to clarify the process.

---

[14] Post (Z) – Pre (A)

[15] This is reflected in the fourth column of Table 3.

[16] The participantive term referring to energy and amplitude.

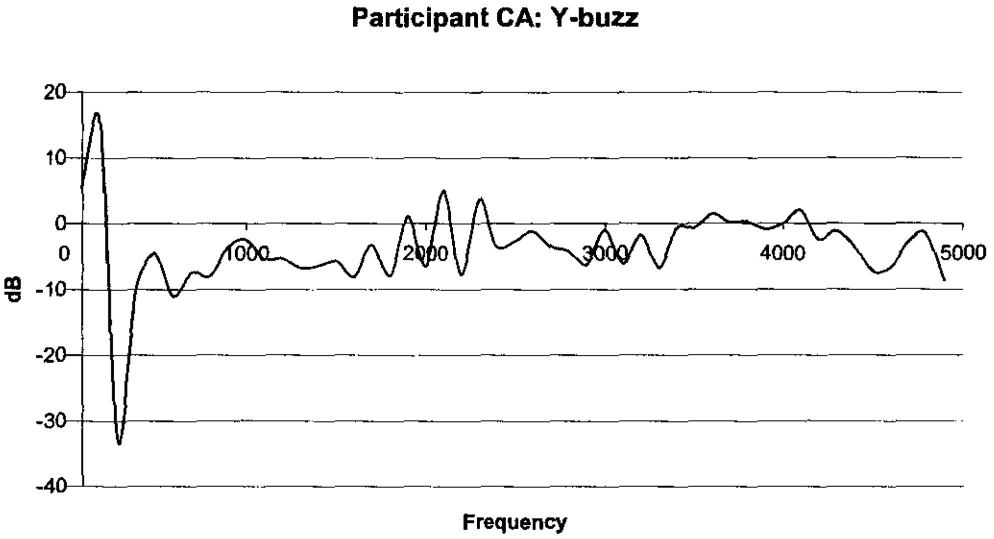


Figure 10: The difference in relation to F0 for all the 100Hz frequency windows, including F0. Participant CA: Y-buzz.

The method of processing the data followed for the one sound sample of the one participant (CA) as demonstrated above was duplicated for the whole group of four participants for each sound mode.

The results for the four different sound modes of this are graphically presented in Figures 11,12,13 and 14.

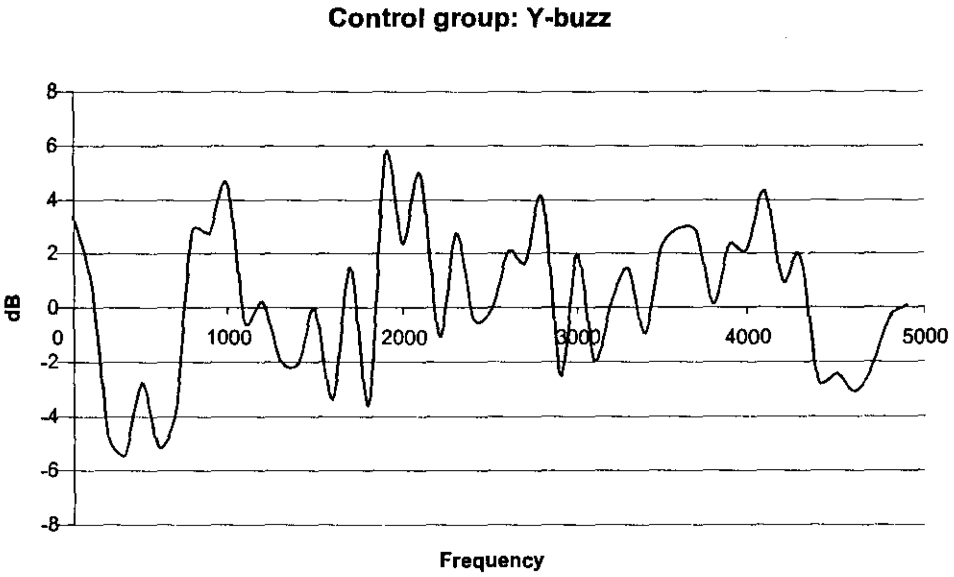


Figure 11: The difference in relation to F0 for all the 100Hz frequency windows, including F0. Y-buzz average of all participants in Control Group.

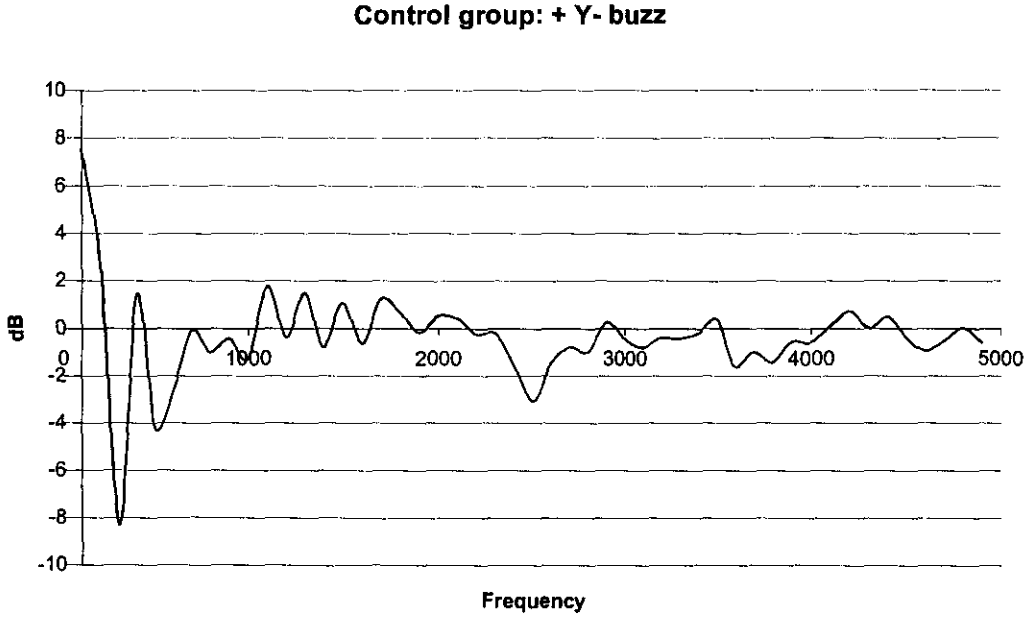


Figure 12: The difference in relation to F0 for all the 100Hz frequency windows, including F0. +Y-buzz average of all participants in Control Group.

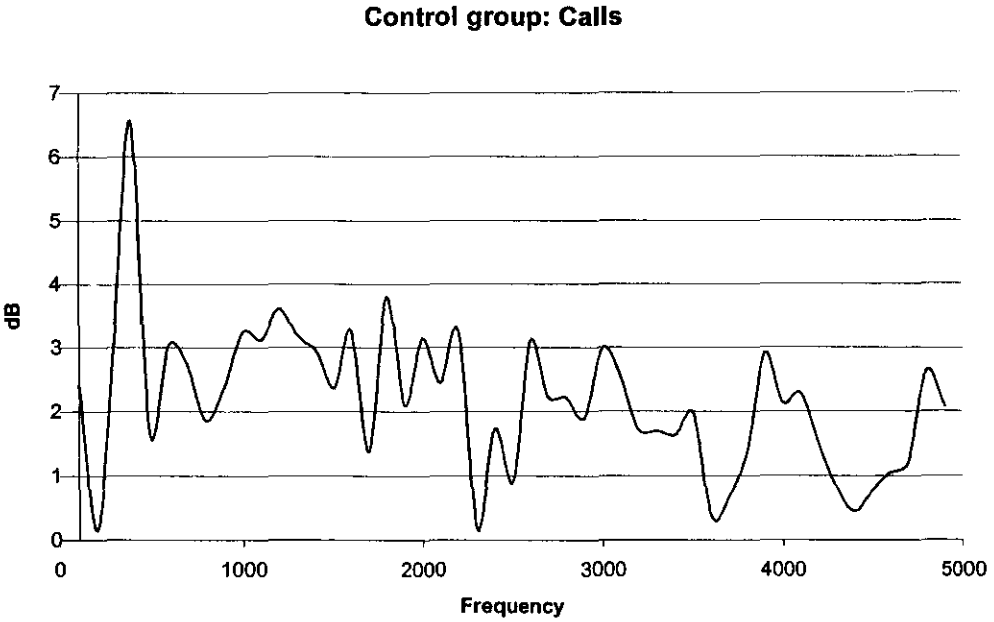


Figure 13: The difference in relation to F0 for all the 100Hz frequency windows, including F0. Calls average of all participants in Control Group.

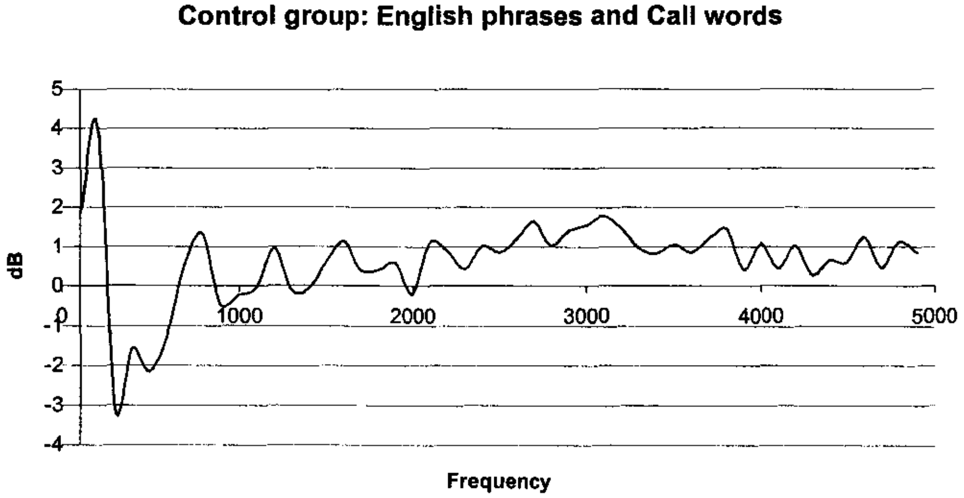


Figure 14: The difference in relation to F0 for all the 100Hz frequency windows, including F0. English phrases and Call words average of all participants in Control Group.

4.4.7. Results and discussion

The size of the control group does not lend itself to inferential statistics. However, general observations can be made. When interpreting the LTAS analysis of the control group of the Y-buzz (sound mode 1) a slight dB increase is observed for F0 between the 1st and 2nd recording. The graphic LTAS presentation (see Figure 5) further indicates a more or less evenly distributed energy increase over the whole spectrum (F0-5000Hz) with a slightly bigger increase between 1500-2900Hz as well as 3400-3800Hz. This is however not enough to change the visual profile of the graphic presentation. The difference in relation to F0 for all the 100Hz frequency windows (including F0), for the Y-buzz of the control group, provides more detail (see Figure 11). From the 1st to the 2nd recording certain frequency windows show a definite increase, albeit small, in energy above the increase in energy that took place in F0[17]: 800-900Hz by 2.85 dB; 1000-1100 by 4.5 dB; 1900-2000 by 5.65 dB; 2100-2200 by 4.93 dB; 2300-2400 by 2.72 dB; 2800-2900 by 4.04 dB; 3600-3700 by 2.9 dB; 3700-3800 by 2.83 dB and 4100-4200 by 4.31 dB. This information supports the interpretation visually gleaned from the LTAS but in a more in-depth manner.

When interpreting the LTAS analysis of the control group of the +Y-buzz (sound mode 2) a slight dB increase is observed for F0 between the 1st and 2nd recording (see Figure 6). The LTAS profile remains basically the same with the energy increase more or less evenly distributed over the spectrum (F0-5000Hz). A greater increase in energy between 800-5000 Hz can be observed. The difference in relation to F0 for all the 100Hz frequency windows including F0, for the +Y-buzz of the control group, provides more detail (see Figure 11). From the 1st to the 2nd recording certain frequency windows show a definite increase in energy above the increase in energy that took place in F0[18]: 400-500 Hz by 1.40; 1100-1200 by 1.74; 1300-1400 by 1.48; 1500-1600 by 1.05; 1700-1800 by 1.25; 2000-2100 by 0.52 and 4200-4300 by 0.73. Again, a clearer picture can be formed of the increase of energy over the whole spectrum analysed.

---

[17] F0 dB increased by 3.34.

[18] F0 dB increased by 7.84.

When interpreting the graphic presentation of the LTAS of the Calls (sound mode 3), as well as the English phrases and Call words (sound mode 4), of the control group, the same pattern that was observed in the Y-buzz and +Y-buzz determines. There may be a slight increase in the dB of F0 and a more or less even increase in the energy of the spectrum up to 5000Hz, but there is no definite change in the graphic profile of the energy display over F0 – 5000 Hz (see Figures 7 and 8). When interpreting the profiles (of both these sound samples) of the difference in relation to F0 for all the 100Hz windows, there is a strong indication that the main energy increase (which is indeed slight) happened in the F0 region and that the increase of energy for the spectrum up to 5000Hz, in relation to F0 was small. These changes may be due to an increase in breath pressure only, as there is no clear indication that the shape of the resonator was adapted to support projection of the voice.

In conclusion, taking the profiles of all four sound samples into account, it is posited that for the untrained female the aim, when asked to speak louder, or make a sound as loud as that which they perceived necessary if they were to be "on stage," is simply to "speak louder" and not to "project" over a distance. Of course, this experiment may be flawed due to the small number of participants. A more detailed study with a bigger participant group will have to done in order to verify these findings.

When reflecting, after completion of Section Two, on the possible areas of improvement identified in Section One, some observations still remain:

- Although Section Two deals with a control group, the number of participants is too small to make a definite contribution to this study and to allow a valid generalization to be made. This can be used only as an indication of the profile of the untrained female voice when asked to speak louder. Furthermore, the size of the control group is too small to analyse any of the findings statistically. A bigger control group will thus improve the quality of the study.
- The visual interpretations of the graphic representations of the LTAS were enhanced by the graphic presentation of the difference in relation to F0 for all the 100 Hz windows, as more detailed information could be gleaned from these but this can perhaps be supported by the use of statistics should the control group be bigger.

- There was no perceptual analysis of the 1st and 2nd recordings. The use of a perception panel would contribute to the validity of the research as "the ear of the audience" is always of great importance, as this leads to aesthetic acceptability.

In the following two sections (Sections Three and Four) an attempt is made to counteract the shortfalls (of design, implementation and analysis) that still remained after Section Two. The investigations will take place through (a) the use of a perception panel to reflect the view of trained voice teachers and theatre specialists, and (b) through acoustic analysis, where the formant patterns of the female voice will be analysed by means of an LTAS method and then graphically represented. Further statistically relevant processing will be used as means of investigation in these two sections. Although the experiments in these two sections have similar designs they will be discussed separately in order to follow the differences of these sections clearly. These differences reflect on the aim of this study as they present:

- different number of contact hours – reflecting on teaching methodology;
- different orientations towards language of instruction– reflecting on teaching methodology and language;
- different first language orientations amongst the two groups.

In Section Three I shall describe, analyse and discuss the experiment with Test Group S. In Section Four I shall describe, analyse and discuss the experiment with Test Group V. As mentioned above, the similarities and differences between the two groups will be indicated appropriately. The names of these groups were randomly assigned in order to protect the identity of the participants as explained under the ethical considerations.

4.5. SECTION THREE: Test Group S

This research with Test Group S is conducted with a group of female students in a voice building class that is part of the curriculum of an Actor training programme at a tertiary institution in South Africa.

4.5.1. Aim

The aim of this section is to investigate whether the use of Lessac's Tonal NRG will enhance the voice quality of the female actor's voice when the Tonal NRG is used in a voice building class in a tertiary institution in South Africa. The female participants had Afrikaans and South African English as well as several indigenous African languages as first language.

4.5.1.1. Sub-aims

- To determine, through the use of a perception panel, whether the preferred sound is contained in the pre-training or post-training recordings.
- To investigate whether any statistically significant difference, as indicated by the perception panel, between the pre-training and post-training recordings exists.
- To compare and interpret the graphic presentations of the LTAS of the pre-training and post-training recordings in relation to the F0 dB difference, in order to investigate whether an improved acoustic profile exists in the post-training recordings.
- To investigate whether any of the acoustic differences between the pre-training and post-training recordings is statistically significant.

4.5.2. Participants

- 15 female students.[19] Age between 18 and 23. Afrikaans: 8; English: 2; Tsonga: 1; Tswana:1; Southern Sotho: 1; Xhosa:1; Zulu:1.[20]
- The same Certified Lessac teacher taught both the groups in Section 3 and Section 4.

---

[19] The total number of students in the class was 45. Some were male, and some females did not participate in this study either because of vocal health issues or because they did not volunteer to participate in this study.

[20] Representative of student profile of the year when teaching was done at this tertiary institution.

4.5.3. Ethical considerations

These female students all voluntarily agreed to participate in this research having been assured that the voice modes will be referred to anonymously and that their identities will be protected.

4.5.4. Training Period

This group had a total of 28 contact hours over a period of 14 weeks.

4.5.5. Training process

The Lessac Approach was used.
Initially three-dimensional breathing and optimal body integration were explored, followed by the introduction and exploration of Tonal NRG as defined and described in the Lessac Approach (see Chapter 2 for details).
An organic developmental flow was crucial so as to follow the holistic nature of the Lessac Approach. The class was thus introduced to the Y-buzz and then proceeded through +Y-buzz to Calls and lastly to Call phrases.

Classes were conducted in English. The use of self-developed first language phrases as equivalents to the English Calls was encouraged.
Students were expected to work in "buddy groups"[21] on the explorations introduced in class, as preparation for the next class.

4.5.6. Data collection

4.5.6.1. Recordings

- Pre- and post-training recordings were done in a sound-treated room at the tertiary institute.

---

[21] Buddy groups are used within the Lessac Workshop set-up where two students are responsible for each other and work together during their preparation times between classes. This is an excellent tool, which guides the students towards discussion about the work and provides ear training.

- Recordings were done onto a DAT (Sony ZA5ES Super bit mapping) recorder. Rec. level: 5; mode rate of recording: 44.1kHz; Input: analogue microphone.
- Microphone used: Shure SM48, Dynamic LO~~Z~~ Unidirectional.
- Microphone-to-mouth distance: 40cm.
- A sound pressure level (SPL) meter was used as guidance to not overload DAT. Students were asked to not go lower than 65 dB nor higher than 75 dB. It seemed to be easy for the students to stay within these parameters during the pre-training recordings. They struggled to maintain this in the post-training recordings, even going as high as 90dB. When this occurred, they were then asked to repeat the recording.

4.5.6.2. Tasks expected from participants

In the pre-training recording students were asked to do an "uh-uh" sound. This was used to determine the pitch given for the Y-buzz sound in pre- as well as post- training recordings. Instruction given for modes 1-3: "Please do the (different sounds inserted – named and demonstrated) as long and loud as is comfortably possible while staying in the parameters 65-75dB on the SPL meter."

Instruction given for modes 4 and 5:[22] "Please read the following (either English Sentences and Call words or first language texts) in a comfortable volume for performers in a speaking voice."

Different modes named and/or demonstrated:[23]

1) Y-buzz on certain pitch as determined. Pitch given on keyboard as determined during pre-training recording.
2) +Y-buzz
3) Calls
4) English phrases and Call words
5) First language text readings of approximately one minute without the use of words containing an "s" sound.[24]

---

[22] Please note that mode 5 is the first language texts recorded.

[23] Seeing that the instruction indicates that the utterances were to be done as long as what is comfortably possible, the recordings had different lengths of time reflecting on the participants' levels of competence. The first language text readings were all more or less one minute.

[24] As a high frequency noise pattern.

4.5.7. Perceptual evaluation: design

- Nine randomly chosen samples from each different mode (three different modes of each subject) were played across subjects to five theatre experts. Three of the five had at least five years of training in performer's voice. The other two work fulltime with students in tertiary institutions, as well as working professionally as directors and performers. The latter two have more than ten years experience each.
- The evaluation questionnaire was provided to panellists with the definition of good voice quality as used in this study.[25] The questionnaire was set up according to a 5 point rating scale: 1 = very poor, 2 = poor, 3 = average, 4 = good, 5 = very good.
- Pre- and post-training recorded modes were played to panellists in random order. Samples of the different modes of Test Group V used in Section 4, and Test Group S used in this section were randomised. The randomised samples were still ordered according to the structured teaching progression – random samples of the Y-buzz mode were kept together, and so forth.
- Modes were played from a CD through a PC using a high quality amplifier and stereo speakers: Creative Inspire 2.1 2400[26] in order to maintain good sound quality.
- Evaluation sheets were filled in anonymously in the presence of the researcher.
- Cross tabulation tables will be used for descriptive statistics. These will provide an indication of the preference of the perception panel should it exist. These tables will also further serve as preparation for the inferential statistics.
- For triangulation more than one inferential statistics analysis were done on the data. 1) Paired t-test: H_o: $\mu_{diff(pre\text{-}post)} = 0$; H_A: $\mu_{diff(pre\text{-}post)} \neq 0$; $\alpha = 0.01$. H_o accepted when the p-value > α; H_o rejected in favour of H_A when the p-value < α. In this study H_o will imply that the scores given to the pre-training recordings minus the scores given to the post training recordings will equal null. Should this be true it will indicate that there was no change, according to the perceptual panel, in the sound quality of the post-training recordings. But should H_0 be rejected in favour of H_A, especially if the pre-scores minus

---

[25] As defined in 1.1, good voice quality (sustainable over periods of time and repeatable) for theatre may be "defined as a result of an optimal use of the vocal organ in order to establish the maximum possible acoustic output by minimal muscular effort" (Laukkanen, 1995:18).

[26] Technical specifications: Satellite: 4.5 watts. 12 ms power per channel (2 channels); subwoofer power: 12 watts RMS; Freq.response: 42 Hz-20kHz; SNR: >75dB; Dimensions: Satellites (LxWxH) - 8.7cm x 9.5 cm x 9.5 cm, Subwoofer – 21.1cm x 19.2 cm x 19.2 cm.

the post-scores provide a negative numerical, it will imply that the total of the post scores was higher than the total of the pre-scores. This will indicate that the utterances contained in the post-training recordings have improved in voice quality according to the perceptual panel. 2) Chi-square: H_0: 2 variables (phase and score) are independent; H_A: 2 variables (phase and score) are dependent. $\alpha = 0.01$. H_0 accepted when the p-value > α; H_0 rejected in favour of H_A when the p-value < α. In this study this H_0 will mean that the scores were randomly attributed to the phases (pre- and post- training recordings) and that there is no correlation between the scores given to the voice quality and whether it is a pre- or post-recording. H_A in this case means that the pre- recordings have lower scores attributed to them and the post-recordings have higher scores attributed to them. This will indicate that the training had a positive influence over the voice quality as perceived by the perception panel.

4.5.7.1. Data analysis and processing of the perceptual evaluation

The evaluation scores of the perception group were entered into Excel spread sheets and statistically processed through the use of SAS[27] (Statistical Analysis Software package, version 8). A paired t-test for independence was done to compare pre-score means with post-score means for all the sound modes combined. For this paired t-test, null hypothesis testing was done where the mean difference equals zero (H_0: $\mu_{\text{diff(pre-post)}} = 0$), against the alternative where the mean does not equal zero (H_A: $\mu_{\text{diff(pre-post)}} \neq 0$).

Conclusively the paired t-test for all the sound modes combined, has the null hypothesis (H_0) as:

No statistically significant difference between the pre- and post-training recordings (referred to as phase in the t-test) of Test Group S (H_0: $\mu_{\text{diff(pre-post)}} = 0$) as indicated by the scores allocated by the perception panellists;

and the alternative hypothesis (H_A) as:

A statistically significant difference between the pre- and post-training recordings (phase) of Test Group S (H_A: $\mu_{\text{diff(pre-post)}} \neq 0$) as indicated by the scores allocated by the perception panellists. For this study to indicate that the training did affect the voice quality positively, the post-training recordings' scores have to be higher than those of the pre-training recordings.

---

[27] Credit to the Department of Statistical Support of the Pretoria Technikon for assistance and guidance in this matter.

Frequency procedures were carried out and cross-tabulation tables were made for the different variables of all the different sound modes. These procedures and tables provide a descriptive profile of the weighting of the scores in relation to the pre- and post-training recordings as phase.

For the sake of triangulation chi-square tests were done on appropriate cross tabulations but due to expected cell frequencies being less than 5, regroupings were done. Sound modes 1,2 and 3 (being Y-buzz, +Y-buzz and Calls) were grouped as Lessac Tonal explorations, whilst sound modes 4 and 5 (being English Phrases and Call words as well as first language texts) were grouped as applications; seeing that these modes represent the carry-over of the explorations into functional speech. In the H_0 (null hypothesis) of the Chi-square test the two variables (phase and score in this study) are independent. The question to determine is whether this H_0 will be accepted or rejected for the alternative where the two variables are dependent.

Conclusively the chi-square test for the Lessac explorations as well as the applications, has the null hypothesis (H_0) as:

The scores (ratings 1-5) were assigned/attributed to the utterances of the pre- and post-training recordings (phase) of Test Group S at random, without any statistically significant relations between the ratings and the 2 phases (being pre- and post-recordings). This will be indicated as the p-value being > 0.01;

and the alternative hypothesis (H_A) as:

The scores (ratings 1-5) were assigned/attributed to the utterances of the pre- and post-training recordings (phase) of Test Group S according to a statistically relevant relationship between the ratings and the 2 phases (pre- and post-recordings). For this study to indicate an improvement of voice quality, the ratings allocated to the post-training recordings have to be higher than the ratings allocated for the pre-training recordings. This will be indicated as the p-value being < than 0.01.

The different language groups were sometimes too small, and as such cannot be used to provide any statistically significant indication. Descriptive statistical analyses were conducted for the different first language groups.

4.5.7.2. Perceptual evaluation results

As mentioned above, a t-test for independence[28] was done for all the sound modes combined, as part of this study. The rule used was that H_o (H_o: $\mu_{diff}=0$) will be rejected if the p-value < α (alpha) with alpha =0.01. It was decided on alpha as 1% to make the Type 1 error as small as possible.[29] The null-hypothesis (H_o) where the mean difference equals zero (H_o: $\mu_{diff}=0$) was thus rejected in favour of the alternative where the mean difference is significantly different from zero (H_A: $\mu_{diff} \neq 0$) seeing that the p-value was shown as <.0001 and this implies that the p-value is < α (see Table 1). Since the mean differences were negative,[30] it indicates that the pre-scores are significantly less than the post-scores ($\mu_{diff} < 0$). The perceptual evaluators, in total, thus preferred the voice quality of the post-recordings of the Test Group.

Table 1: T test for independence, Test Group S.

Analysis Variable: DIFF						
N	Mean	Std Dev	Minimum	Maximum	t Value	Pr> /t/
225	-1.2755	1.2191	-4.0000	2.0000	-15.69	<.0001

The cross tabulation tables for all the different sound modes (Y-buzz, +Y-buzz, Calls, English Phrases and Call words, as well as first language text readings) indicated a clear difference between the pre- and post-training recordings where the post-recordings[31] were reflective of the preferred voice quality. In Table 2 the cross tabulation table of the Y-buzz of Test Group S[32] is provided as an example.

Table 2: Group S, Y-buzz: Table of phase by score.

	Score					
Phase	1	2	3	4	5	Total
A	16	12	12	3	2	45
Z	1	8	11	19	6	45
Total	17	20	23	22	8	90

---

[28] Paired data

[29] Alpha (α) = P(Type 1 error); Alpha = P(reject H_0 when H_0 true)

[30] For example observe the mean in Table 1.

[31] In the cross tabulation tables the pre-training recording is indicated as A and the post-training recording as Z.

[32] In frequency of score per phase.

Table 2 indicates that the perceptual evaluators rated the Y-buzz modes of the pre-training recording (A) primarily as very poor –score 1 (16/45 divided by 100 = 35.56%), poor - score 2 (26.67%) and average – score 3 (26.67%). The weighting of the post-training recording (Z) perceptual evaluation leans strongly towards average - score 3 (24.44%), good – score 4 (42.22%) and very good – score 5 (13.33%). This pattern is basically followed in all four the other sound mode groups[33] with the score weighting moving from being centred on score 1 and 2 to score 4 and 5.[34]

As mentioned under "Data analysis and processing of the perceptual evaluation," chi-square tests were done on appropriate cross tabulations. For chi-square the H_o is that the two variables (phase and score) are independent. This will be tested against the alternative (H_A), which indicates the two variables as dependent. Regrouping had to be done due to expected cell frequencies being less than five. Sound modes 1, 2 and 3 (Y-buzz, +Y-buzz and Calls) were grouped as Lessac Tonal explorations, whilst sound modes 4 and 5 (English Phrases and Call words as well as first language texts) were grouped as applications because these sound modes represent the carry-over of the explorations into functional speech. The score possibilities were also combined so that scores 1 and 2 on the perception panel questionnaire are in the chi-square test regrouped as score 1; score 3 on the perception panel questionnaire remains 3 in the chi-square test and scores 4 and 5 on the perception panel questionnaire are regrouped as score 5 in the chi-square tests.

In the chi-square the H_o (null hypothesis) has it that the two variables (phase and score in this study) are independent. This will be accepted should the p-value be $> \alpha$ and rejected should the p-value be $< \alpha$. Alpha (α) equals 0.01[35] in order to make the type 1 error as small as possible. The alternative (H_A) is that the two variables (phase and score) are dependent. Should the null hypothesis be rejected it will thus be an indication that a relationship exists

---

[33] The +Y-buzz, Calls, English phrases and Call words as well as first language texts all reflected the same pattern and are in the possession of the researcher.

[34] As previously mentioned, the cross tabulation tables are descriptive statistics and don't have a null-hypothesis. This profile thus does not feed directly into the acceptance or rejection of an H_O, but, seeing that these tables are used in preparation for the chi-square test, it is already obvious that a positive relationship exists between score and phase, seeing that the weighting of the scores for the post-recordings is higher than the weighting of the scores for the pre-recordings.

[35] Please note that the numerical contributed to Alpha is only applicable for the perceptual evaluation. Should the p-value be $< \alpha$ in this case, it would be statistically very significant.

between the rating received by the perception panel and the phase (pre- or post-training recording). This relationship will be proven to be positive or negative pending on the outcome of the cross tabulation and the chi-square test.

Table 3a below provides the cross tabulation (of phase by score) in preparation for the chi-square test. Table 3b provides the chi-square test results of the Lessac explorations – sound modes 1, 2 and 3.[36]

Table 3a: Cross tabulation in preparation for chi-square of the Lessac explorations.

Frequency Expected Percent Row Pct Col Pct	Table of phase by nscore			
phase	nscore			Total
	1	3	5	
A	85 50.5 31.48 62.96 84.16	37 39.5 13.70 27.41 46.84	13 45 4.81 9.63 14.44	135 50.00
Z	16 50.5 5.93 11.85 15.84	42 39.5 15.56 31.11 53.16	77 45 28.52 57.04 85.56	135 50.00
Total	101 37.41	79 29.26	90 33.33	270 100.00

Table 3b: The chi-square results of the Lessac explorations.

Statistic	DF	Value	Prob
Chi-Square	2	92.9662	<.0001

---

[36] Y-buzz, +Y-buzz and Calls.

Table 4a below provides the cross tabulation (of phase by score) in preparation for the chi-square test. Table 4b provides the chi-square test results of the applications – sound modes 4 and 5.[37]

Table 4a. Cross tabulation in preparation for chi-square of the applications.

Frequency Expected Percent Row Pct Col Pct

Table of phase by nscore				
phase	nscore			Total
	1	3	5	
A	52 29.5 28.89 57.78 88.14	26 29 14.44 28.89 44.83	12 31.5 6.67 13.33 19.05	90 50.00
Z	7 29.5 3.89 7.78 11.86	32 29 17.78 35.56 55.17	51 31.5 28.33 56.67 80.95	90 50.00
Total	59 32.78	58 32.22	63 35.00	180 100.00

Table 4b. The chi-square results of the applications.

Statistic	DF	Value	Prob
Chi-Square	2	59.0856	<.0001

The tables indicate a preference of the post-training recordings. The Lessac explorations pre-training recordings only have 9.63% of their total amount for scores allocated as score 5, but the post-training recordings have 57% of their total scores allocated to score 5. Similarly, the application modes only have 13.33% of the scores being 5 for the pre-training recordings but 56.67% of the scores being 5 in the post-training recordings. Since the p-

---

[37] English Phrases and Call words and First language texts.

value is, in both the Lessac explorations[38] (Y-buzz, +Y-buzz and Calls – see Figure 3b), and the applications[39] (English Phrases and Calls words and first language – see Figure 4b), less than .0001. The p-value of both these cases is smaller than α (α=0.01). The null hypothesis[40] is rejected. It is therefore evident that phase does effect score. This indicates that H_0 where phase and score are independent is rejected in favour of the alternative where phase and score are dependent. The perception panellists rated the post-recordings in both cases (explorations and applications) significantly higher than the pre-recordings, indicating that the post-training recordings contained the preferred sounds. A dependant positive relationship thus exists between the phase and the score. This pattern indicates that the perception panellists are of the opinion that the outcome of the training is in line with the definition provided to them of what good voice quality is (see 1.1 p.3 for details).

Although, as already stated, one has to be careful to make final conclusions about the various first language groups,[41] it can be used as in indication of a trend to be explored in further research. As such the cross tabulation (of phase by score) for the Calls of the three language groups[42] in Test Group S[43] are provided in Tables 5.

Table 5a: Group S, Calls, African languages: Table of phase by score.

	Score					
Phase	**1**	**2**	**3**	**4**	**5**	**Total**
A	4	5	5	1	0	15
Z	0	0	5	8	2	15
Total	4	5	10	9	2	30

---

[38] See Figure 3b.
[39] See Figure 4b.
[40] The null hypothesis is that the two variables are independent.
[41] This is due to the small amount of participants in this study
[42] Afrikaans is taken as one language group, English as another and the 5 participants having an African language as first language is combined into a third group namely "African". The researcher admits that this is not an ideal situation. Further research into the use of the Lessac Approach for various African languages is planned.
[43] The cross tabulation in indicated in frequency of score by phase.

Table 5b: Group S, Calls, Afrikaans: Table of phase by score.

	Score					
Phase	**1**	**2**	**3**	**4**	**5**	**Total**
A	8	11	6	0	0	25
Z	0	0	11	11	3	25
Total	8	11	17	11	3	50

Table 5c: Group S, Calls, English: Table of phase by score.

	Score					
Phase	1	**2**	3	4	**5**	**Total**
A	1	4	0	0	0	5
Z	0	0	0	0	5	5
Total	1	4	0	0	0	10

The cross tabulation tables for the Calls of the three language groups (Tables 5 a, b, and c) illustrate a reflective profile of the actual ratings being assigned by the perception panel to the pre- (A) and post- (Z) training recordings. In all these profiled examples the post-training recordings were preferred, with a higher rating (of score) than the pre-training recordings.[44] Again, this is also the case with all four the other sound modes for all three the language groups.[45]

Cross tabulation tables were made for the scores of the three pre-training recordings, as well as three post-training recordings for each participant in Test Group S that were played to the perception panel. Table 6 depicts this cross tabulation for one of the participants (called AA in the table), selected at random.

---

[44] When comparing score 4 and 5 of A (Pre-training) and Z (post-training) the preference for the post-training recordings are eminent.

[45] Specific cross tabulations in possession of the researcher.

Table 6: Cross Tabulation Table: Group S, participant AA, all sound modes.

Phase	Score 1	Score 2	Score 3	Score 4	Score 5	Row Totals
A	1	11	2	1	0	15
Z	0	0	3	11	1	15
Total	1	11	5	12	1	30

Seeing that this table is typical, it functions as an example to indicate that each participant's voice has, according to the perception panel, improved during the training process, as the post-training recordings ratings centre around the higher scores.[46]

Although an indication of the profiles of the perception panellists has been provided before, it will contribute to the effectiveness of this study to reflect on their reliability as a group. Seeing that the various samples taken from the sound modes of Test Group S were played randomly with various samples of the sound modes of Test Group V, statistical analysis on the inter reliability of the perception panel was done according to the information gathered from both test groups. As such, an in-depth discussion about this will take place in the comparative section of this study (see Chapter 5). Reflecting on this forthcoming discussion, it is noted that all the perception panellists (raters) separately indicated a significant improvement in the post-training recordings versus the pre-training recordings with the results of the paired t-tests done for each rater indicating a p-value of <.0001. It is thus very clear that the H_o ($\mu_{diff} = 0$) is rejected in favour of the HA ($\mu_{diff} \neq 0$) with $\mu_{diff} < 0$ by each rater.

4.5.8. Acoustic investigation

4.5.8.1. Design, data analysis and processing of the acoustic investigation

Pre- and post-training recordings were transposed from DAT into CSL as NSP files. This was put onto rewritable CD's using a Hewlett Packard – Sonic Foundy –ACID (1998) programme. Sound modes were drawn into Multi-Speech programme. A Long Term Average Spectrum

---

[46] See footnote 40.

was done for each of the five sound modes[47] of each participant, as well as a group average LTAS for each sound mode. Settings used: FFT size: 512 (thus modes/frame: 257); pre-emphasis: 0.000; Window weighting: Hanning; Smoothing level: None.

The numerical data of the acoustical profiles of pre- and post-training recordings was then entered into MS Excel.

- F0 dB pre-training was subtracted from F0 dB post-training.
- The highest amplitude within windows of 100 Hz pre-recording was subtracted from the highest amplitude within the same windows of 100 Hz post-recording. Thus the highest amplitude reading between 1500-1600 Hz pre-recording, was deducted from the highest amplitude reading between 1500-1600 Hz post-recording and so on.
- The difference between F0 dB post-recording and F0 dB pre-recording was then subtracted from all these answers leaving the amount of amplitude increase over and above the difference of F0. The differences were thus calculated in relation to F0 dB difference.[48]

This data was statistically processed by the use of SAS (Statistical Analysis Software programme, version 8). The null hypothesis testing whether the group mean, or median, of the difference between post- and pre- is equal, was tested against the alternative that the difference is positive ($H_o: \mu_{post-pre} = 0$; $H_A: \mu_{post-pre} > 0$). The first step was to determine whether a paired t-test (mean) or a Wilcoxon test (median) should be applied to this data, depending whether the data was normally distributed or not. Shapiro Wilk was used to test for this assumption of normality. Data was accepted as normally distributed when the p-value was >0.05. If the data was distributed normally, the t-test was used with $H_o: \mu=0$; $H_a: \mu>0$ and $\alpha=0.05$. H_o was accepted when the p-value (divided by 2) ≥ 0.05. H_o was rejected in favour of H_A when the p-value (divided by 2) ≤ 0.05. Otherwise the Wilcoxon distribution free method was used.[49]

Conclusively only the H_o and H_A of the t-test for the acoustic analyses, in general, are provided seeing that all the statistically significant differences occurred when the data was

---

[47] As previously mentioned these modes are the 5 utterances: Y-buzz, +Y-buzz, Calls, English Phrases and Call words and first language texts.

[48] See Section 2 for more details.

[49] $H_o: \eta = 0$; $H_A: \eta > 0$; $\alpha = 0.05$; When p-value (divided by 2) of S ≥ 0.05 accepted H_o. Reject H_o in favour of H_A when p-value (divided by 2) of S ≤ 0.05 and S value is positive.

normally distributed and the t-test was thus used. The t-test was used for each of the 5 sound modes respectively and within these sound modes separately for each 100Hz window. However, to provide the t-test hypotheses for each 100 Hz window of each of the 5 sound modes will result in repetition and duplication.[50]

The null hypothesis (H_o) has that:

No statistically significant difference between the acoustic profiles of the pre- and post-training recordings (referred to as phase in the t-test) of Test Group S (H_O: $\mu_{diff(post-pre)} = 0$) exists for each specific 100Hz window for each one of the sound modes;

and the alternative hypothesis (H_A) has that:

A statistically significant positive difference between the acoustic profile of the pre- and post-training recordings (phase) of Test Group S (H_A: $\mu_{diff(post-pre)} > 0$) exists for each specific 100Hz window of each one of the sound modes.

For this study to indicate that the training did affect the voice quality positively, the post-training recordings' mean has to be statistically significantly higher over certain frequencies than the mean of the pre-training recordings. These frequency clusters will be referred to in the discussion and interpretation of the acoustical profile of the post-training recording modes.

Further qualitative observations about the frequency of F0 of the +Y-buzz, Calls, English Phrases and Call words, as well as first language text readings were done.[51] Observations were also made in regards to F1 and F2 of the Y-buzz when applicable. The change in the energy profile, over the 2500-5000 Hz spectrum in all the modes, has been commented on as necessary.

4.5.8.2. Acoustic investigation results

The group average LTAS of the pre-training and post-training recordings of each sound mode of Test Group S is graphically presented:

---

[50] This data is in possession of the researcher and can be obtained from her.

[51] Seeing that the Y-buzz mode had a given pitch, the F0 frequency differences will only be commented on.

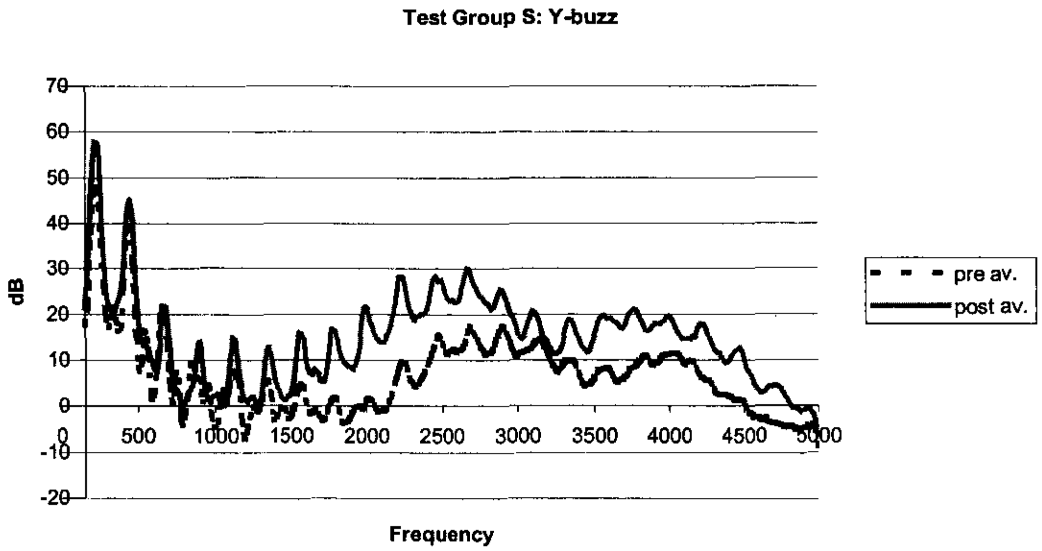


Figure 1. The average LTAS of the pre- and post-training of the Y-buzz of Test Group S.

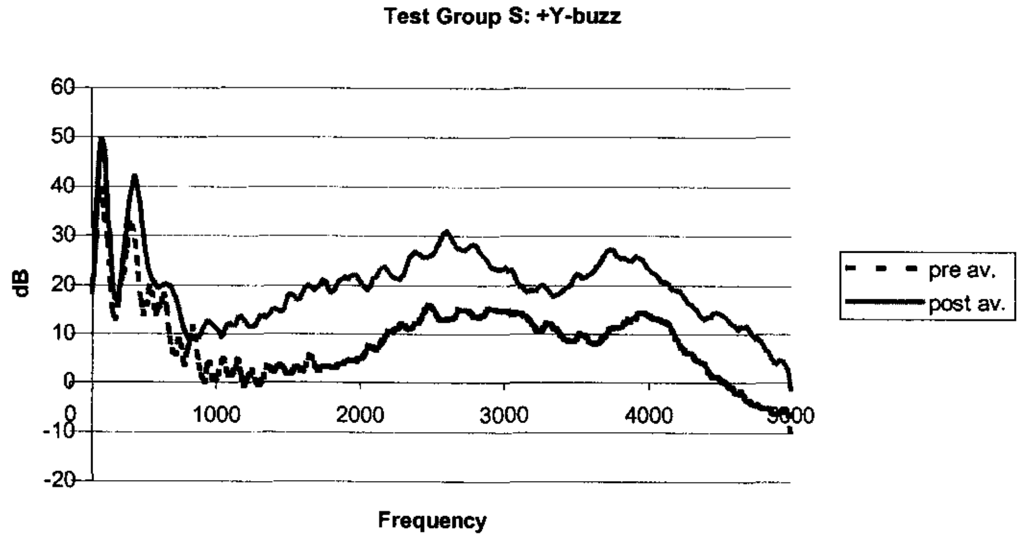


Figure 2: The average LTAS of the pre- and post-training of the +Y-buzz of Test Group S.

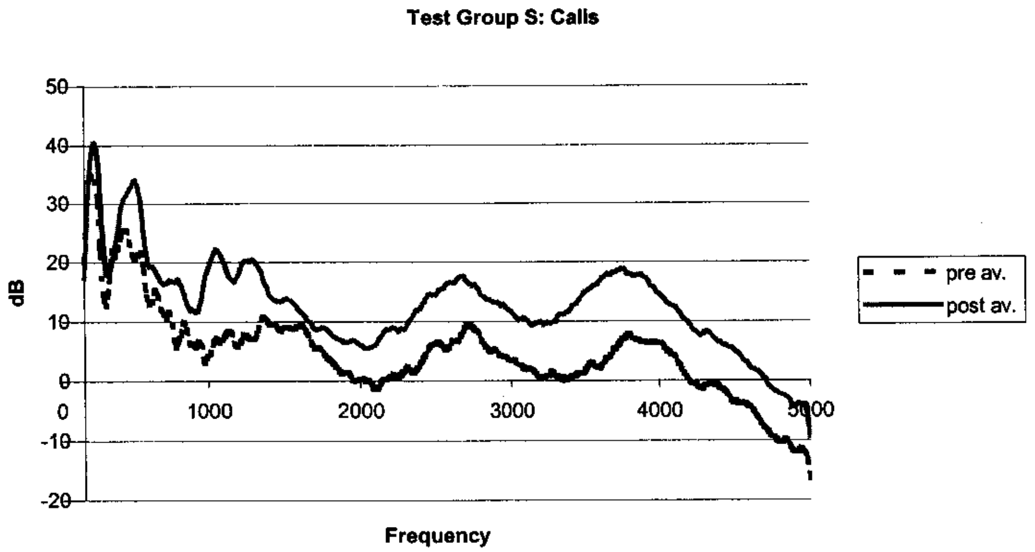


Figure 3: The average LTAS of the pre- and post-training of the Calls of Test Group S.

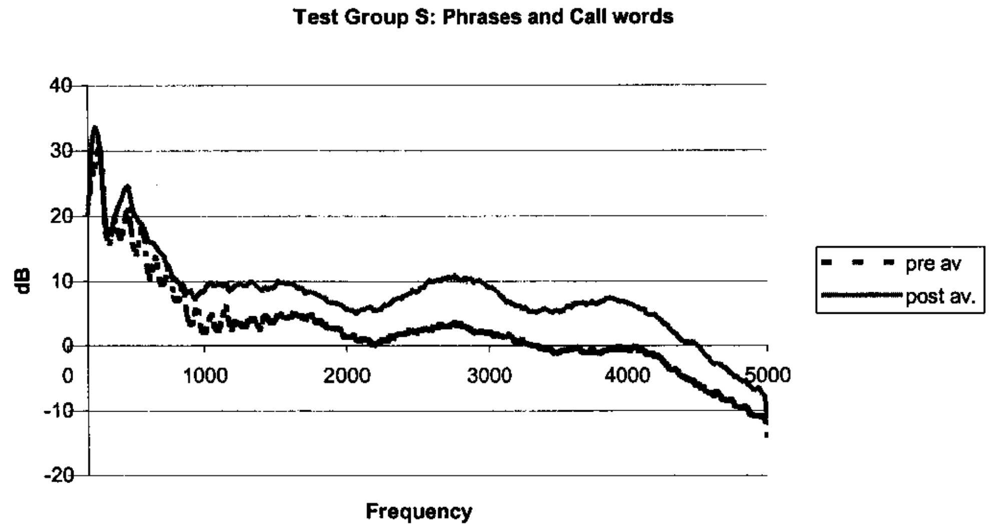


Figure 4: The average LTAS of the pre- and post-training of the English phrases and Call words of Test Group S.

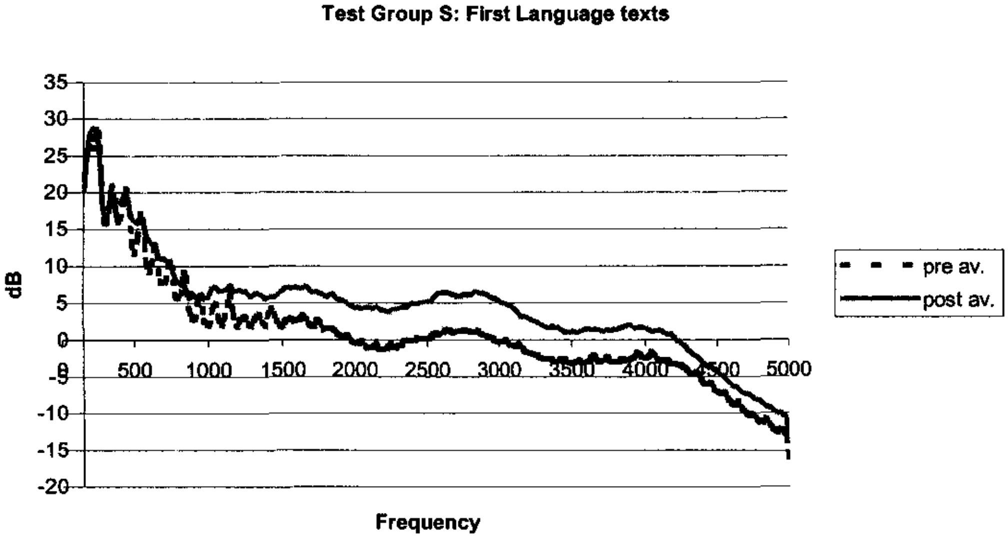


Figure 5: The average LTAS of the pre- and post-training of the first language texts of Test Group S.

As mentioned in the "Data analysis and processing of the acoustic investigation" the numerical results of the LTAS of each of these modes were entered into Excel. As in Section 2, an approximation of the numerical results was done, where the highest dB within a window of each 100Hz was chosen. Following this approximation, the amplitude of each 100 Hz window of the first recording was subtracted from the amplitude of each 100 Hz window of the second recording,[52] providing the amplitude difference between each frequency window of the pre- training and post-training recordings. In an attempt to standardise these differences and interpret them in relation to the "loudness"[53] difference of F0, the difference between amplitude of the F0 of the post-training recording, and amplitude of the F0 of the pre-training recording, was subtracted from the differences of each of the 100 Hz windows.

Table 7 posits an extract as example of this process that was followed for each sound mode of each participant.[54] The participant referred to in table 7 is referred to as AA and was randomly chosen.

---

[52] Post (Z) – Pre (A)

[53] The subjective term referring to energy and amplitude.

[54] See Section 2:4.4.6 in this chapter for a step-by-step demonstration of this process.

Table 7: An extract example of the processing of the data. Test Group S, participant AA, +Y-buzz.

Freq.	Pre	Post	Diff.	F0 diff	Diff in relation to F0
F0	43.02	56.51	13.49	13.49	13.49
200-300	23.49	56.51	33.02	13.49	19.53
300-400	43.04	22.19	-20.85	13.49	-34.34
400-500	18.03	54.84	36.81	13.49	23.32
500-600	32.51	29.08	-3.43	13.49	-16.92

The "difference in relation to F0" (including F0) data of the separate participants' sound modes were then grouped according to mode type (Y-buzz, +Y-buzz etc) in order to process the data statistically. A mean and median were assigned to each 100 Hz window of each of these groups. The Shapiro Wilk test was used to test for normality. Data were normally distributed when the p-value was > 0.05. The data that was statistically significant was all normally distributed. For these the mean is used for the null hypothesis testing. In this particular case the null hypothesis is $\mu_{post-pre} = 0$. The mean difference (for post - pre - F0 dB difference) was tested against the alternate ($\mu_{post-pre} \neq 0$). H_0 was rejected in favour of H_A as $\mu_{post-pre} \neq 0$. Seeing that SAS automatically provides two-sided testing and this study is interested in one-sided testing, the p-value of the mean has to be divided by two, before comparing to Alpha.[55] Alpha in this case is 0.05. When the p-value divided by 2[56] is smaller than Alpha the increase will be statistically significant. The t-test outcome is graphically presented in the different bar graphs below (Figures 6,7 8, 9 and 10) where the mean/median differences are indicated (for the highest amplitude difference in each 100Hz window) for each sound on the following graphs. In each of these graphs, it is indicated with a dot where the mean dB in each 100Hz window[57] is statistically significant on a 5% level of significance. The p-values (when divided by 2) of the means of those 100Hz windows were thus smaller

---

[55] The p-value divided by 2 < α; p-value divided by 2 < 0.05.

[56] SAS automatically provides a two-sided p-value. This study is only focussing on the increase of the mean and thus one-sided testing will be used. For one-sided testing, the p-value must be divided by 2 before comparing it with Alpha ($\alpha = 0.05$).

[57] After standardisation, e.g.: Highest dB between 300-400Hz Post minus highest dB 300-400Hz Pre minus difference dB of post minus pre F0.

than Alpha (0.05), which is statistically significant (p-value < 0.05). The Shapiro Wilk, mean, t-test, and p-value of each of the 100Hz windows that show a statistically significant increase, of each mode, are provided in the addendum to this study.

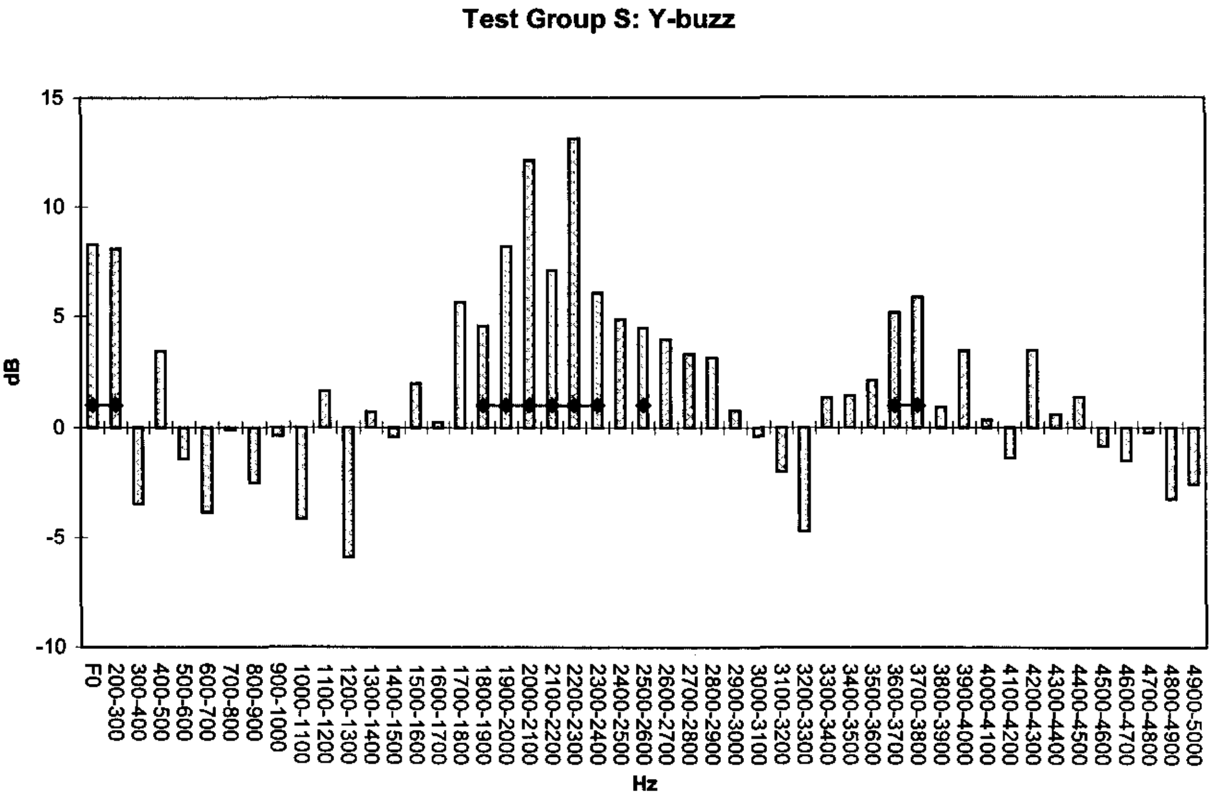


Figure 6: Bar graph of mean/median differences of the Y-buzz. Statistically significant differences (of mean) are indicated with dots. Statistically significant clusters are indicated by lines connecting the dots.

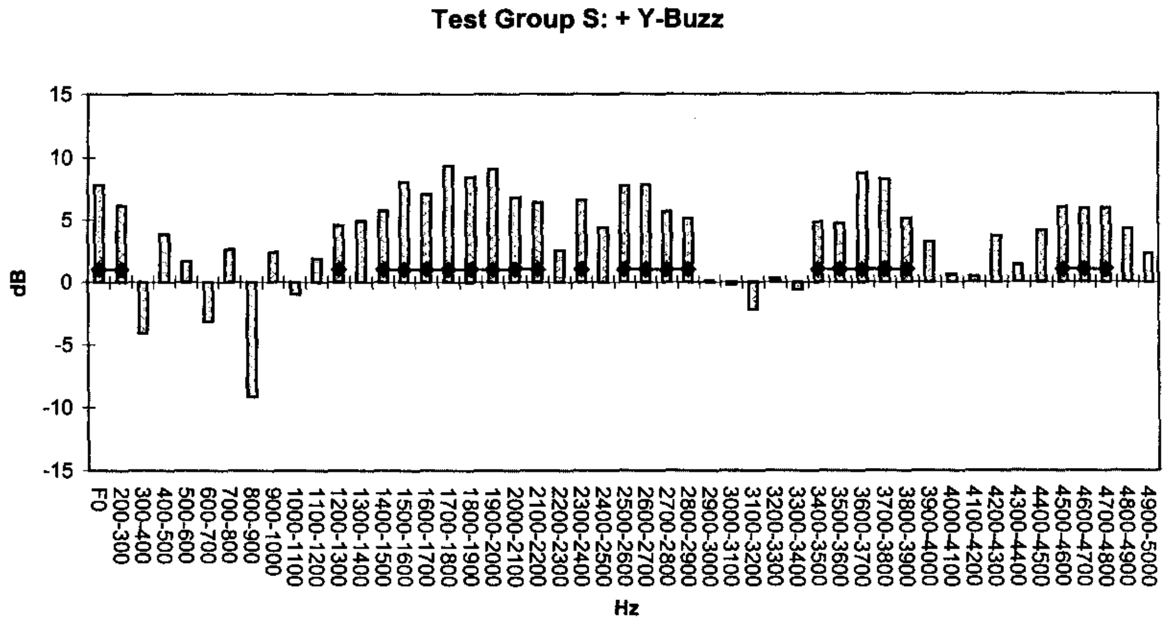


Figure 7: Bar graph of mean/median differences of the + Y-buzz. Statistically significant differences (of mean) are indicated with dots. Statistically significant clusters are indicated by lines connecting the dots.

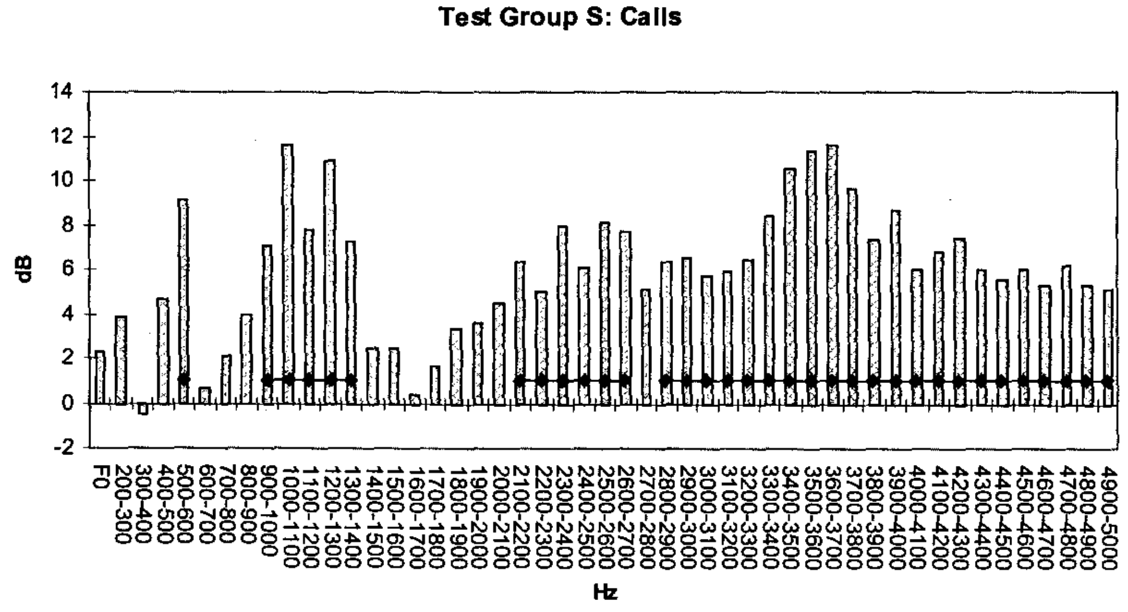


Figure 8: Bar graph of mean/median differences of the Calls. Statistically significant differences are indicated with dots. Statistically significant clusters are indicated by lines connecting the dots.

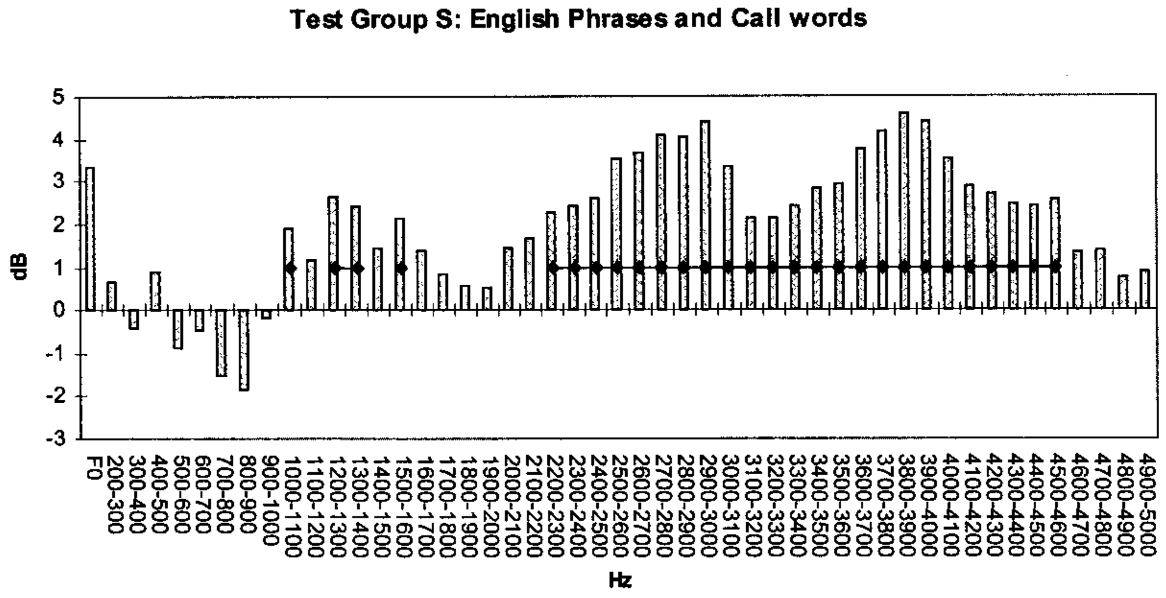


Figure 9: Bar graph of mean/median differences of the English phrases and Call words. Statistically significant differences are indicated with dots. Statistically significant clusters are indicated by lines connecting the dots.

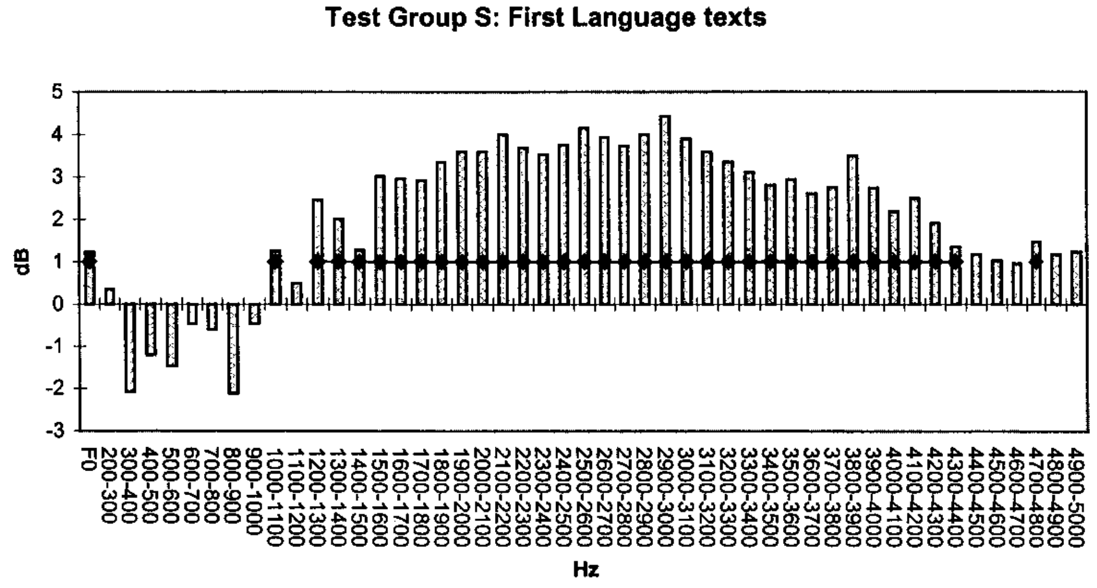


Figure 10: Bar graph of mean/median differences of the First language texts. Statistically significant differences (of mean) are indicated with dots. Statistically significant clusters are indicated by lines connecting the dots.

Results based on the comparative profiles of the pre- and post-training recordings, as depicted in the LTAS (Figures 1-5) and the bar graphs[58] (Figure 6-10) of each mode group follows.

Test Group S: Y-buzz:

In order to demonstrate this process of accumulating the data, the LTAS and Bar Graph of the Test Group S: Y-buzz are placed next to each other in Figure 11 below.[59]

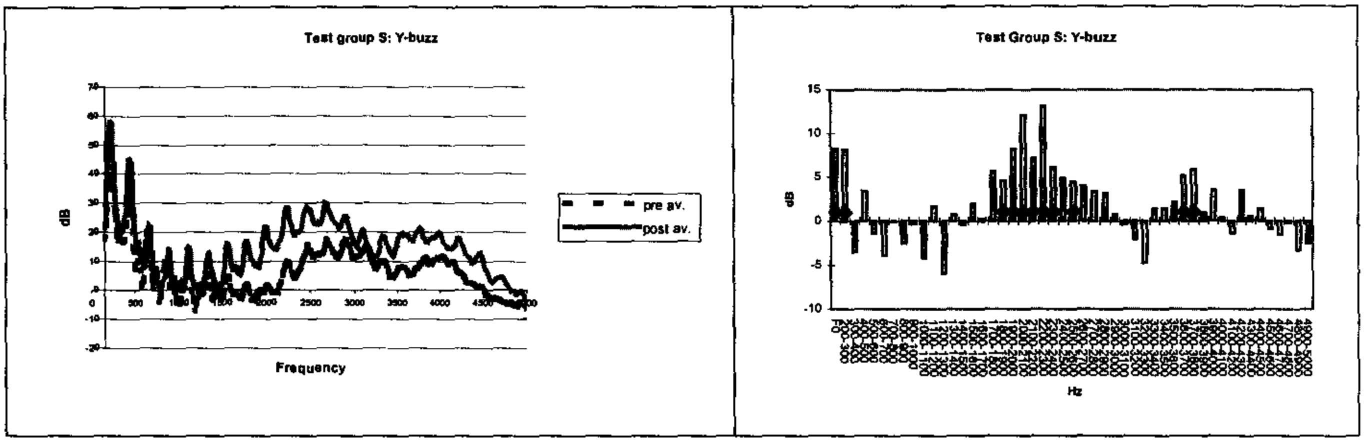


Figure 11. The LTAS and Bar Graph of the Test Group S: Y-buzz.

Post-training recordings reflect:

- a statistically significant increase in the F0 dB. Also observed is an increase[60] in the 200-300 Hz region. For most of the participants this is their F0 range and as such it will be treated as a F0 increase.
- an increase, in relation to the F0dB difference, as indicated in the 400-500Hz window.
- a more definite (clearer) valley between 500-1000 Hz and possibly up to 1700Hz.
- a clear, statistically significant,[61] energy increase between 1800-2400Hz, above the F0 dB difference. The amplitude peak of this cluster moved down in frequency, in comparison to the peak observed in the LTAS of the pre-training recording. This is

---

[58] The mean/median differences of each 100 Hz window of each sound mode of Test Group S in relation to the F0 dB difference.

[59] Previously these Figures appeared in Section 3 as Figure 1 and Figure 6.

[60] Please note that this increase is not statistically significant and can thus not be used for generalization but is definitely recognizable. When an increase is statistically significant it will be indicated as such.

[61] As indicated with the dots on the bar graphs.

possibly due to the use of the forward facial orientation, which has as effect a longer resonator.

- a dB increase between 3600-3800Hz where the cluster peak again moved down in frequency.

The smallest mean difference in the post training recording[62] that is statistically significant, lies in the 1800-1900Hz window where the mean is 4.59, the student's t is 1.97, and the p-value is 0.0687.[63]

The biggest mean difference in the post-training recording that is statistically significant lies in the 2200-2300Hz window where the mean is13.09, the student's t is 3.15 and the p-value is 0.0070.

Test Group S: +Y-buzz:

Post-training recordings reflect:[64]

- a statistically significant increase in the F0 dB. An energy increase is observed in the 200-300 Hz region. As mentioned previously this is the F0 window for most of the participants and will be treated as such;
- in relation to the F0dB difference, a statistically significant dB increase between 1200-2100 Hz (except for the 1300-1400Hz window that indicates an increase but not statistically significant);
- in relation to the F0dB difference, a statistically significant dB increase between 2300-2900 Hz (except for the 2400-2500Hz window that indicates an increase but not statistically significant);
- a clear cluster, with a statistically significant dB increase, between 2300-2900 Hz. The peak frequency is lower than in the pre-training LTAS.
- a clear valley between 2900-3400 Hz;
- a clear cluster, with a statistically significant dB increase, between 3400-3900 Hz. The peak frequency is lower than in the pre-training LTAS;
- a statistically significant dB increase between 4500-4800Hz.

---

[62] This is above the difference of the dB of F0 – thus more than the dB increase observed in the F0 post training recording.

[63] As explained in footnote 52, the p-value needs to be divided by 2 before being compared to Alpha ($\alpha = 0.05$).

[64] Compare Figures 2 and 7.

The smallest mean difference in the post-training recording that is statistically significant lies in the 1200-1300Hz window where the mean is 4.55, the student's t is 2.01, and the p-value is 0.0641.[65] The biggest mean difference in the post-training recording that is statistically significant lies in the 2000-2100 Hz window where the mean is 6.77, the student's t is 2.46 and the p-value is 0.0273.

Test Group S: Calls:

Post-training recordings reflect:[66]

- a slight increase in the F0 dB;
- a energy increase between 400-600 Hz, with the dB increase between 500-600Hz being statistically significant;
- an increase of energy between 1900-5000Hz, although not all is statistically significant;
- a statistically significant increase of energy between 900-1400Hz with the peak cluster lower than in the pre-training recording;
- between 2100-2700Hz a statistically significant dB increase. The cluster peak remains around the same frequency as in that of the pre- training recording;
- that from 2800-5000Hz all energy increases are statistically significant. The highest energy increase happening between 3500-3700Hz.

The LTAS of the group of Call modes indicates a cluster between 3500-4000Hz. This is a slight move downwards towards a lower peak frequency for this specific cluster. The smallest mean difference in the post-training recording that is statistically significant lies in the 2200-2300 Hz window where the mean is 5.06, the student's t is 2.85, and the p-value is 0.0138.The biggest mean difference in the post-training recording that is statistically significant lies in the 1000-1100Hz and 3600-3700 Hz windows where the mean for the 1000-1100Hz is 11.59, the student's t is 5.62 and the p-value <.0001. For the 3600-3700Hz window the mean is 11.59, the student's t is 4.007, and the p-value is 0.0015.

---

[65] Divided by 2 before comparing with Alpha ($\alpha = 0.05$) – see footnote 52.

[66] Compare Figures 3 and 8.

Test Group S: English Phrases and Call words:

Post-training recordings reflect:[67]

- a slight increase in the F0 dB. An energy increase is again observed in the 200-300 Hz window. This will be treated as mentioned before;
- basically the same visual pattern of the LTAS of the pre-training recording;
- the cluster between 2500-3200Hz as more pronounced. This also correlates with a statistically significant increase in dB in this area;
- the cluster between 3500-4200Hz as more pronounced. Again, this correlates with a statistically significant increase in dB in this area. The peak frequency of this cluster is lower in the post-recording, possibly indicating the lengthening of the resonator due to the "forward facial orientation" and a relaxed use of the larynx.

The smallest mean difference in the post-training recording that is statistically significant lies in the 1000-1100Hz window where the mean is 1.89, the student's t is 2.12, and the p value is 0.0550.[68] The biggest mean difference in the post- training recording that is statistically significant lies in the 3800-3900Hz window where the mean is 4.60, the student's t is 4.32, and the p-value is 0.0010.

Test Group S: First language texts:

Post-training recordings reflect:[69]

- a very small increase in the F0 dB;
- in relation to F0dB, a decrease in energy between 300-1000Hz;
- a statistically significant energy increase between 1000-1100 Hz, 1200-4300Hz and 4700-4800Hz;
- a more pronounced cluster between 2500-3200 Hz with the cluster peak at 2800-3000 Hz.

The smallest mean difference in the post-training recording that is statistically significant lies in the 1000-1100Hz and 1400-1500Hz windows where the mean for the 1000-1100 Hz

---

[67] Compare Figures 4 and 9.

[68] Divided by 2 before comparing with Alpha.

[69] Compare Figures 5 and 10.

window is 1.25, the student's t is 1.96, and the p-value is 0.0701.[70] For the 1400-1500HZ window, the mean is 1.27, the student's t is 2.18 and the p-value is 0.0466. The biggest mean difference in the post-training recording that is statistically significant lies in the 2900-3000Hz window where the mean is 4.60, the student's t is 4.433, and the p-value is 0.0002.

As previously mentioned, the groups of first language speakers were in the cases of English and African languages too small to do any statistical analysis. The LTAS was used to determine whether the first language groups reflect the patterns observed in the combined language groups. The same basic pattern primarily occurs in all different language groups.

For argumentation, the graphic LTAS presentation of the Call mode of the first language African speakers (Figure 12), as well as the Afrikaans first language text mode (Figure 13) will be discussed.

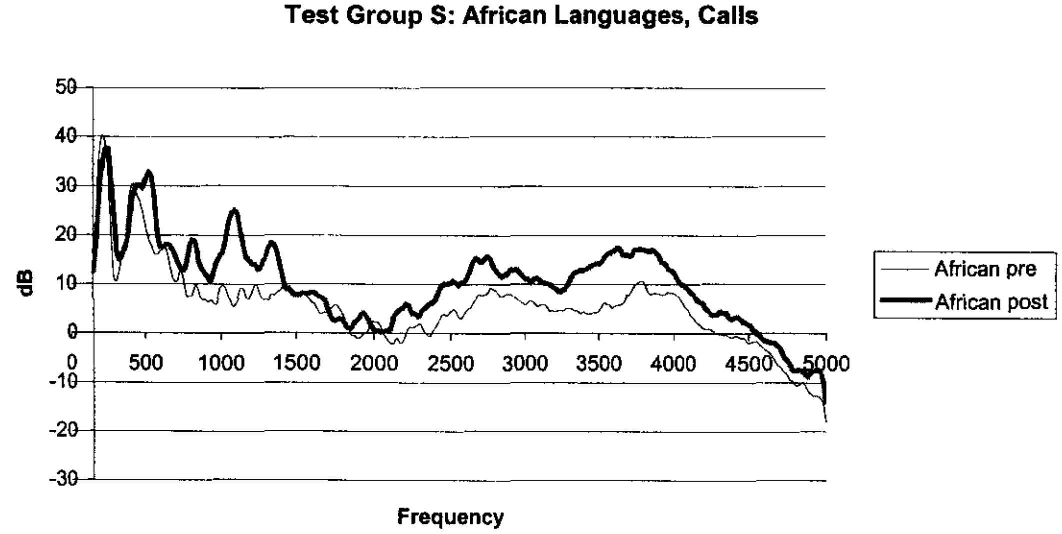


Figure 12. The average LTAS of the pre- and post recordings of the Calls of Test Group S, First language African.

---

[70] Divided by 2 before comparing with Alpha.

Test Group S: Calls, First language: African:

The average LTAS of pre- and post-training recordings reflect that the post-training recording has:

- a slightly higher frequency of F0;
- a slightly lower F0 dB;
- a change of the visual profile, specifically up to 2000 Hz. This may be due to a change in the diphthong pronunciation. The diphthong might have been shaped more clearly – thus the two vowels involved are both pronounced "more overtly". This is in line with the findings of Van Rooy and Van Huyssteen (2000) in relation to "Black South African English." This may also be due to a possible difference in the length of the two vowels being different in the post-training recordings. This, however, reflects a change in the diphthong characteristics and should be investigated in a research project where the length and shaping of the vowels can be controlled;
- a clear cluster between 2597-3046 Hz;
- a clear cluster between 3339-4003 Hz with its peak frequency lower than the peak frequency observed in the pre-training recording.

When comparing the LTAS of the Calls of Test Group S (Figure 3) and the LTAS of the sub-group first language African group only (Figure 12), strong resemblances exist:

- an energy increase between 400-600 Hz, as well as between 900-1400 Hz;
- a cluster between 2100-3000 Hz (although the first language African cluster is wider in frequency range);
- a cluster from 3500 Hz to 4000 Hz with its peak frequency slightly lower than the peak frequency of the cluster observed in the pre-training LTAS.

Test Group S: First Language texts, First language: Afrikaans.

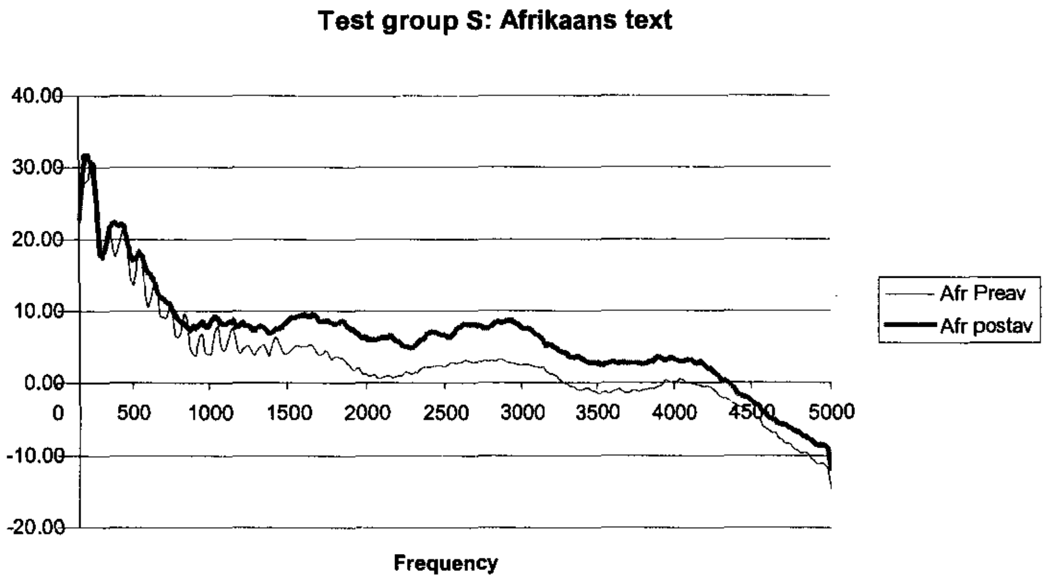


Figure 13. The average LTAS of the pre- and post-recordings of the first language texts, Test Group S, first language Afrikaans.

The average LTAS of pre- and post-training recordings reflect that the post-training recording has:

- the frequency of F0 slightly lower than that of the pre-recording LTAS;
- a very small increase of F0 dB;
- a cluster at 371-449 Hz;
- a clear increase in energy, in relation to F0 dB difference, over the spectrum of 1562-4062 Hz;
- more prominent clusters are observed between 1464 –1972 Hz as well as 2539-3222 Hz.

Although it has to be kept in mind that the Afrikaans first language speakers are 53% of the total number of participants of Test Group S, it may be valuable to reflect on a comparison between the LTAS of the first language texts of Test Group S, (Figure 5) and the LTAS of the first language texts, sub-group first language, first language Afrikaans speakers only (Figure 13). Strong resemblances exist:

- basically the same visual profile of the LTAS;

- both have a slightly lower F0 Hz;
- both cluster around more or less 300-500 Hz, 1400-2000 Hz and 2500-3200 Hz;

These two examples reflect the consistent pattern of the various languages in all five the different modes.

Further qualitative observations were made from individual LTAS' of each participant for each sound mode. These observations were done in relation to the F0 dB difference unless otherwise commented. These observations are not exact in nature but act as a mere indication of what tendencies are observed in the post-training recordings.

Y-buzz:

The F0 of the Y-buzz was given to the participants in both the pre-training and post-training recordings. The majority of post- training recordings reflect the same F0 frequency. 67% of the participants increase the F0dB whilst 33% used the same amount of energy and thus the dB stayed the same. 87% of the LTAS post-training recordings (Y-buzz) indicated that the frequency for F1 stayed the same, whilst 13% was lower in frequency. 53% of the dB of F1 post-training recording stayed, in relation to F0 dB, the same, whilst 7% dropped in energy and 40% increased in energy. When reflecting on the frequency of F2 of the various individuals' Y-buzz, 67% stayed the same, 13% was higher and 20% lower in relation to the F0 dB difference. Energy-wise 43% of the post-training recordings indicate an increased dB, whilst 57% remained constant.

+Y-buzz:

Although the F0 of the +Y-buzz was not given to the individuals, 47% used the same F0 in the post-training recording, 13% used a lower frequency and 33% a higher frequency. The dB of F0 were either the same (33%) or higher with 67%.

40% of the participants displayed a lower F1 frequency on the post-training recordings and 60% had the same F1 frequency. In 47% of the individual comparisons, the dB of the F1 increased, 20% had the same energy and 33% had a decrease of energy. Energy increase over the higher spectrum 2500-5000 Hz was observed as presented below in Figure 14.

These energy increases indicated are tendencies referred to in general, and the specific dB increase is not addressed, as it is taken care of in the LTAS and bar graph analysis already discussed. This thus acts as a further, possibly subjective, means of interpretation.

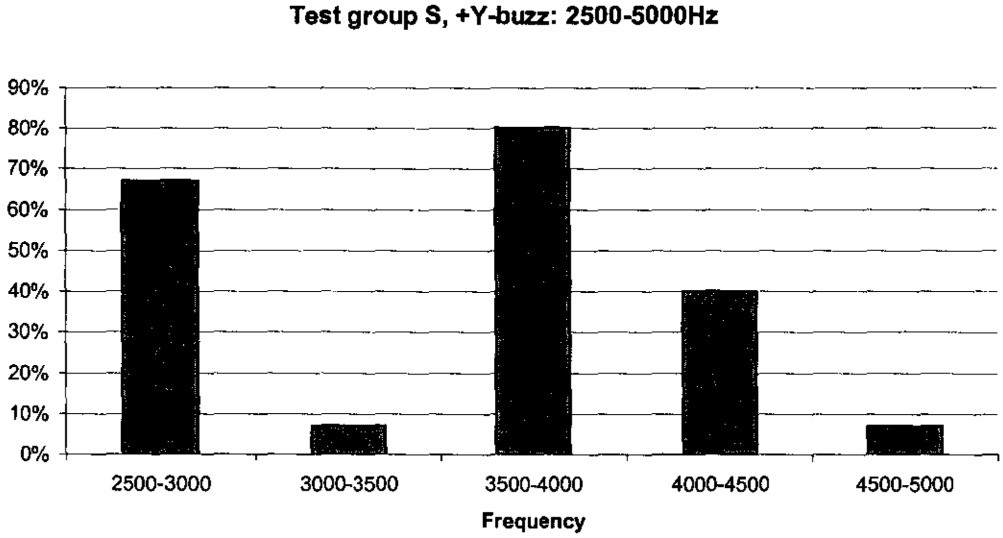


Figure 14: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group S: +Y-buzz.

The figure indicates an increase of energy in the 2500-3000Hz window for 67% of the participants and an increase of energy in the 3500-4000 Hz window for 80% of the participants. This correlates with the increases in energy that were found as statistically significant, as reflected in the bar graph of Figure 7.

Calls:

Although the F0 of the Calls was not given to the individuals, 57% used the same F0 in the post-training recording, 14% used a lower frequency and 29% a higher frequency. The dB of F0 were the same for only 14%, higher with 64% and lower fro 21%. Only 14% of the participants displayed a lower F1 frequency on the post- training recordings and 86% had the same F1 frequency. In 57% of the individual comparisons, the dB of the F1 increased, 36% had the same energy and 7% had a decrease of energy. Energy increase over the higher spectrum 2500-5000 Hz was observed as presented below in Figure 15.

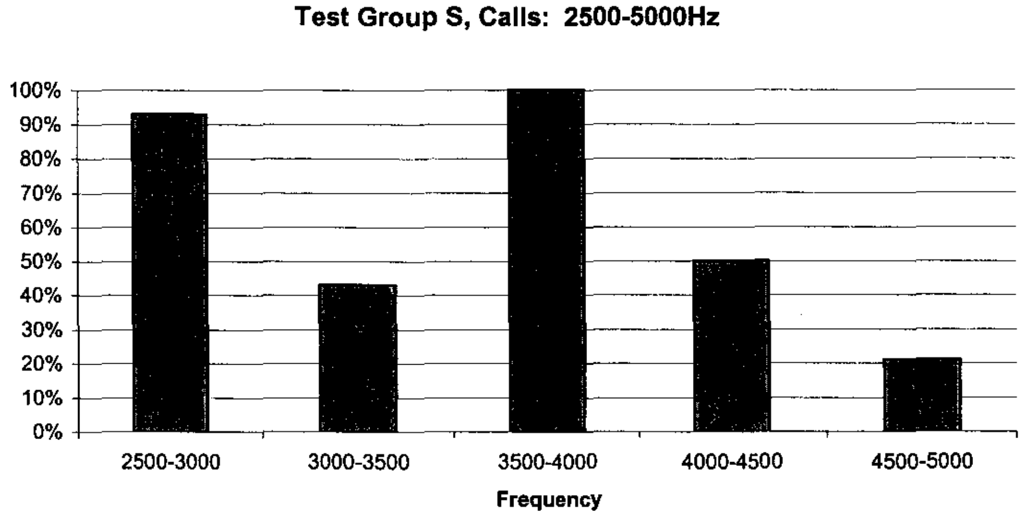


Figure 15: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group S: Calls.

The figure indicates an increase of energy in the 2500-300Hz window for 93% of the participants and an increase of energy in the 3500-4000 Hz window for 100% of the participants. This correlates with the increases in energy that were found as statistically significant, except for the 2700-2800 Hz window where the increase was not significant in relation to the F0 dB difference (see Figure 8).

English Phrases and Call words.

Although the F0 of the English Phrases and Call words was not given to the individuals, 46% used the same F0 in the post-training recording, 23% used a lower frequency and 31% a higher frequency. The dB of F0 were the same for only 15%, higher with 62% and lower fro 23%. Only 8% of the participants displayed a lower F1 frequency on the post-training recordings and 92% had the same F1 frequency. In 62% of the individual comparisons, the dB of the F1 increased and 54% had the same energy in relation to F0 dB difference. Energy increase over the higher spectrum 2500-500 Hz was observed as presented below in Figure 16.

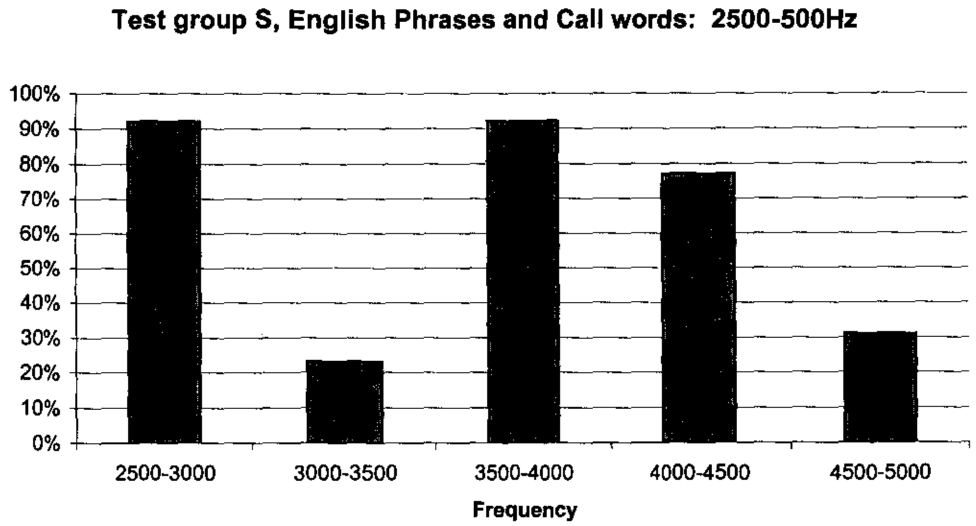


Figure 16: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group S: English phrases and Call words.

The figure indicates an increase of energy in the 2500-3000Hz window for 92% of the participants, an increase of energy in the 3500-4000 Hz window for 92% of the participants and an increase in the 4000-4500 Hz window of 77%. This correlates with the increases in energy that were found as statistically significant (see Figure 9).

First language texts:

Obviously the F0 of the First language texts was not given to the individuals. Interesting to note, 73% used the same F0 in the post-training recording, 20% used a lower frequency and only 7% a higher frequency. The dB of F0 was the same for 60%, higher with only 27% and lower for 14%. Energy increase over the higher spectrum 2500-5000 Hz was observed as presented below in Figure 17.

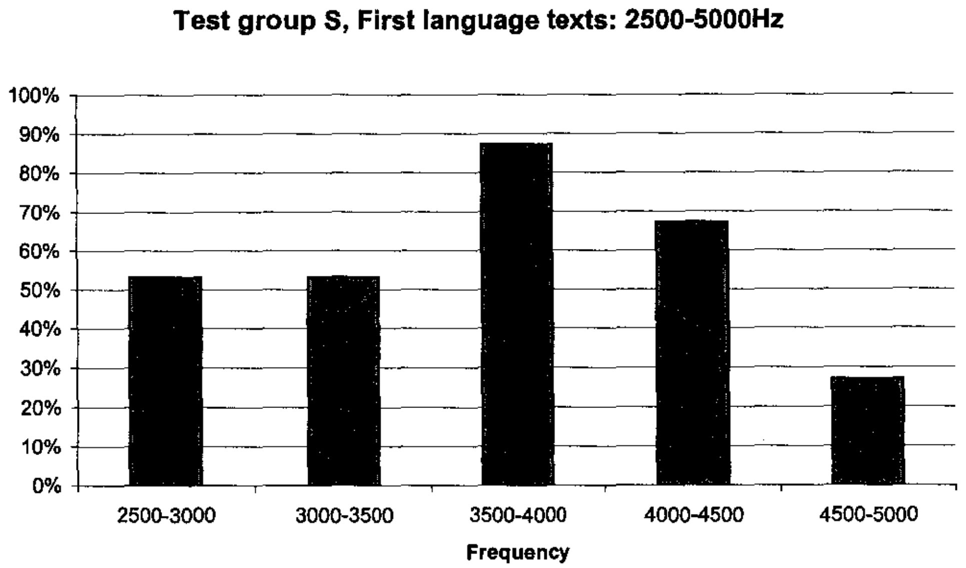


Figure 17: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group S: First language texts.

The figure indicates an increase of energy in the 2500-3000 Hz and 3000-3500 Hz windows for 53% of the participants, an increase of energy in the 3500-4000 Hz window for 87% of the participants and an increase in the 4000-4500 Hz window of 67%. This correlates with the increases in energy that were found as statistically significant as indicated in the bar graph of Figure 10.

4.5.9. Discussion

The perception panel preferred those utterances contained in the post-training recordings. The preference is statistically very significant. This preference can possibly be explained by the acoustic profile changes that took place.

Various means of processing the acoustical data have been used. An LTAS of the pre- and post-training recordings has been graphically presented. Visual interpretations have been made from these graphic displays of the LTAS of the pre- and post-training recordings of each sound mode. The numerical results from the LTAS were used to do an extrapolation of the highest amplitude in each 100 Hz window, for pre- and post-training recordings. This was then normalised in relation to F0dB. Inferential statistics as well as descriptive statistics were used to investigate the differences between the pre- and post-training recordings. A comparison was done between the LTAS and the statistical significant data. Further qualitative observations were done on each individual's sound modes and reported on in this Section as qualitative observations. The individual changes were reflected in group percentages. These were compared and discussed.

Although each of these various methods of processing highlight the differences and specifically improvements between the pre- and post-training recordings in diverse ways, the various methods support each other to provide a clear view of the over-all differences between the pre- and post-training recordings.

All five these post-training sound mode groups indicate an F0dB increase. In all of these five post-training sound mode groups an increase in energy can be observed in the 2500-3000 Hz region. This is especially true for the + Y-buzz, Calls, and the English Phrases and Call words. Another frequency region that indicates a clear increase in energy is the 3500-4000 Hz region. This is true for all the sound mode groups.

All five sound mode groups reflect differences in the lower spectrum (200-2500 Hz). This is possibly due to the vowels being enunciated "more overtly." An argument can thus be made that the Lessac Tonal NRG influences the shaping of the oral cavity in such a way that vowels are pronounced "better." This can however not be stated without doubt due to the fact that in attempting to keep the holistic flow during the actual recordings, the exact length of the vowels (especially in diphthongs) could not be controlled. Further research focussing on this specific phenomenon needs to be undertaken.

It is observed that F0 is sometimes lower in the post-training recordings. This may be due to relaxation of the larynx. The F0 is in a few cases higher in the post- recording. The probability exists that the voice was placed too low before. This is, however mere speculation but opens the door to future research where an investigation regarding the effect of the Lessac Tonal NRG as a laryngeal "relaxer" can be researched.

F1 did not, in general, decrease to a lower frequency as expected, but this drop of the F1 frequency did sometimes occur. The drop of F1 frequency would be in line with previous research by Scherer and Raphael (1987). The F1dB, in general, did increase in relation to F0dB difference. A general tendency to have a cluster peak drops in frequency in the post-training recording LTAS was observed. Due to the energy increase that did take place in the higher frequencies, a less steep slope over the LTAS did surface. This will, in line with existing literature, lead to improved projection. Further comparisons between the findings of this section and existing literature will be presented in Chapter 5.

In Section 4 a certain amount of repetition will take place seeing that the basic design and process of the research undertaken with Test Group V is exactly the same as that with Test Group S as reported on in Section 3. It has to be noted that the contact time, language and cultural proportions in Section 4 is different than those of Section 3.

4.6. SECTION FOUR: Test Group V

This research with Test Group V is conducted with a group of female students in a voice building class that is part of the curriculum of an Actor training programme at a tertiary institution in South Africa.

4.6.1. Aim

To investigate whether the use of Lessac's Tonal NRG will enhance the voice quality of the female actor's voice when it is used in a voice building class in a tertiary institution in South Africa, where the majority of females had either Afrikaans or South African English as first language.

4.6.1.1. Sub-aims

- To determine, through the use of a perception panel, whether the preferred sound is contained in the pre- training or post-training recordings.
- To investigate whether any statistically significant difference, as indicated by the perception panel, between the pre-training and post-training recordings exists.
- To compare and interpret the graphic presentations of the LTAS of the pre- training and post-training recordings in relation to the F0 dB difference, in order to investigate whether an improved acoustic profile exists in the post-training recordings.
- To investigate whether any of the acoustic differences between the pre-training and post-training recordings is statistically significant.

4.6.2. Participants

- 13 females.[70] Age between 18 and 23. Afrikaans: 8, English: 4, Northern Sotho: 1.[71]
- Taught by the same Certified Lessac teacher as the test group in Section 3.

---

[70] The total number of students in the class was 32. Some were men and some females who either had vocal health issues and were seeing a speech therapist or did not volunteer to be involved in this study.

[71] This is representative of student profile of this tertiary institution in the year that this teaching was done.

4.6.3. Ethical considerations

These female students all voluntarily agreed to participate in this research having been assured that the voice modes will be referred to anonymously and that their identities would be protected.

4.6.4. Training Period

This group had a total of 14 contact hours over a period of 14 weeks.

4.6.5. Training process

The Lessac Approach was used.
Initially three-dimensional breathing and optimal body integration were explored, followed by the introduction and exploration of Tonal NRG as defined and described in the Lessac Approach (see Chapter 2 for details).
An organic developmental flow was crucial so as to follow the holistic nature of the Lessac Approach. The class was thus introduced to the Y-buzz and then proceeded through +Y-buzz to Calls and lastly to Call phrases.

Classes were conducted in English and Afrikaans. The use of self-developed first language phrases as equivalents to the English Calls was encouraged. Students were expected to work, in buddy groups,[72] on the explorations introduced in class, as preparation for the next class.

4.6.6. Data collection

4.6.6.1. Recordings

- Pre- and post-training recordings were done in a sound-treated room at the tertiary institute.

---

[72] The concept of "buddy groups" is explained in Section 3.

- Recordings done onto a DAT (Sony ZA5ES Super bit mapping) recorder. Rec. level: 5;sampling rate of recording: 44.1kHz; Input: analogue microphone.
- Microphone used: Shure SM48. Dynamic ~~LOZ~~- Unidirectional.
- Microphone-to-mouth distance: 40cm.
- A SPL meter was used as guidance to not overload DAT. Students were asked to not go lower than 65 dB nor higher than 75 dB. It seemed to be easy for the students to stay within these parameters during the pre training recordings. They struggled to maintain this in the post training recordings even going as high as 90 dB. When this occurred, they were then asked to repeat the recording.

4.6.6.2. Tasks expected from participants

In the pre-training recording students were asked to do an "uh-uh" sound. This was used to determine the pitch given for the Y-buzz sound in pre- as well as post-training recordings.

Instruction given for modes 1-3: "Please do the (different sounds inserted – named and demonstrated) as long and loud as is comfortably possible while staying in the parameters 65-75 on the SPL meter."

Instruction given for modes 4 and 5: "Please read the following (either English Sentences and Call words or first language texts) in a comfortable volume for performers in speaking voice."

Different sounds named and demonstrated:[73]

1) Y-buzz on certain pitch as determined. Pitch given on keyboard as determined during the pre-training recording.
2) +Y-buzz
3) Calls
4) English Phrases and Call words
5) First language text readings of approximately one minute without use of words containing an "s" sound.[74]

---

[73] Seeing that the instruction indicates that the utterances were to be done as long as what is comfortably possible, the recordings had different lengths of time reflecting on the participants' levels of competence. The first language readings were all more or less on minute.

[74] As a high frequency noise pattern.

4.6.7. Perceptual evaluation: design

- 8 randomly chosen modes of each different sound (3 different modes of each subject) were played to 5 theatre experts. Three of the 5 had at least 5 years of training in performer's voice. The other two work fulltime with students in tertiary institutions, as well as working professionally as directors and performers. The latter 2 each have more than 10 years experience.
- The evaluation questionnaire was provided to panellists with the definition of good voice quality as used in this study.[75] The questionnaire was set up according to a 5 point rating scale: 1 = very poor, 2 = poor, 3 = average, 4 = good, 5 = very good.
- Pre- and post-training recorded modes where played to panellists in random order. Samples of the different modes of Test Group V used in Section 4, and Test Group S used in this section were randomised. The randomised samples were still ordered according to the structured teaching progression – random samples of the Y-buzz mode were kept together, and so forth.
- Modes were played from a CD through a PC using a high quality amplifier and stereo speakers: Creative Inspire2.1 2400[76] in order to maintain good sound quality.
- Evaluation sheets were filled in anonymously in the presence of the researcher.
- Cross tabulation tables will be used for descriptive statistics. These will provide an indication of the preference of the perception panel should it exist. These tables will also further serve as preparation for the inferential statistics.
- For triangulation more than one inferential statistics analyses were done on this data. 1) Paired t-test: H_O: $\mu_{diff(pre-post)} = 0$; H_A: $\mu_{diff(pre-post)} \neq 0$; $\alpha = 0.01$. H_o accepted when the p-value > α; H_o rejected in favour of H_A when the p-value < α. In this study H_o will imply that the scores given to the pre-training recordings minus the scores given to the post training recordings will equal null. Should this be true it will indicate that there was no change, according to the perceptual panel, in the sound quality of the post-training recordings. But should H_0 be rejected in favour of H_A, especially if the pre-

---

[75] As defined in 1.1, good voice quality (sustainable over periods of time and repeatable) for theatre may be "defined as a result of an optimal use of the vocal organ in order to establish the maximum possible acoustic output by minimal muscular effort" (Laukkanen, 1995:18).

[76] Technical specifications: Satellite: 4.5 watts. 12 ms power per channel (2 channels); subwoofer power: 12 watts RMS; Freq.response: 42 Hz-20kHz; SNR: >75dB; Dimensions: Satellites (LxWxH) - 8.7cm x 9.5 cm x 9.5 cm, Subwoofer – 21.1cm x 19.2 cm x 19.2 cm.

scores minus the post-scores provide a negative numerical, it will imply that the total of the post-scores was higher than the total of the pre-scores. This will indicate that the utterances contained in the post-training recordings have improved in voice quality according to the perceptual panel. 2) Chi-square: H_o: 2 variables (phase and score) are independent; H_A: 2 variables (phase and score) are dependent. $\alpha = 0.01$. H_o accepted when the p-value > α; H_o rejected in favour of H_A when the p-value < α. In this study this H_o will mean that the scores were randomly attributed to the phases (pre- and post- training recordings) and that there is no correlation between the scores given to the voice quality and whether it is a pre- or post-recording. H_A in this case means that the pre- recordings have lower scores attributed to them and the post-recordings have higher scores attributed to them. This will indicate that the training had a positive influence over the voice quality as perceived by the perception panel.

4.6.7.1. Data analysis and processing of the perceptual evaluation

The evaluation scores of the perception group were entered into Excel spread sheets and statistically processed through the use of SAS (Statistical Analysis Software package, version 8). A paired t-test for independence was done to compare pre-score means with post-score means for all the sound modes combined. For this paired t-test, null hypothesis testing was done where the mean difference equals zero (H_o: $\mu_{\text{diff(pre-post)}} = 0$), against the alternative where the mean does not equal zero (H_A: $\mu_{\text{diff(pre-post)}} \neq 0$).

Conclusively the paired t-test for all the sound modes combined, has the null hypothesis (H_o) as:

No statistically significant difference between the pre- and post-training recordings (referred to as phase in the t-test) of Test Group S (H_o: $\mu_{diff(pre\text{-}post)} = 0$) as indicated by the scores allocated by the perception panellists;

and the alternative hypothesis (H_A) as:

A statistically significant difference between the pre- and post-training recordings (phase) of Test Group S (H_A: $\mu_{diff(pre\text{-}post)} \neq 0$) as indicated by the scores allocated by the perception panellists. For this study to indicate that the training did affect the voice quality positively, the post-training recordings' scores have to be higher than those of the pre-training recordings.

Frequency procedures were carried out and cross-tabulation tables were done on the different variables for all the different sound modes. These procedures and tables provide a descriptive profile of the weighting of the scores in relation to the pre- and post-training recordings as phase.

For the sake of triangulation, chi-square tests were also done on appropriate cross tabulations but due to expected cell frequencies being less than 5, regroupings were done. Sound modes 1, 2 and 3 (being Y-buzz, +Y-buzz and Calls) were grouped as Lessac Tonal explorations, whilst sound modes 4 and 5 (being English Phrases and Call words as well as first language texts) were grouped as applications seeing that these sound modes represent the carry-over of the explorations into functional speech. In the H_o (null hypothesis) of the chi-square test the two variables (phase and score in this study) are independent. The question to determine is whether this H_o will be accepted or rejected for the alternative where the two variables are dependent.
Conclusively the chi-square test for the Lessac explorations as well as the applications, has the null hypothesis (H_o) as:
The scores (ratings 1-5) were assigned/attributed to the utterances of the pre- and post-training recordings (phase) of Test Group S at random, without any statistically significant relations between the ratings and the 2 phases (being pre- and post-recordings). This will be indicated as the p-value being > 0.01;
and the alternative hypothesis (H_A) as:
The scores (ratings 1-5) were assigned/attributed to the utterances of the pre- and post-training recordings (phase) of Test Group S according to a statistically relevant relationship between the ratings and the 2 phases (pre- and post-recordings). For this study to indicate an improvement of voice quality, the ratings allocated to the post-training recordings have to be higher than the ratings allocated for the pre-training recordings. This will be indicated as the p-value being < than 0.01.

The different language groups were sometimes too small, and as such cannot be used to provide any statistically significant indication. Descriptive statistical procedures were done for the different first language groups.

4.6.7.2. Perceptual evaluation results

As mentioned above, a t-test for independence[77] was done for all the sound modes combined, as part of this study. The rule used was that H_o will be rejected if the p-value $< \alpha$ (alpha) with alpha =0.01. It was decided on alpha as 1% to make the Type 1 error as small as possible.[78] The null-hypothesis where the mean difference equals zero (H_o: $\mu_{diff} = 0$) was thus rejected in favour of the alternative where the mean difference is significantly different from zero (H_A: $\mu_{diff} \neq 0$) seeing that the p-value was shown as <.0001 and it implies that the p-vlaue is $< \alpha$ (see Table 1). Since the mean differences were negative,[79] it indicates that the pre-scores are significantly less than the post- scores. The perceptual evaluators, in total, thus preferred the voice quality of the post recordings of the test group.

Table 1: T test for independence, Test Group V.

Analysis Variable: DIFF						
N	Mean	Std Dev	Minimum	Maximum	t Value	Pr> /t/
200	-1.2850	1.0339	-4.0000	2.0000	-17.58	<.0001

The cross tabulation tables for all the different sound modes (Y-buzz, +Y-buzz, Calls, English phrases and Call words, as well as first language text readings) indicated a clear difference between the pre- and post-training recordings where the post-recordings[80] were reflective of the preferred voice quality. In Table 2 the cross tabulation table of the Calls of Test Group V[81] is provided.

---

[77] Paired data

[78] Alpha (α) = P(Type 1 error); Alpha = P(reject H_0 when H_0 true)

[79] Observe the mean in Table 1.

[80] In the Cross tabulation Tables the pre-training recording is indicated as A and the post-training recording as Z.

[81] In frequency of score per phase.

Table 2: Group V, Calls: Table of phase by score.

	Score					
Phase	**1**	**2**	**3**	**4**	**5**	**Total**
A	7	29	4	0	0	40
Z	0	0	19	16	5	40
Total	7	29	23	16	5	80

Table 2 indicates that the perceptual evaluators rated the Call modes of the pre-training recording (A) primarily as very poor –score 1 (7/40 divided by 100 = 17.5%), poor - score 2 (72.5%) and average – score 3 (10%). The weighting of the post-training recording (Z) perceptual evaluation leans strongly towards average - score 3 (47.5%), good – score 4 (40%) and very good – score 5 (12.5%). This pattern is basically followed in all 4 the other sound mode groups[82] with the score weighting moving from being centred on score 1 and 2 to score 4 and 5.[83]

As mentioned under "Data analysis and processing of the perceptual evaluation", chi-square tests were done on appropriate cross tabulations. For chi-square the H_0 is that the two variables (phase and score) are independent. This will be tested against the alternative (H_A), which indicates the two variables as dependent. Regrouping had to be done due to expected cell frequencies being less than 5. Sound modes 1, 2 and 3 (Y-buzz, +Y-buzz and Calls) were grouped as Lessac Tonal explorations, whilst sound modes 4 and 5 (English Phrases and Call words as well as first language texts) were grouped as applications because these sound modes represent the carry-over of the explorations into functional speech. The score possibilities were also combined so that scores 1 and 2 on the perception panel questionnaire are in the chi-square test regrouped as score 1; score 3

---

[82] The Y-buzz, +Y-buzz, English phrases and Call words as well as the first language texts all reflected the same pattern and are in the possession of the researcher.

[83] As mentioned in Section 3, the cross tabulation tables are descriptive statistics and don't have a null-hypothesis. This profile thus does not feed directly into the acceptance or rejection of an H_O, but, seeing that these tables are used in preparation for the chi-square test, it is already obvious that a positive relationship exists between score and phase, seeing that the weighting of the scores for the post-recordings is higher than the weighting of the scores for the pre-recordings.

on the perception panel questionnaire remains 3 in the chi-square test and scores 4 and 5 on the perception panel questionnaire are regrouped as score 5 in the chi-square tests. In the chi-square the H_0 (null hypothesis) has it that the two variables (phase and score in this study) are independent. This will be accepted should the p-value be $> \alpha$ and rejected should the p-value be $< \alpha$. Alpha (α) equals 0.01[84] in order to make the type 1 error as small as possible. The alternative (H_A) is that the two variables (phase and score) are dependent. Should the null hypothesis be rejected it will thus be an indication that a relationship exists between the rating received by the perception panel and the phase (pre- or post-training recording). This relationship will be proven to be positive or negative pending on the outcome of the cross tabulation and the chi-square test.

Table 3a below provides the cross tabulation (of phase by score) in preparation for the chi-square test. Table 3b provides the chi-square test results of the Lessac explorations – sound modes 1, 2 and 3.[85]

Table 3a: Cross tabulation in preparation for chi-square of the Lessac explorations.

Frequency Expected Percent Row Pct Col Pct

Table of phase by nscore				
phase	**nscore**			**Total**
	1	**3**	**5**	
A	102 56.5 42.50 85.00 90.27	16 32.5 6.67 13.33 24.62	2 31 0.83 1.67 3.23	120 50.00
Z	11 56.5 4.58 9.17 9.73	49 32.5 20.42 40.83 75.38	60 31 25.00 50.00 96.77	120 50.00
Total	113 47.08	65 27.08	62 25.83	240 100.00

---

[84] Please note that the numerical contributed to Alpha is only applicable for the perceptual evaluation. Should the p-value be $< \alpha$ in this case, it would be statistically very significant.

[85] Y-buzz, +Y-buzz and Calls.

Table 3b: The chi-square results of the Lessac explorations.

Statistic	DF	Value	Prob
Chi-Square	2	144.2951	<.0001

Table 4a below provides the cross tabulation (of phase by score) in preparation for the chi-square test. Table 4b provides the chi-square test results of the applications – sound modes 4 and 5.[86]

Table 4a: Cross tabulation in preparation for chi-square of the applications.

Frequency Expected Percent Row Pct Col Pct

Table of phase by nscore				
phase	nscore			Total
	1	3	5	
A	34 21 21.25 42.50 80.95	36 32 22.50 45.00 56.25	10 27 6.25 12.50 18.52	80 50.00
Z	8 21 5.00 10.00 19.05	28 32 17.50 35.00 43.75	44 27 27.50 55.00 81.48	80 50.00
Total	42 26.25	64 40.00	54 33.75	160 100.00

Table 4b: The chi-square results of the applications.

Statistic	DF	Value	Prob
Chi-Square	2	59.0856	<.0001

---

[86] English phrases and Call words and First language texts.

The tables indicate a preference of the post-training recordings. The Lessac explorations pre-training recordings only have 1.67% of their total amount for scores allocated as score 5, but the post-training recordings have 50% of their total scores allocated to score 5. Similarly, the application modes only have 12.50% of the scores being 5 for the pre-training recordings but 55% of the scores being 5 in the post-training recordings. Since the p-value is, in both the Lessac explorations[87] (Y-buzz, +Y-buzz and Calls), and the applications[88] (English Phrases and Calls words and first language), less than .0001, the null hypothesis[89] is rejected. It is therefore evident that phase does effect score. This indicates that H_0 where phase and score are independent is rejected in favour of the alternative where phase and score are dependent. The perception panellists rated the post-recordings in both cases (explorations and applications) significantly higher than the pre-recordings, indicating that the post training recordings contained the preferred sounds. A dependant positive relationship thus exists between the phase and the score. This pattern indicates that the perception panellists are of the opinion that the outcome of the training is in line with the definition provided to them of what good voice quality is (see 1.1 p.3 for details).

Although, as already stated, one has to be careful to make final conclusions about the various first language groups,[90] it can be used as in indication of a trend to be explored in further research. As such the cross tabulation (of phase by score) for the + Y-buzz of the three language groups[91] in Test Group V[92] are provided in Tables 5.

Table 5a: Group V, + Y buzz, African languages: Table of phase by score.

Phase	1	2	3	4	5	Total
A	2	2	1	0	0	5
Z	0	0	3	2	0	5
Total	2	2	4	2	0	10

---

[87] See Figure 3b.

[88] See Figure 4b.

[89] The null hypothesis is that the two variables are independent.

[90] This is due to the small numbers of participants in this study

[91] Afrikaans is taken as one language group, English as another and the 5 participants having an African language as first language is combined into a third group namely "African". The researcher admits that this is not an ideal situation. Further research into the use of the Lessac Approach for various African languages is planned.

[92] The cross tabulation in indicated in frequency of score by phase.

Table 5b: Group V, + Y-buzz, Afrikaans: Table of phase by score.

Phase	1	2	3	4	5	Total
A	5	16	3	1	0	25
Z	1	1	6	13	4	25
Total	6	17	9	14	4	50

Table 5c: Group V, + Y-buzz, English: Table of phase by score.

Phase	1	2	3	4	5	Total
A	3	6	1	0	0	10
Z	0	1	4	2	3	10
Total	3	7	5	2	3	20

The cross tabulation tables for the Calls of the three language groups (Tables 3 a, b, and c) illustrate a reflective profile of the actual ratings being assigned by the perception panel to the pre- (A) and post- (Z) training recordings. In all these profiled examples the post-training recordings were preferred, with a higher rating (of score) than the pre- training recordings.[93] Again, this is also the case with all 4 the other sound modes for all 3 the language groups.[94]

Cross tabulation tables were made for the scores of the three pre-training recordings, as well as three post-training recordings for each participant in test group V that were played to the perception panel. Table 6 depicts this cross tabulation for one of the participants, selected at random.

Table 6: Cross Tabulation Table: Group V, participant AA, all sound modes.

Phase	Score 1	Score 2	Score 3	Score 4	Score 5	Row Totals
A	0	14	1	0	0	15
Z	0	0	1	13	1	15
Total	0	14	2	13	1	30

---

[93] When comparing score 4 and 5 of A (Pre-training) and Z (post-training) the preference for the post- training recordings are eminent.

[94] Specific cross tabulations in possession of the researcher.

Seeing that this table is typical, it functions as an example to indicate that each participant's voice has, according to the perception panel, improved during the training process, as the post-training recordings ratings centre around the higher scores.[95]

Although an indication of the profiles of the perception panellists has been provided before, it will contribute to the effectiveness of this study to reflect on their reliability as a group. Seeing that the various samples of the sound modes of Test Group S were randomly interspersed with various samples of the sound modes of Test Group V, statistical analysis on the inter reliability of the perception panel was done according to the information gathered from both test groups. As such, an in-depth discussion about this will take place in the comparative section of this study (see Chapter 5). Reflecting on this forthcoming discussion, it is noted that all the perception panellists (raters) separately indicated a significant improvement in the post-training recordings versus the pre-training recordings with the results of the paired t-tests done for each rater indicating a p-value as <.0001. It is thus very clear that the H_o ($\mu_{diff} = 0$) is rejected in favour of the HA ($\mu_{diff} \neq 0$) with $\mu_{diff} < 0$ by each rater.

4.6.8. Acoustic investigation

4.6.8.1. Design, data analysis and processing of the acoustic investigation

Pre- and post-training recordings were transposed from DAT into CSL as NSP files. This was put onto rewritable CD's using a Hewlett Packard – Sonic Foundy –ACID (1998) programme. Sound modes were drawn into Multi-Speech programme. A Long Term Average Spectrum was done for each of the 5 sound modes[96] of each participant, as well as a group average LTAS for each sound mode. Settings used: FFT size: 512 (thus modes/frame: 257); pre-emphasis: 0.000; Window weighting: Hanning; Smoothing level: None.

The numerical data of the acoustical profiles of pre- and post-training recordings was then entered into MS Excel.

- F0 dB pre-training was subtracted from F0 dB post-training.

---

[95] See footnote 75.

[96] As previously mentioned these modes are the 5 utterances: Y-buzz, +Y-buzz, Calls, English phrases and Call words and first language texts.

- The highest amplitude within windows of 100 Hz pre-recording was subtracted from the highest amplitude within the same windows of 100 Hz post-recording. Thus the highest amplitude reading between 1500-1600 Hz pre-recording, was deducted from the highest amplitude reading between 1500-1600 Hz post- recording and so on (see Section 2 for more details).
- The difference between F0 dB post-recording and F0 dB pre-recording was then subtracted from all these answers leaving the amount of amplitude increase over and above the difference of F0. The differences were thus calculated in relation to F0 dB difference.

This data was statistically processed by the use of SAS (Statistical Analysis Software programme, version 8).[97] The null hypothesis testing whether the group mean, or median, of the difference between post- and pre- is equal, was tested against the alternative that the difference is positive ($H_o:\mu_{post\text{-}pre} = 0$; $H_A: \mu_{post\text{-}pre} > 0$). The first step was to determine whether a paired t-test (mean) or a Wilcoxon test (median) should be applied to this data, depending whether the data was normally distributed or not. Shapiro Wilk was used to test for this assumption of normality. Data was accepted as normally distributed when the p-value was >0.05. If the data was distributed normally, the t-test was used with H_o: $\mu=0$; H_a: $\mu>0$ and $\alpha=0.05$. H_o was accepted when the p-value (divided by 2) ≥ 0.05. H_o was rejected in favour of H_A when the p-value (divided by 2) ≤ 0.05. Otherwise the Wilcoxon distribution free method was used.[98]

Conclusively only the H_o and H_A of the t-test for the acoustic analyses, in general, are provided seeing that all the statistically significant differences occurred when the data was normally distributed and the t-test was thus used. The t-test was used for each of the 5 sound modes respectively and within these sound modes separately for each 100Hz window. However, to provide the t-test hypotheses for each 100 Hz window of each of the 5 sound modes will result in repetition and duplication.[99]

---

[97] Credit to the Department of Statistical Support of the Pretoria Technikon for assistance and guidance in this matter.

[98] H_o: $\eta = 0$; H_A: $\eta > 0$; $\alpha = 0.05$; When p-value (divided by 2) of S ≥ 0.05 accepted H_o. Reject H_o in favour of H_A when p-value (divided by 2) of S ≤ 0.05 and S value is positive.

[99] This data is in possession of the researcher and can be obtained from her.

The null hypothesis (H_o) has that:

No statistically significant difference between the acoustic profiles of the pre- and post-training recordings (referred to as phase in the t-test) of Test Group S (H_o: $\mu_{diff(post-pre)} = 0$) *exists for each specific 100Hz window of each one of the sound modes;*

and the alternative hypothesis (H_A) has that:

A statistically significant positive difference between the acoustic profile of the pre- and post-training recordings (phase) of Test Group S (H_A: $\mu_{diff(post-pre)} > 0$) *exists for each specific 100Hz window of each one of the sound modes.*

For this study to indicate that the training did affect the voice quality positively, the post-training recordings' mean has to be statistically significantly higher over certain frequencies than the mean of the pre-training recordings. These frequency clusters will be referred to in the discussion and interpretation of the acoustical profile of the post-training recording modes.

Further qualitative observations about the frequency of F0 of the +Y-buzz, Calls, English phrases and Call words, as well as first language text readings were done.[100] Observations were also made in regards to F1 and F2 of the Y-buzz when applicable. The change in the energy profile over the 2500-5000 Hz spectrum in all the modes has been commented on as necessary.

4.6.8.2. Acoustic investigation results

The group average LTAS of the pre-training and post-training recordings of each sound mode of Test Group V is graphically presented:

---

[100] Seeing that the Y-buzz modes had a given pitch, the F0 frequency differences will only be commented on.

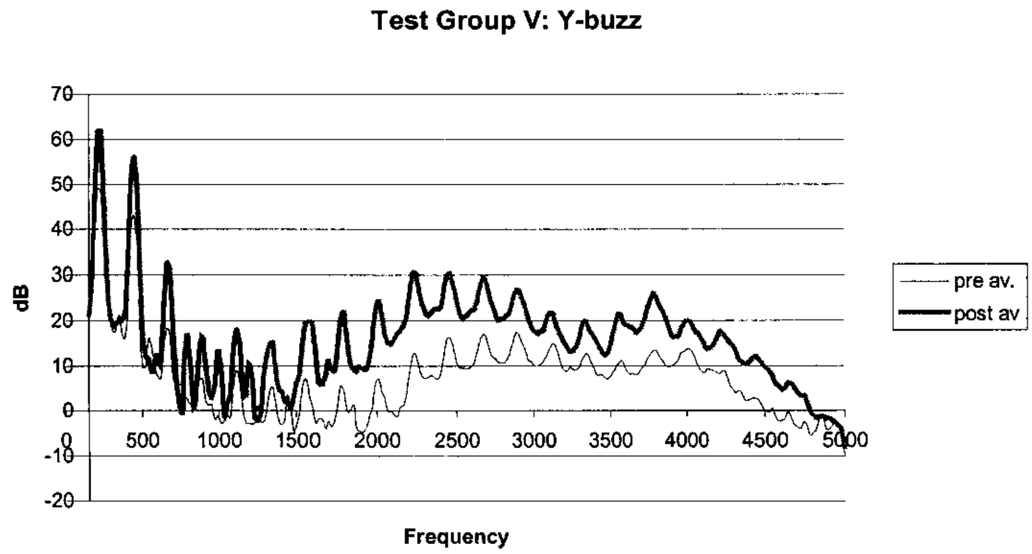


Figure 1. The average LTAS of the pre- and post-training of the Y-buzz of Test Group V.

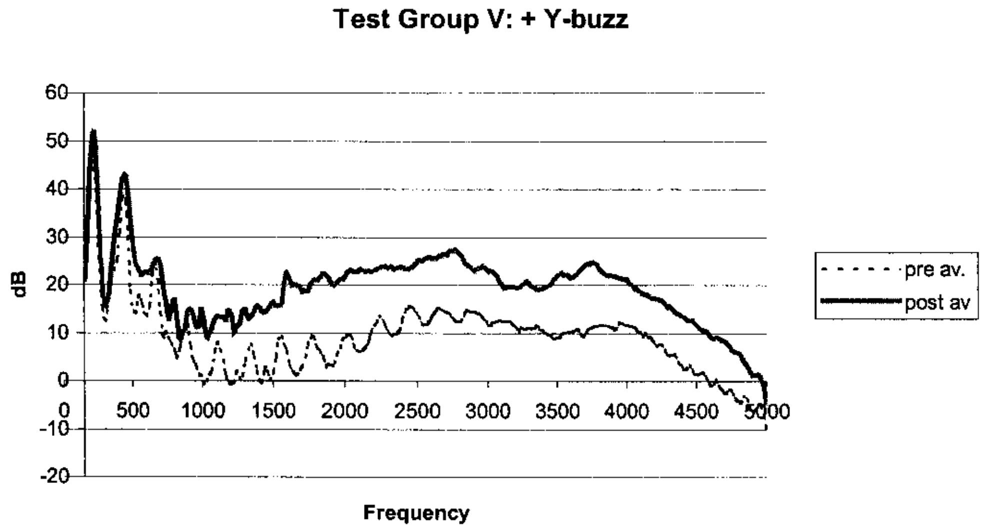


Figure 2: The average LTAS of the pre- and post-training of the +Y-buzz of Test Group V.

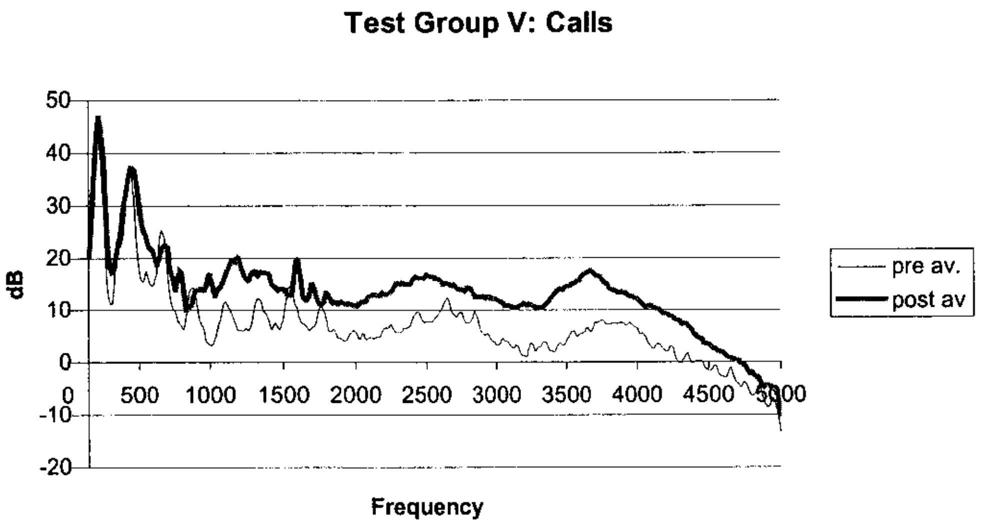


Figure 3: The average LTAS of the pre- and post-training of the Calls of Test Group V.

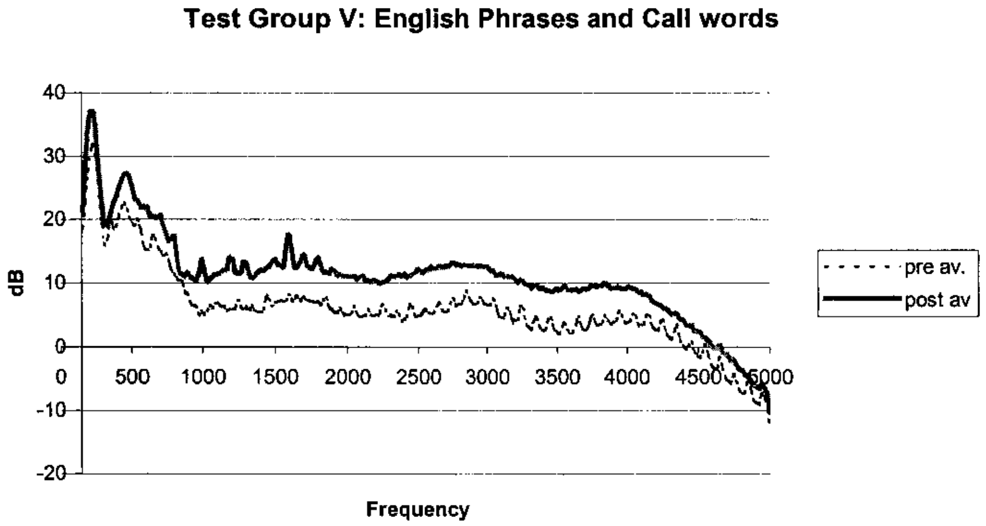


Figure 4: The average LTAS of the pre- and post-training of the English phrases and Call words of Test Group V.

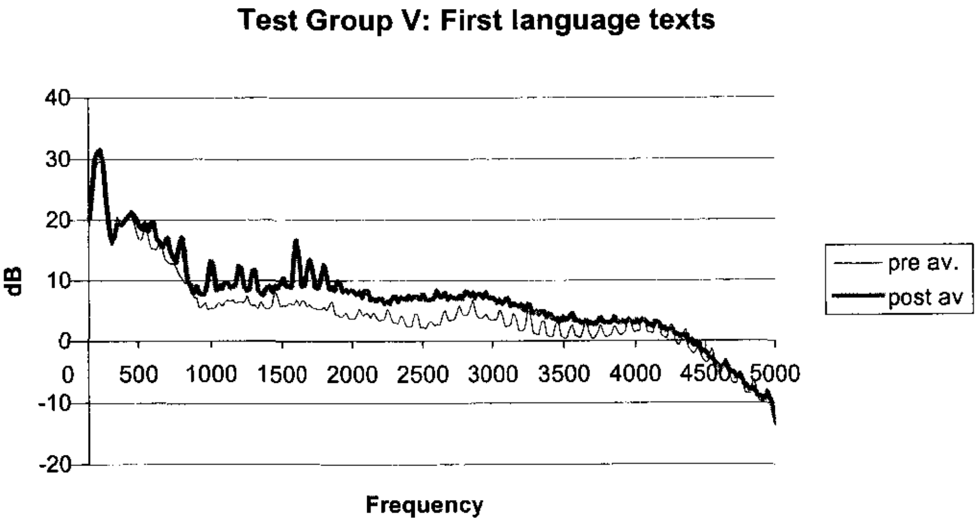


Figure 5: The average LTAS of the pre- and post-training of the first language texts of Test Group V.

As mentioned in the "Data analysis and processing of the acoustical investigation" the numerical results of the LTAS of each of these modes were entered into Excel. As in Sections 2 and 3, an approximation of the numerical results was done, where the highest dB within a window of each 100Hz was chosen. Following this approximation, the amplitude of each 100 Hz window of the first recording was subtracted from the amplitude of each 100 Hz window of the second recording,[101] providing the amplitude difference between each frequency window of the pre-training and post-training recordings. In an attempt to standardise these differences and interpret them in relation to the "loudness"[102] difference of F0, the difference between amplitude of the F0 of the post-training recording, and amplitude of the F0 of the pre-training recording, was subtracted from the differences of each of the 100 Hz windows.

Table 7 posits an extract as example of this process that was followed for each sound mode of each participant.[103] The participant referred to in Table 7 is referred to as AA and was randomly chosen.

---

[101] Post (Z) – Pre (A)

[102] The subjective term referring to energy and amplitude.

[103] See Section 2:4.4.6 in this chapter for a step-by-step demonstration of this process.

Table 7: An extract example of the processing of the data. Test Group V, participant AA, +Y-buzz.

Freq.	Pre	Post	Diff.	F0 diff	Diff in relation to F0
F0	50.34	50.63	0.29	n.a	0.29
200-300	50.34	50.63	0.29	n.a	0.29
300-400	20.2	19.93	-0.27	0.29	-0.56
400-500	49.6	46.41	-3.19	0.29	-3.48
500-600	15.57	46.08	30.51	0.29	30.22
600-700	30.76	25	-5.76	0.29	-6.05
700-800	6.87	25.49	18.62	0.29	18.33
800-900	15.47	27.55	12.08	0.29	11.79
900-1000	2.71	16.19	13.48	0.29	13.19

The "difference in relation to F0" (including F0) data of the separate participant's sound modes were then grouped according to mode type (Y-buzz, +Y-buzz etc) in order to process the data statistically. A mean and median were assigned to each 100 Hz window of each of these groups. The Shapiro Wilk test was used to test for normality. The data that was statistically significant was all normally distributed. For these the mean is used for the null hypothesis testing. In this particular case the null hypothesis is $\mu_{\text{post-pre}} = 0$. The mean difference (for post - pre - F0 dB difference) was tested against the alternate ($\mu_{\text{post-pre}} \neq 0$). H_0 was rejected in favour of H_A as $\mu_{\text{post-pre}} \neq 0$. Seeing that SAS automatically provides two-sided testing and this study is interested in one-sided testing, the p-value of the mean has to be divided by two, before comparing to Alpha.[104] Alpha in this case is 0.05. When the p-value divided by 2[105] is smaller than Alpha the increase will be statistically significant. The t-test outcome is graphically presented in the different bar graphs below (Figures 6,7 8, 9 and 10) where the mean/median differences are indicated (for the highest amplitude difference in each 100Hz window) for each sound on the following graphs. In each of these graphs, it is indicated with a dot where the mean dB in each 100Hz window[106] is

---

[104] The p-value divided by 2 < α; p-value divided by 2 < 0.05.

[105] SAS automatically provides a two-sided p-value. This study is only focussing on the increase of the mean and thus one-sided testing will be used. For one-sided testing, the p-value must be divided by 2 before comparing it with Alpha ($\alpha = 0.05$).

[106] After standardisation, e.g.: Highest dB between 300-400Hz Post minus highest dB 300-400Hz Pre minus difference dB of post minus pre F0.

statistically significant on a 5% level of significance.[107] The p-values (when divided by 2) of the means of those 100Hz windows were thus smaller than Alpha (0.05), which is statistically significant (p-value < 0.05). The Shapiro Wilk, mean, t-test, and p-value of each of the 100Hz windows that has a statistically significant increase, of each mode, are provided in the addendum to this study.

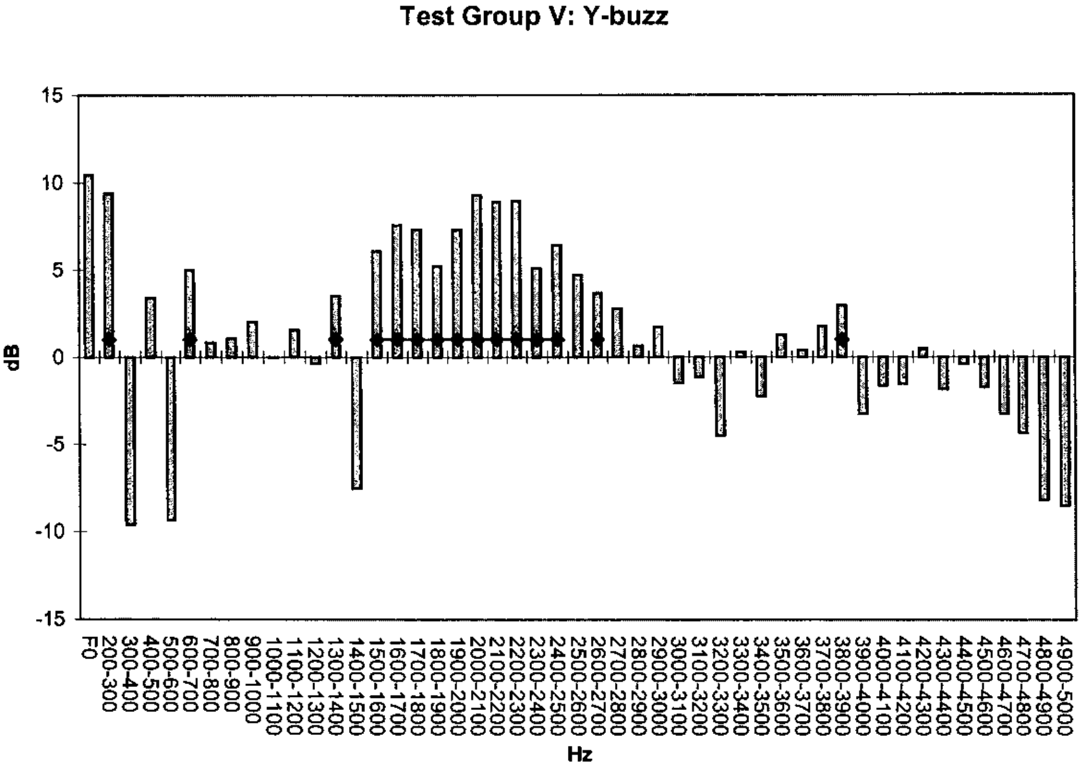


Figure 6: Bar graph of mean/median differences of the Y-buzz. Statistically significant differences (of mean) are indicated with dots. Statistically significant clusters are indicated by lines connecting the dots.

---

[107] SAS provides a two-sided p-value. Thus for one sided testing, the p-value must be divided by 2 before comparing it with Alpha ($\alpha = 0.05$).

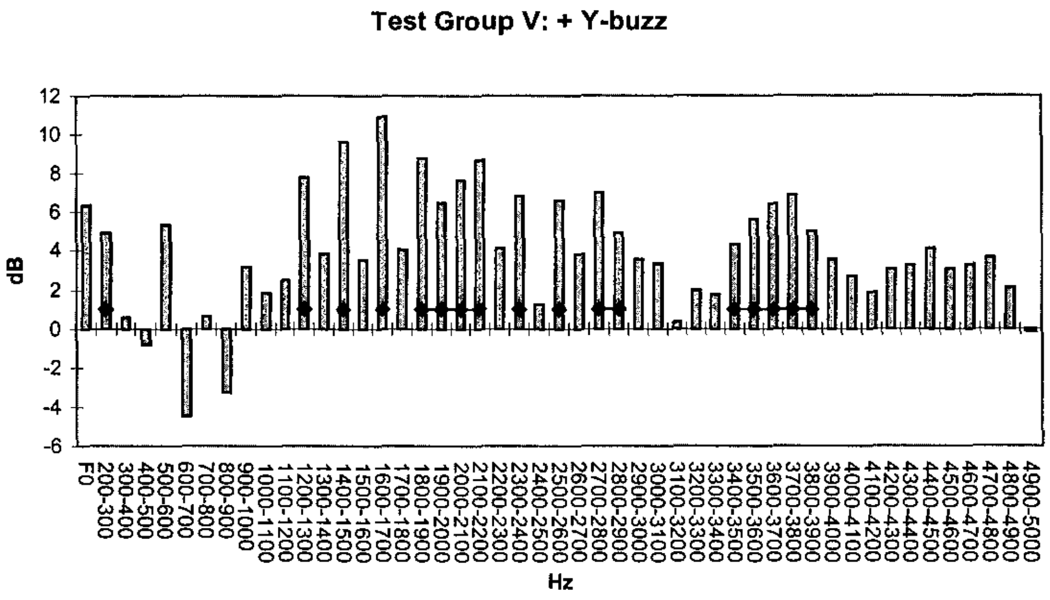


Figure 7: Bar graph of mean/median differences of the + Y-buzz. Statistically significant differences (of mean) are indicated with dots. Statistically significant clusters are indicated with lines connecting the dots.

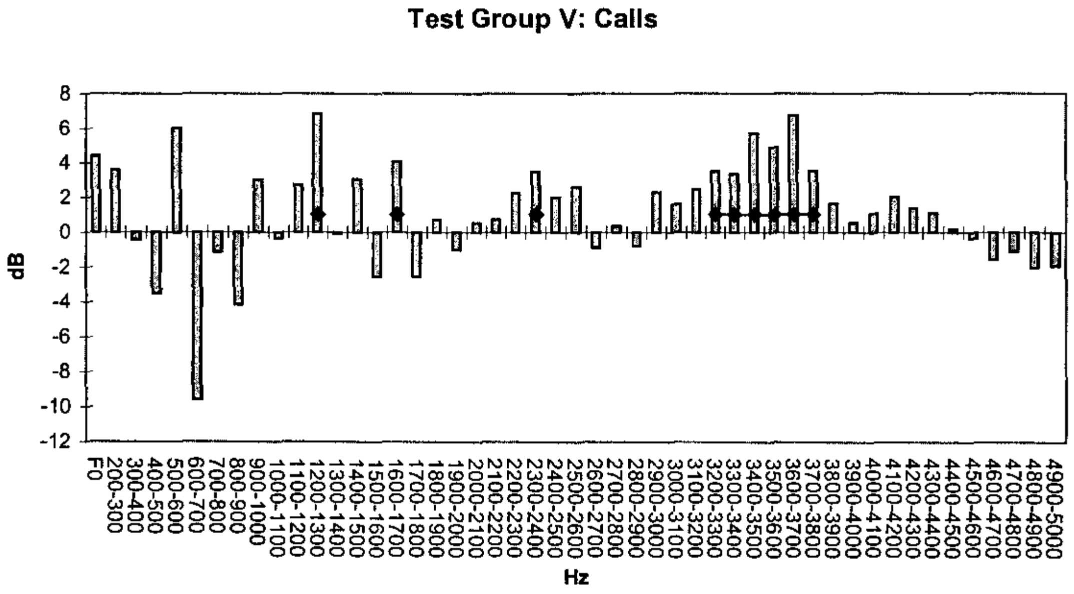
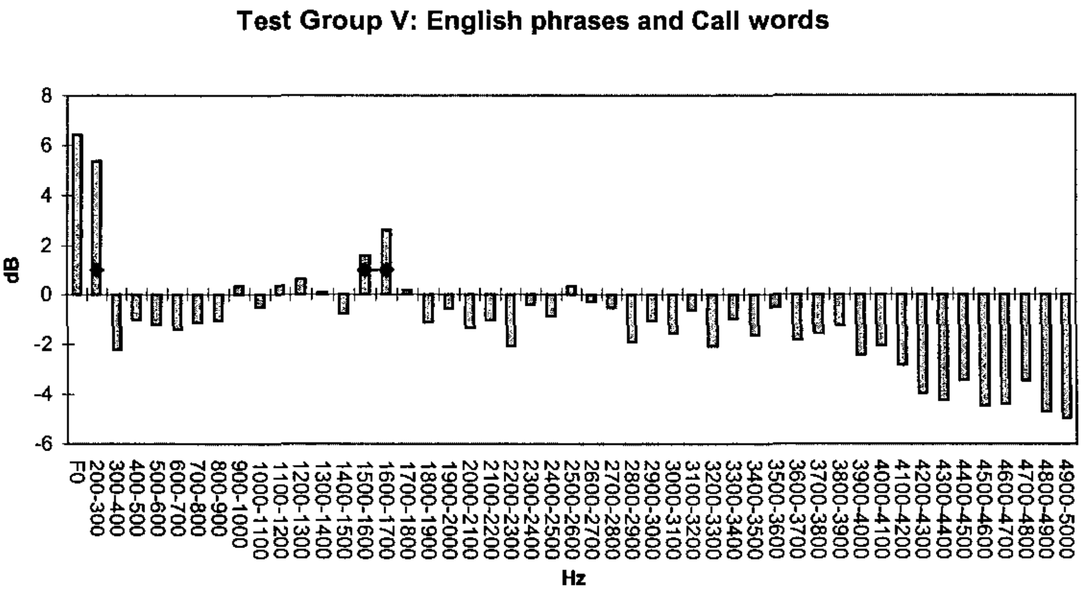
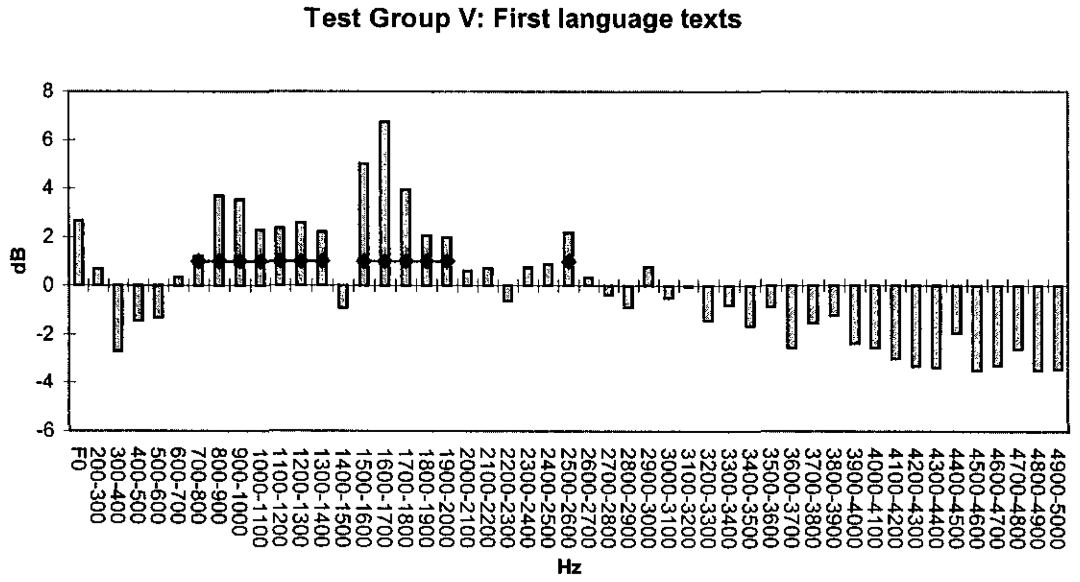


Figure 8: Bar graph of mean/median differences of the Calls. Statistically significant differences are indicated with dots. Statistically significant clusters are indicated with lines connecting the dots.



Test Group V: English phrases and Call words

Figure 9: Bar graph of mean/median differences of the English phrases and Call words. Statistically significant differences are indicated with dots. Statistically significant clusters are indicated with lines connecting the dots.



Test Group V: First language texts

Figure 10: Bar graph of mean/median differences of the First language texts. Statistically significant differences (of mean) are indicated with dots. Statistically significant clusters are indicated with lines connecting the dots.

A discussion based on the comparative profiles of the pre- and post-training recordings, as depicted in the LTAS (Figures 1-5) and the bar graphs[108] (Figure 6-10) of each mode group follows:

Test Group V: Y-buzz:

In order to demonstrate this process the LTAS and Bar Graph of the Test Group V: Y-buzz are placed next to each other in Figure 11 below.[109] This is the same process followed in Section 3 with Test Group S.

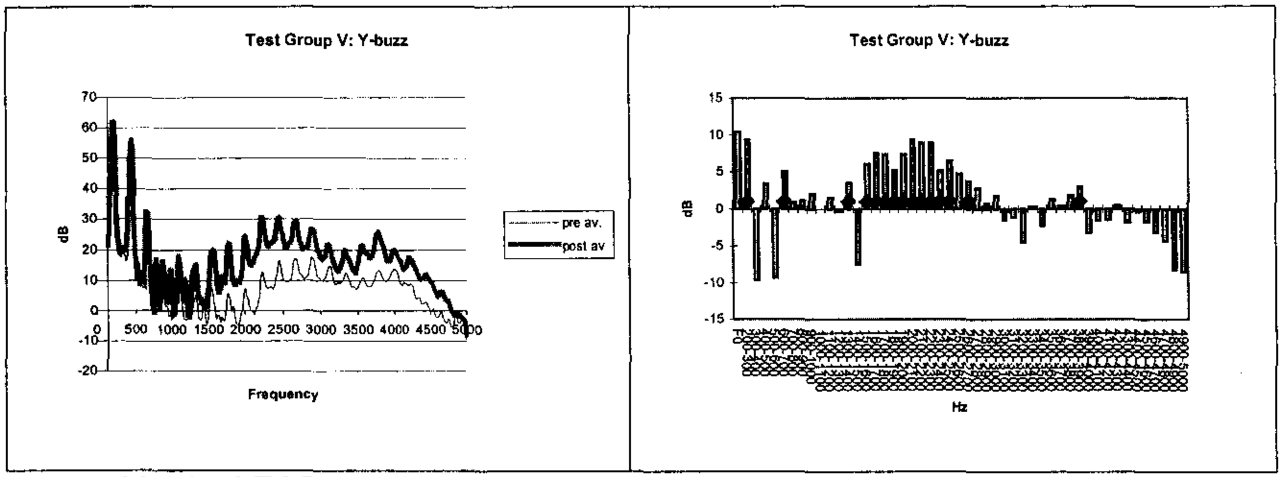


Figure 11. The LTAS and Bar Graph of the Test Group V: Y-buzz.

Post-training recordings reflect the following:

- A more than 10dB increase in the F0 dB.
- An increase in the 200-300 Hz region. For most participants this is their F0 range and as such it will be treated as a F0 dB increase.
- An increase in the 400-500 Hz range and a statistically significant increase of energy in the 600-700 Hz window.
- Should 2226.56 Hz be taken as the post-training recording's F2, then a slightly higher F2 frequency is observed in the pre-training recording. This reflects the use of a longer resonator, possibly due to the "forward facial orientation" in the post-training recordings.

---

[108] The mean/median differences of each 100 Hz window of each sound mode of test group V in relation to the F0 dB difference.

[109] Previously these Figures appeared in Section 4 as Figure 1 and Figure 6.

- A statistically significant increase, in relation to F0dB, in the region of 1500-2500 Hz. This increase is clearly indicated on the bar graph, as well as LTAS presentation, as being in the 1500-2800 Hz region – although not statistically significant for each 100 Hz window.
- In the LTAS presentation of the post-training recordings, an increase of energy in the 3574-3900 Hz region. This correlates with the bar graph, which indicates an energy increase in the 3500-3600 Hz and 3700-3900 Hz windows, although not statistically significant. The peak frequency in the 3500-4000 Hz region moved down in frequency in the post-training recordings, from 4023 Hz (13.56 dB) in the pre-training recording to 3789 Hz (25.82 dB) in the post training recording.[110] This is possibly due to the use of a longer resonator.[111]
- In the higher frequencies energy increase, in relation to the F0 dB difference, in the 2500-2800 Hz and 3500-4000 Hz regions.
- A sharp decrease of energy in the spectrum of 4000-5000 Hz, especially above 4300 Hz.

The smallest mean difference in the post-training recording[112] that is statistically significant, lies in the 2600-2700 Hz window where the mean is 3.711, the student's t is 2.030, and the p-value is 0.0652.[113]

The biggest mean difference in the post-training recording that is statistically significant lies in the 2000-2100 Hz window where the mean is 9.320, the student's t is 2.988 and the p-value is 0.0113.

Test group V: +Y-buzz:

Post-training recordings reflect[114] the following:

- An increase in the F0 dB. An energy increase that is statistically significant is observed in 200-300 Hz in the bar graph. As mentioned previously this is the F0 window for most of the participants and will be treated as such.

---

[110] This is the actual dB at these frequencies and not given in relation to F0 dB difference.

[111] Which, in turn, correlates with the use of the "forward facial orientation."

[112] This is above the difference of the dB of F0 – thus more than the dB increase observed in the F0 post-training recording.

[113] As explained in footnote 37, the p-value needs to be divided by 2 before being compared to Alpha (α = 0.05).

[114] Compare Figures 2 and 7.

- An energy increase, in relation to the F0dB difference, from 900-4900 Hz, although the increase is not statistically significant in each window.
- Strong increases, which may be related to the shaping of the oral tract to pronounce the diphthong, occur at 500-600 Hz, 1200-1300 Hz and 1600-1700 Hz. This interpretation is pure speculation and this phenomenon will have to be investigated in a separate study.
- There is an observable cluster between 2500-3000 Hz. The bar graph indicated a statistically significant increase in the 2500-2600 and 2700-2900 Hz windows.
- A second cluster between 3400-4100 Hz exists with a prominent peak at 3750 Hz. This peak is lower in frequency and more prominent than the one possibly observed in the pre-training recording (at 3984 Hz). This may reflect the use of the "forward facial orientation".
- Although not statistically significant, an increase in energy over the spectrum 4200-4900 Hz is eminent.

The smallest mean difference in the post-training recording that is statistically significant lies in the 3400-3500 Hz window where the mean is 4.304, the student's t is 2.296, and the p-value is 0.0405.[115] The biggest mean difference in the post-training recording that is statistically significant lies in the 1600-1700 Hz window where the mean is 10.929, the student's t is 4.077 and the p-value is 0.0013.

Test group V: Calls:

Post-training recordings reflect[116] the following:

- A slight energy increase in the F0 region.
- Changes in the LTAS profile indicates energy changes over the F0 – 2000 Hz spectrum. This is possibly due to the "more overt" pronunciation of the diphthong used in the Call sound.[117]
- On the LTAS presentation a cluster in the 2000-3000 Hz region is observed. The bar graph indicates an increase, in relation to F0dB difference, between 2000-2600Hz,

---

[115] Divided by 2 before comparing with Alpha ($\alpha = 0.05$) – see footnote 37.

[116] Compare Figures 3 and 8.

[117] Seeing that the length of the diphthongs used was not controlled during the recordings no definite claims can be made about this phenomenon. The length of the diphthongs was not regulated due to the aim, which was to keep in line with the free holistic essence of the Approach.

with the energy increase between 2300-2400 Hz, as statistically significant. This cluster peaks at 2519 Hz, which is a lower frequency than that of the slight peak observed in the pre training recording at 2656 Hz.

- An increase in energy, as reflected in the bar graph, happens at 2900-4400 Hz, with a statistically significant increase at 3200-3800 Hz. This correlates with the visual profile presented of the LTAS. A cluster in the region of 3400-4000 Hz with a prominent peak at 3671 Hz is clear. This falls in the 3600-3700 Hz window that has the biggest energy increase.
- A decrease in energy, in relation to F0dB difference, is eminent form 4500-5000 Hz.

The smallest mean difference in the post-training recording that is statistically significant lies in the 3300-3400 Hz window where the mean is 3.344, the student's t is 1.793, and the p-value is 0.0981.The biggest mean difference in the post-training recording that is statistically significant lies in the 3600-3700 Hz window where the mean is 6.713, the student's t is 4.275 and the p-value 0.0011.

Test group V: English Phrases and Call words:

Post-training recordings reflect[118] the following:

- A slight increase in the F0 dB. An energy increase that is statistically significant is also observed in the 200-300 Hz window but seeing that this is the region where most of the participants' F0 lies it will be interpreted as such.
- A statistically significant energy increase is presented between 1500-1700 Hz. This phenomenon is reflected in the Y-buzz of this test group. The +Y-buzz and the Call outcomes both reflect a significant increase between in the 1600-1700 Hz window.
- The visual profile of the LTAS indicates a move towards a cluster between 2500-3000 Hz but the bar graph shows that energy decreased in this area when interpreted in relation to the F0dB difference.
- Again the LTAS reflects a possible move towards a cluster between 3500-4000 Hz, but the bar graph indicates a decrease of energy in relation to the F0dB difference. Speculation may have it that the pattern of clustering occurs first and only then the

---

[118] Compare Figures 4 and 9.

energy increase. Of course this will have to be tested in a research project where same test group is recorded with intervals during the training process.

The smallest mean difference in the post-training recording that is statistically significant lies in the 1500-1600 Hz window where the mean is 1.588, the student's t is 2.369, and the p-value is 0.0372.[119] The biggest mean difference in the post-training recording that is statistically significant lies in the 1600-1700 Hz window where the mean is 2.619, the student's t is 3.441, and the p-value is 0.0055.

Test group V: First language texts:

Post training recordings reflect[120] the following:

- A very small increase in the F0 dB and the average F0 Hz is 195.35.
- The LTAS reflects clearer peaks and valleys in the F0 – 1800 Hz spectrum of the post-training recording. As interpreted previously, this may be due to improved pronunciation of the vowels that may reflect an improved awareness of the sensation of the bone conduction on the hard palate. The bar graph indicates a statistically significant energy increase in the regions of 700-1400 Hz and 1500-2000 Hz.
- The bar graph indicates only one other statistically significant energy increase, at 2500-2600 Hz. The overall profile in the 2500-5000 Hz spectrum is that if a decrease in energy.

The smallest mean difference in the post-training recording that is statistically significant lies in the 700-800 Hz window where the mean is 1.200, the student's t is 1.836, and the p-value is 0.0935.[121] The biggest mean difference in the post-training recording that is statistically significant lies in the 1600-1700 Hz window where the mean is 6.751, the student's t is 7.391, and the p-value is <.0001.

As previously mentioned, the groups of first language speakers were in the cases of English and African languages too small to do any statistical analysis. The LTAS was used to determine whether the first language groups reflect the patterns observed in the

---

[119] Divided by 2 before comparing with Alpha.

[120] Compare Figures 5 and 10.

[121] Divided by 2 before comparing with Alpha.

combined language groups. The same basic pattern primarily occurs in all different language groups.

For argumentation, the graphic LTAS presentation of the Call modes of the first language African speaker (Figure 12), as well as the + Y-buzz average of the English first language group (Figure 13) will be discussed. The focus will be on these two language groups as the Afrikaans first language group was in the majority in this test group.

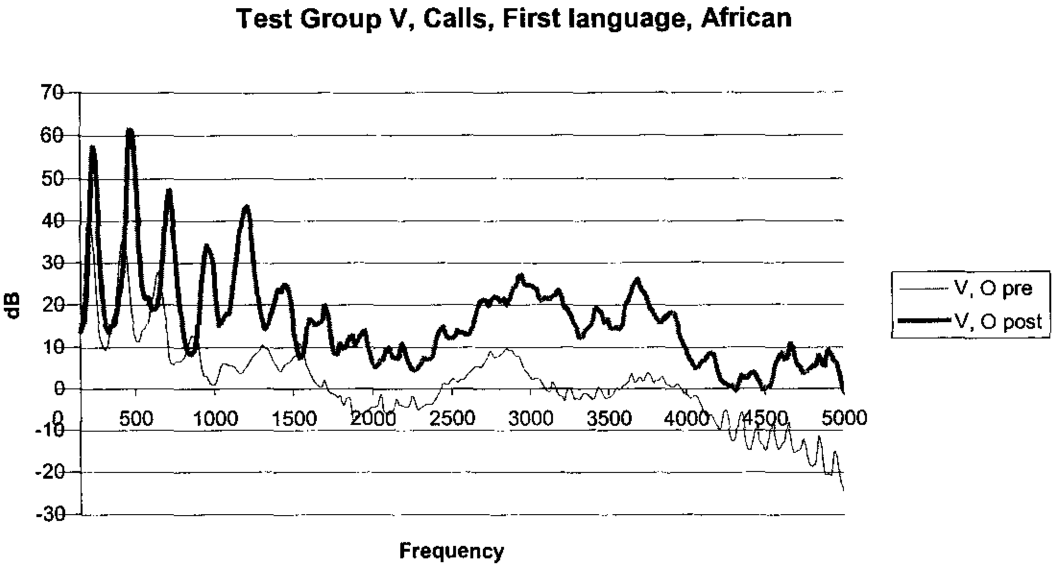


Figure 12. The LTAS of the Calls of test group V, First language African.

Test group V: Calls, First language: African:

The average LTAS of pre- and post-training recordings reflect that the post-training recording has:

- A strong increase of F0 dB.
- An increase of energy around F1, in relation to the F0dB difference.
- A change of the visual profile of the spectrum 500-2000 Hz with more pronounced peaks and valleys. This may refer to a "more overt" pronunciation of the diphthong used in the Call.
- Clearly defined clusters in the 2500-3200 Hz and the 3500-4000 Hz regions, with the latter reflecting a lowering of the peak frequency.
- An increase in energy in the 4500-5000 Hz region.

When comparing the LTAS of the Calls of test group V (Figure 3) and the LTAS of the sub-group first language African group only (Figure 12), strong resemblances exist, especially in the higher frequency regions:

- A tendency towards more pronounced clustering in the 2500-3000 Hz region , although the first language African participant's profile does not indicate a lowering of the peak frequency.
- A tendency towards a more pronounced cluster in the 3500-4000 Hz region with a lowering of the peak frequency.

Test group V: +Y-buzz, First language: English.

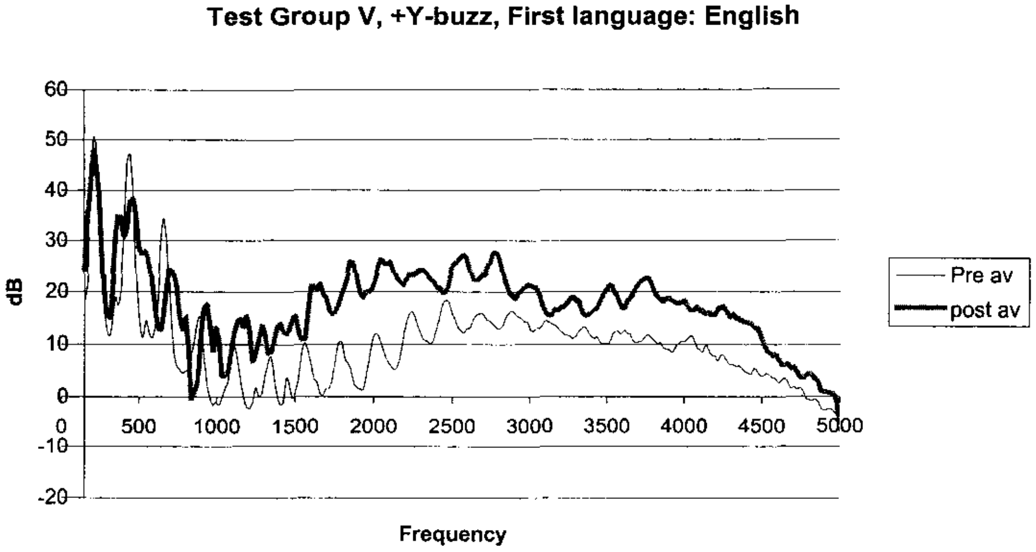


Figure 13. The average LTAS of the +Y-buzz, Test Group V, first language English.

The average LTAS of pre- and post-training recordings reflect that the post-training recording has:

- A decrease of energy at F0.
- An energy increase from more or less 1000 Hz up to 5000 Hz.
- A visual profile change between F0 and 2500 Hz that may indicate pronunciation adaptations as previously mentioned.
- A tendency towards clustering at 2500-3000 Hz and again between 3400-4000 Hz.

Although it has to be kept in mind that the English first language speakers are 31% of the total number of participants of Test Group V, it may be valuable to reflect on a comparison between the LTAS of the +Y-buzz of Test Group V, (Figure 2) and the LTAS of the +Y-buzz of the sub-group first language, English (Figure 13). Strong resemblances exist in the higher frequencies:

- A tendency to cluster between 2500-3000 Hz.
- A second possible cluster between 3400-4100 Hz.

These two examples reflect a relatively consistent pattern of the various languages in all 5 the different modes.

Further qualitative observations were made from individual LTAS of each participant for each sound mode. These observations were done in relation to the F0dB difference unless otherwise commented. These observations are not exact in nature but act as a mere indication of what tendencies are observed in the post training recordings. It is realized that only after 14 hours of training, the differences in the sound modes of Test Group V are still very tentative. These qualitative observations will thus rather be made to record the process of change.

Y-buzz:

The F0 of the Y-buzz was given to the participants in both the pre-training as well as post-training recordings. 84% of the post-training recordings reflect the same F0 frequency. 85% of the participants increase the F0dB whilst 15% used the same amount of energy and thus the dB stayed the same. 92% of the LTAS post-training recordings (Y-buzz) indicated that the frequency of F1 stayed the same, whilst 8% was lower in frequency. Only 15% of the dB of F1 post-training recordings stayed, in relation to F0dB the same, whilst 23% dropped in energy and 62% increased in energy. Whilst reflecting on the 2500-5000 Hz spectrum 92% of the post-training recordings have a movement towards clustering and sometimes an increase of energy (however small or large) in the 2500-3000 Hz region; 23% in the 3000-3500 Hz region; 69% in the 3500-4000 Hz region and 38% in the 4000-50000 Hz region. This is graphically displayed in Figure 14.

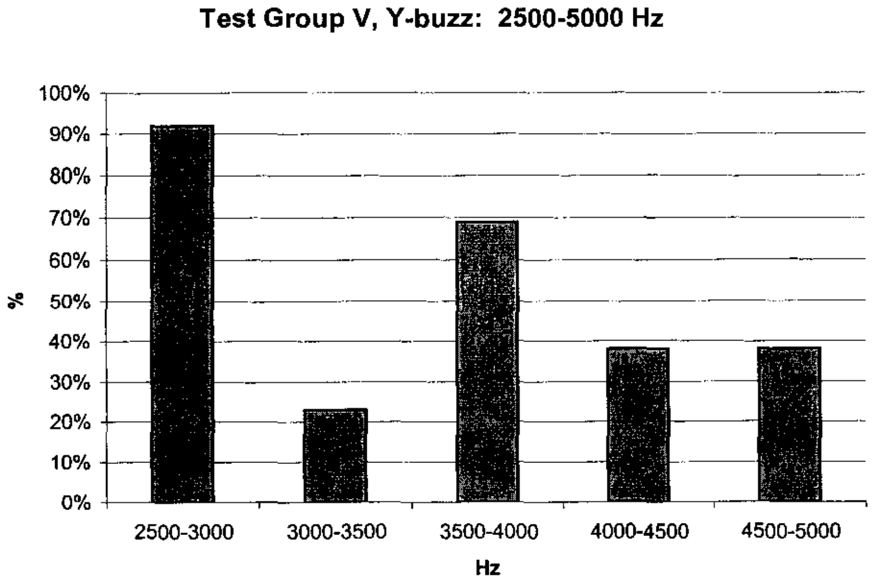


Figure 14: A bar graph demonstrating the percentage of change between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: Y-buzz.

+Y-buzz:

Although the F0 of the +Y-buzz was not given to the individuals, 62% used the same F0 in the post-training recording, 15% used a lower frequency and 23% a higher frequency. The dB of F0 were either the same (16%) or higher with 84%. Only 15% of the participants displayed a lower F1 frequency on the post-training recordings and 62% had the same F1 frequency whilst 23% displayed a higher F1 frequency. In 31% of the individual comparisons, the dB of the F1 increased, 15% had the same energy and 54% had a decrease of energy. Movement towards clustering and sometimes a slight energy increase over the higher spectrum 500-5000 Hz was observed as presented below in Figure 15. These tendencies referred to in general and the specific dB increase is not addressed, as it is taken care of in the LTAS and bar graph analysis already discussed. This thus acts as a further, possibly subjective, means of interpretation.

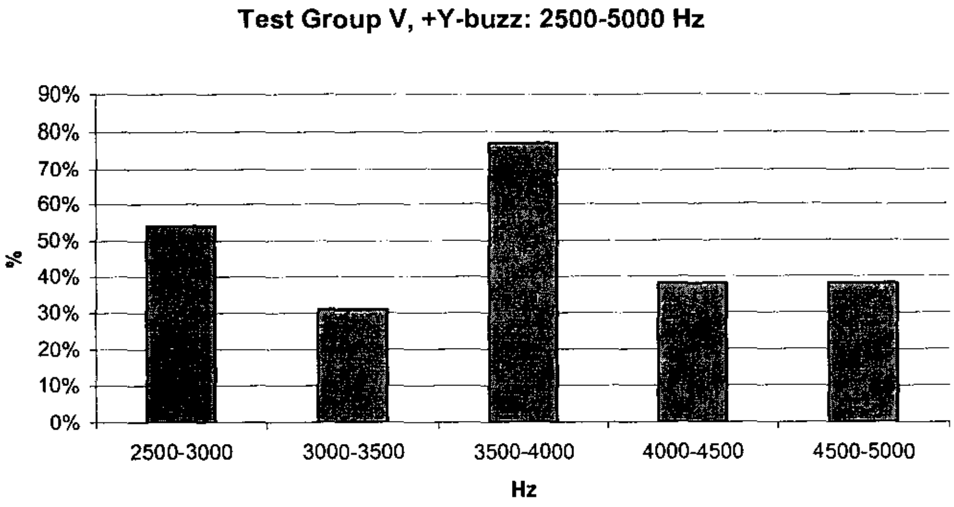


Figure 15: A bar graph demonstrating the percentage increase of the energy between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: +Y-buzz.

The figure indicates an increase of energy in the 2500-300Hz window for 54% of the participants and an increase of energy in the 3500-4000 Hz window for 77% of the participants. This correlates with the increases in energy that were found as statistically significant, as reflected in the bar graph of Figure 7.

Calls:

Although the F0 of the Calls was not given to the individuals, 69% used the same F0 in the post-training recordings, 8% used a lower frequency and 23% a higher frequency. The dB of F0 was the same for 23%, higher for 54% and lower for 23%. 77% of the participants displayed the same F1 frequency on the post-training recordings, with only 8% indicating a lower F1 frequency and 15% a higher F1 frequency. In only 23% of the individual comparisons, the dB of the F1 increased, none had the same energy and 77% had a decrease of energy.

Movement towards clustering and sometimes a slight energy increase over the higher spectrum 2500-500 Hz were observed as presented below in Figure 16.

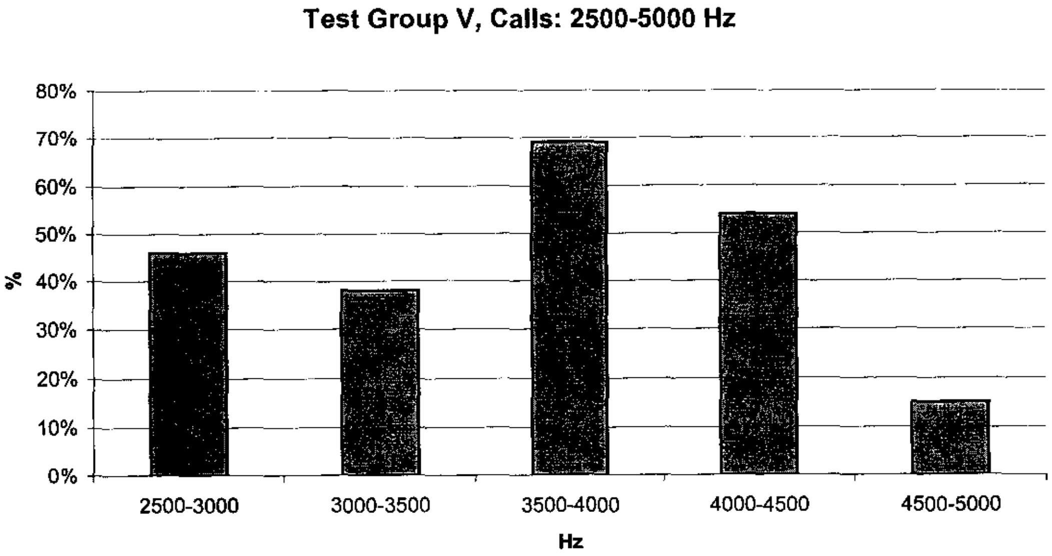


Figure 16: A bar graph demonstrating the percentage of change between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: Calls.

The figure indicates a movement towards clustering and sometimes an increase in energy in the 2500-3000 Hz window for 46% of the participants, in the 3500-4000 Hz window for 69% of the participants and in the 4000-4500 Hz window for 54% of the participants.

English phrases and Call words.

Although the F0 of the English phrases and Call words was not given to the individuals, 83% used the same F0 in the post-training recording. The dB of F0 was higher for 92% and lower for 8% of the participants. Only 16.5% of the participants display a lower F1 frequency on the post-training recordings and 67% had the same F1 frequency. In only 17% of the individual comparisons, the dB of the F1 increased and 75% had the same energy in relation to F0 dB difference. Movement towards clustering and sometimes a slight energy increase over the higher spectrum 2500-500 Hz was observed as presented below in Figure 17.

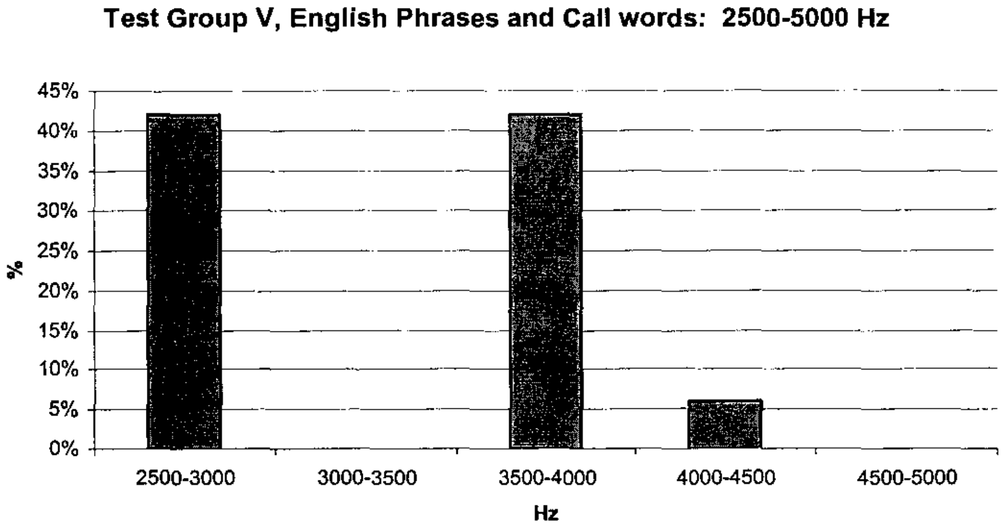


Figure 17: A bar graph demonstrating the percentage of change between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: English Phrases and Call words.

The figure indicates a movement towards clustering and an increase of energy in the 2500-3000 Hz window for 42% of the participants, and in the 3500-4000 Hz window for 42% of the participants.

First language texts:

Obviously the F0 of the first language texts was not given to the individuals. Interesting to note, 83% used the same F0 in the post-training recording, 17% used a lower frequency. The dB of F0 was the same for 25%, higher for 67% of the participants and lower for 8%. Movement towards clustering and sometimes a slight energy increase over the higher spectrum of 2500-5000 Hz were observed as presented below in Figure 18.

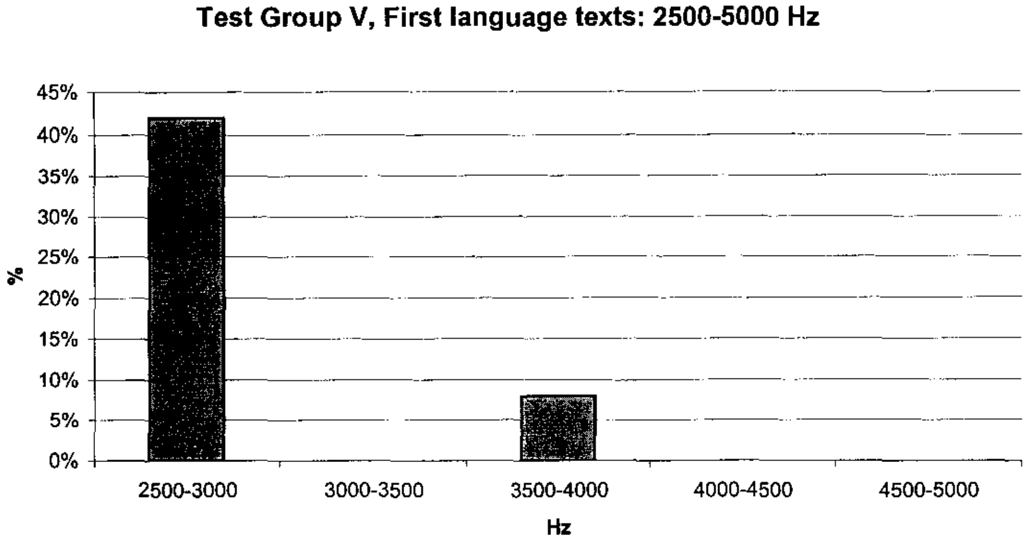


Figure 18: A bar graph demonstrating the percentage of change between 2500-5000 Hz as observed in each individual LTAS of the Test Group V: First language texts

The figure indicates a movement towards clustering and a possible increase of energy in the 2500-3000 Hz window for 42% of the participants.

4.6.9. Discussion

Although the perception panel preferred the post-training recordings with a statistically significant difference, the picture becomes more complicated when evaluating the acoustic profiles of the post-training recordings in comparison to the pre-training recordings. It has to be taken into account that 14 contact hours is very little time to learn any approach that leads to different behavioural patterns. This is more so when it has to influence neuro-muscular patterning. Interpreting the differences that do appear in the post-training recordings, comparative to the pre-training recordings, may however, be important, as this knowledge may contribute to the planning of a voice curriculum for performers in training and here specifically for the female performer in training. This can further make a contribution when used as an addition to the findings of Test Group S (discussed in Section 3), which had 28 contact hours during their training process.

The positive contributions of the Lessac Approach on the female performers' voice after 14 contact hours are:

Y-buzz:

- A F0dB increase;
- A change in the LTAS profile over the spectrum F0-2500 Hz. This may, as previously mentioned, be due to the change in vowel pronunciation. This in turn, may explain why the perception panel preferred the post-training recordings although there is no clear enhancement of energy in the higher frequencies. The trained ears of the perception panellists preferred the pronunciation in the post-training recordings and possibly detected the movement towards clustering as well as the small energy increases that happened over the higher spectrum frequencies (2500-5000 Hz);
- The lowering of the frequency of the cluster peaks;
- An energy increase in the 2500-3000 Hz and 3500 – 4000 Hz regions.

+Y-buzz

- A F0dB increase;
- A change in the LTAS profile over the spectrum F0-2500 Hz – this is possibly due to pronunciation changes.
- A tendency towards clustering as well as an energy increase in the 2500-3000 Hz and 3500-4000 Hz regions with the latter reflecting a lower peak frequency.

Calls:

- A slight energy increase of F0;
- A change in the LTAS profile over the spectrum F0-2500 Hz – this is possibly due to pronunciation changes;
- An eminent cluster between 2000 – 3000 Hz with a peak in the 2500 Hz region;
- An energy increase between 3500 – 4000 Hz with a lower peak frequency.

English phrases and Call words as well as First language texts.

Nothing definite is observed in these sound modes. This is possibly due to the small amount of contact hours, as the application of the change of the habitual patterning was still too vulnerable.

- Most observable are the changes from F0- 2500 Hz.
- A decrease of energy is observed in the higher frequencies, which may indicate a process of shift between the old habitual pattern of resonator shaping and the new learned pattern.
- Pure speculation may have it, though, that there is a tendency to cluster in certain higher frequencies although not definite energy increase has taken place after 14 contact hours.

Where Chapter 4 reported on the empirical studies done for this project, Chapter 5 will compare the outcomes of these projects and compare them with the existing trends in relevant literature.

Chapter 5

Summary, Conclusions and Shortfalls

5.1. Restating the field of investigation

As mentioned in Chapter 1 the actress has the same vocal demands placed on her as is placed on the actor, namely to communicate. This communication includes the use of the voice in such a way that the sound can carry over a distance without the voice losing flexibility. This furthermore leads to the need for a sustained healthy voice usage. Various voice-training systems claim that "they" can cater for such demands. One of these is the Lessac Approach. Previous research indicated that the Lessac Approach contributes positively to the development of the actor's voice.

This study set out specifically to investigate the effectiveness of 1) the Tonal NRG of the Lessac Approach to 2) enhance projection of 3) the female actor's voice
4) irrespective of language differences of the participants 5) or the idiosyncrasies of the certified Lessac teachers and training methodologies.

This aim had to a greater or lesser extent some variables included in the investigative process:

- Language (Teacher's first language, language of instruction and first language of the participants);
- Training methodology differences by means of
- 4 different teachers teaching the one American group and one teacher teaching both the South African groups;
- Contact hours during the training process – where the control group had no training, Test Group V had 14 contact hours in 14 weeks, Test Group S had 28 contact hours in 14 weeks and the American Group had a six week intensive workshop.

Another variable that had to be dealt with was the adaptation of the research design and process as the study progressed, thus leaning toward action research. Although this hampers direct correlations between the groups, it adds to the value of the study as it strongly indicates that the process followed with the analysis and interpretation of Test Groups S and V provides richer material.

The various empirical investigations were:

- the experiment with and use of the American Group where only LTAS were analysed and interpreted;
- the use of a Control Group where the LTAS were analysed and interpreted with the added dimension of the subtraction of the F0 db difference from the dB difference observed in each 100 Hz window;
- the use of Test Groups S and V where a perception panel was used. Their evaluations were statistically analysed by means of t-tests, cross tabulations and chi-square tests. For the acoustic investigation the LTAS graphic presentation as well as inferential statistics were used to determine the difference between the pre- and post-training recordings. Triangulation to determine veracity and accuracy thus took place, which provided solid material for interpretation.

5.2. Structuring the discussion of the findings

Firstly a compilation of the findings of the four different empirical experiments will be made and then discussed, with an extrapolation of the most consistent changes found in the post- or second recordings. Secondly, an evaluation of the Tonal NRG of the Lessac Approach will be given according to the three categories used in voice analysis based, namely, on acoustics, perception and the physiology. This will lead to a conclusive argument of the value of the Lessac Approach with specific emphasis on the Tonal NRG in the building of the female actor's voice.

In order to trace the most consistent changes that exist in the post-training recordings, a comparison between the four experiments discussed in Chapter 4, is necessary. It has, to be kept in mind, however, that a direct comparison is not possible due to the different acoustic analysis procedures followed. Strong indications exist that the various and different ways of analysis assist and reflect off each other to provide an in-depth picture of the acoustic differences of the pre- and post-training recordings and should be used together, rather than relying on only one specific analysis process. Tables 1-5[1] (see below), where all the differences are indicated, provide an overview of the changes.

The information gleaned from the post-/second recordings of the four different experiments have been ordered in these tables according to the amount of actual training and contact hours. The Control Group with no training is first, followed by Test Group V with 14 contact hours, then Test Group S with 28 contact hours and lastly the American Test Group, with a six-week intensive workshop.[2] Having the order in the tables in this way, it is possible to follow the process of change and development that took place during the training. The argument is that the Control Group had no training and as such the differences observed in the second recording will reflect what happens with the voice during the described modes, when increased breath pressure is used. Test Group V will reflect on the next step of changes, with Test Group S the next level and the American Test Group possibly reflecting on the changes that occur when not only the Tonal NRG has been introduced and practised, but also the other elements of the Lessac Approach. Seeing that the other elements of the Lessac Approach fall outside the scope of the study further research will have to be done in this regard.

It would obviously have been more beneficial if the analytical proceedings used for Test Groups S and V had been used throughout this study but reasons for this not occurring have already been discussed in Chapter 4 and the adaptation of the design has certain values in itself. Language or teacher specificity is not addressed in these tables (1-5) as it has been reflected on indirectly in Chapter 4 and will be discussed later in this chapter.

---

[1] Each different sound mode appears separately on a table.

[2] This is not the order used in Chapter 4 where the order is chronologically established.

5.3 Findings

5.3.1 Acoustical findings

The concept of the "actor's formant" or "speaker's ring" is used as guideline to investigate any possible improvements that were observed in the post-training recordings. Only observations relevant to the improvement of the voice quality and projectability are indicated in Tables 1-5. A block may actually be blank because the observation of what happened in that section does not necessarily contribute to the argument. Table 1 reflects the changes found in the Y-buzz mode, Table 2 those of the +Y-buzz mode, Table 3 the Call mode, Table 4 the English phrases and Call words mode and Table 5 the First language texts mode. As such the Control Group and the American Test Group will not reflect any information in Table 5 as these groups did not have first language texts. Table 1-5 provide (as mentioned above) additional insight into the process of repatterning, seeing that an indication of the chronological process of change that occurs during the voice building process when using the Tonal NRG is given.

Table 6 reflects an extrapolation of the most consistent changes observed in the different training groups for all modes, possibly indicating what areas over the spectrum are the most "sensitive" to change.

It is important to remember, when interpreting these tables, that there are areas where an increase did take place but the design of the research with that specific group did not include inferential statistics (Control Group and the American Test Group). These energy increases will be referred to only as an increase in energy. When referring to Test Groups S and V, some frequency windows may reflect and increase in energy that was observed but not statistically significant – these will only be referred to as an increase in energy. When an increase in energy occurred that was statistically significant, it will be clearly indicated as such.

Table 1: Y-buzz Mode. A comparative table providing results that demonstrate tendencies.

Modes	Analysis		Control group -No Training-	Test group V -14 contact hours-	Test group S -28 contact hours-	American test group 6 week intensive workshop
Y-buzz	F0	Hz	same	same	same	same
		dB	slight increase	More than 10dB increase	Statistically significant increase	increase
	F1	Hz	same	same	same	same
		dB	same	same	Same or increased	increase
	F0-2500 Hz		slight dB increase in: 800-900; 1000-1100, 1900-2000, 2100-2200, 2300-2400 Hz	• Profile change; • DB increase 400-500 Hz; • Statistically significant increase 600-700, 1300-1400Hz; • Mostly statistically significant difference 1500-2500 Hz. • Lowering of F2 peak freq.	• Profile change; • dB increase 400-500 Hz; • statistically significant increase 1800-2400 Hz;	Profile change Due to system of analysis no specifics can be provided
	2500-3000 Hz		Slight dB increase in: 2800-2900 Hz	dB increase 2500-3000 Hz	dB increase 2500-3000 Hz	dB increase
	3000-3500 Hz					dB increase; valley
	3500-4000 Hz		Slight dB increase in: 3600-3800 Hz	dB increase 3500-3900	• dB increase 3500-4000 Hz, • statistically significant between 3600-3800 Hz	dB increase
	4000-4500 Hz		Slight dB increase in: 4100-4200 Hz	decrease	Some windows reflect a dB increase	DB increase
	4500-5000 Hz			Sharp decrease	decrease	Some participants had dB increase

Table 2: +Y-buzz Mode. A comparative table providing results that demonstrate tendencies.

Modes	Analysis		Control group -No Training-	Test group V -14 contact hours-	Test group S -28 contact hours-	American test group 6 week intensive workshop
+Y-buzz	F0	Hz	same	same	same	same
		dB	decrease	increase	Statistically significant increase	Same or decrease
	F1	Hz	same	same	same	same
		dB		same	increase	Mostly same
	F0-2500 Hz		slight dB increase in: 400-500 1100-1200, 1300-1400 1500-1600 1700-1800 2000-2100 Hz	• Profile change; • dB increase 500-600, 900-2500 Hz, often statistically significant.	• Profile change; • dB increase 1100-2500 Hz – mostly statistically significant.	Profile change Due to system of analysis no specifics can be provided
	2500-3000 Hz			• Cluster observed; • DB increase 2500-2600, 2700-2900 Hz	• Cluster observed with lowered peak freq. • Statistically significant dB increase	Clusters observed dB increase sometimes lowering of peak freq.
	3000-3500 Hz			• Statistically significant increase at 3400-3500 Hz	Clear Valley	valley
	3500-4000 Hz			• Cluster observed; • Statistically significant increase at 3500-3900 Hz	• Cluster observed with lowered peak freq.; • Statistically significant dB increase	Clusters observed dB increase sometimes lowering of peak freq.
	4000-4500 Hz		slight dB increase in: 4200-4300 Hz	dB increase 4200-4500	Slight dB increase	
	4500-5000 Hz			dB increase 4500-4900 Hz	Statistically significant dB increase	

Table 3: Call Mode. A comparative table providing results that demonstrate tendencies.

Modes	Analysis		Control group -No Training-	Test group V -14 contact hours-	Test group S -28 contact hours-	American test group 6 week intensive workshop
Calls	F0	Hz	same	same	same	same
		dB	slight increase	Slight increase	Slight increase	Mostly decrease
	F1	Hz	Same	same	same	same
		dB	Slight increase	Increase or decrease	increase	increase
	F0-2500 Hz		• Slight increase (under 4 dB) up to 2200 Hz. • No clear profile change.	• Profile change; • dB increase 500-600; 900-1000, 1100-1300, 1400-1500,1600-1700, 2200-2500 Hz; • statistically significant in 1200-1300, 1600-1700, 2300-2400 Hz; • cluster peak at ~2500 Hz	• Profile change; • dB increase 400-2500 Hz; • statistically significant at 500-600, 900-1400, 2100-2500 Hz.	Profile change Due to system of analysis no specifics can be provided
	2500-3000 Hz		Slight increase	• Cluster peak lowered in freq. to ~ 2500 Hz; • Decrease of energy 2600-2900 Hz	• dB increase; • Statistically significant at 2500-2700, 2700-3000 Hz • Cluster observed with peak possibly lower in freq.	dB increase Clusters observed Lowering of cluster peak freq.
	3000-3500 Hz		Slight increase	• dB increase; • statistically significant 3200-3500 Hz but still valley	Statistically significant dB increase – although still a clear valley	Valley
	3500-4000 Hz		Slight increase	• dB increase; • statistically significant 3500-3800; • cluster peak moved down to lower freq.	• dB increase; • cluster observed with freq. peak possibly lower	dB increase Clusters observed Lowering of cluster peak freq.
	4000-4500 Hz		Slight increase	Slight dB increase	Statistically significant dB increase	dB increase Clusters observed Lowering of cluster peak freq.
	4500-5000 Hz		Slight increase	Energy decrease	Statistically significant dB increase	

Table 4: English Phrases and Call words Mode. A comparative table providing results that demonstrate tendencies.

Modes	Analysis		Control group -No Training-	Test group V -14 contact hours-	Test group S -28 contact hours-	American test group 6 week intensive workshop
English Phrases And Call words	F0	Hz	same	same	same	same
		dB	Slight dB increase	Slight dB increase	Slight dB increase	Same or decrease
	F1	Hz	same	same	same	same
		dB	Slight increase	same	increase	Same or increase
	F0-2500 Hz		Very small increase	• Mostly a decrease of energy; • Statistically significant increase at 1500-1700 Hz;	• Energy increase at 400-500 and 1000-2500 Hz; • with statistically significant increases in 1000-1100, 1200-1400,1500-1600, 2200-2500 Hz	Due to system of analysis no specifics can be provided
	2500-3000 Hz		Slight increase	Decrease of energy; Cluster peak lower in freq.	• Statistically significant dB increase; • More pronounced cluster	Clusters observed Sometimes dB increase
	3000-3500 Hz		Slight increase	Decrease of energy	Statistically significant dB increase	valley
	3500-4000 Hz		Slight increase	Decrease in energy; Move towards cluster	• Statistically significant dB increase; • More pronounced cluster	DB increase Clusters observed
	4000-4500 Hz		Slight increase	Strong decrease in energy	Statistically significant dB increase	Increased dB sometimes – with clustering
	4500-5000 Hz			Strong decrease in energy	Energy increase	Increased dB sometimes – with clustering

Table 5: First language texts Mode. A comparative table providing results that demonstrate tendencies.

Modes	**Analysis**		**Control group** **-No Training-**	**Test group V** **-14 contact hours-**	**Test group S** **-28 contact hours-**	**American test group** **6 week intensive workshop**
First Language Texts	F0	Hz	n.a	same	same	n.a
		dB	n.a	Slight increase	Slight increase	n.a
	F1	Hz	n.a	same	same	n.a
		dB	n.a	Slight decrease	same	n.a
	F0-2500 Hz		n.a	• dB increase 600-1400, 1500-2100,2300-2500 Hz; • Statistically significant at 700 –1400, 1500-2000 Hz	• Basically same visual pattern; • DB increase 1000-2500 Hz; • all but 1100-1200 Hz statistically significant	n.a
	2500-3000 Hz		n.a	• Statistically significant dB increase at 2500-2600 Hz • Decrease in energy	• Statistically significant dB increase; • Cluster observed	n.a
	3000-3500 Hz		n.a	Decrease in energy	Statistically significant dB increase	n.a
	3500-4000 Hz		n.a	Decrease in energy	Statistically significant dB increase;	n.a
	4000-4500 Hz		n.a	Strong decrease in energy	• Statistically significant dB increase up to 4400 Hz; • dB increase to 4500 Hz	n.a
	4500-5000 Hz		n.a	Strong decrease in energy	• DB increase; • Statistically significant increase 4700-4800 Hz	n.a

Table 6: Most consistent changes observed in all training groups for all modes.

Analysis		**Y-buzz**	**+Y-buzz**	**Calls**	**English Phrases and Call words**	**First language texts** *Only referring to Test groups V and S as a continuum of the training process*
F0	Hz	Same	Same	same	same	same
	dB	Increase	Same or increase	increase	Mostly increase	Slight increase
F1	Hz	same	Same	same	same	same
	dB	increase	Same or increase	increase	Same or increase	Same or decrease
F0-2500 Hz		• Profile change; • Often statistically significant	• Profile change; • Often statistically significant	• Profile change; • Often statistically significant	• Profile change; Often statistically significant	Basically same
2500-3000 Hz		Energy increase	• Energy increase; • Cluster; • Lowering of peak freq.	• Energy increase; • Cluster; • Lowering of peak freq.	• Cluster • Mostly increase • Lowering of peak freq.	Initially increase in small window, rest decrease then increase and cluster follow
3000-3500 Hz		• Some reflected energy increase; • Valley observed	• Some reflected energy increase; • Valley observed	• Some reflected energy increase; • Valley observed	• Some reflected energy increase; • Valley observed	Initially decrease, later increase but still valley
3500-4000 Hz		Energy increase	• Energy increase; • Cluster; • Lowering of peak freq.	• Energy increase; • Cluster; • Lowering of peak freq.	• Cluster • Mostly increase of energy	Initially increase in small window, rest decrease then increase follows
4000-4500 Hz		• Test group V decrease; • Test group S some increases; • American group energy increase;	Energy increase or same	• Energy increase; • Sometimes clustering	• Test group V energy decrease • Test group S statistically significant increase • American test group energy increase, sometimes with clustering	Initially increase in small window, rest decrease then increase follows
4500-5000 Hz		• Test groups V and S decrease; • Some in American group increased.	Energy increase or same	• Test group V energy decrease • Test group S statistically significant increase	• Test group V energy decrease • Test group S increase • American test group energy increase	• Test group V energy decrease • Test group S increase

By extracting information from Table 6 an overall conclusion can be made that the Tonal NRG of the Lessac Approach leads to:

- an increase in energy of F0;
- an increase in energy of F1 – should the duration of training be sufficient - which indicates the tuning of F1;
- a profile change of the LTAS F0-2500 Hz, which possibly indicates improved pronunciation or resonance;
- an energy increase in the 2500-3000 Hz window with a tendency to cluster and to lower the cluster peak frequency which may be due to the forward facial orientation;
- a valley in the 3000-3500 Hz window, although an increase of energy may take place;
- a tendency to cluster in the 3500-4000 Hz window, with an energy increase and a lowering of peak frequency should the duration of training be sufficient;
- a tendency to cluster in the 4000-4500 Hz window, with an energy increase and a lowering of peak frequency should the duration of training be sufficient. Speculatively, this researcher is of the opinion that this may happen only when the changes in the lower frequency windows have happened;
- in most cases a decrease of energy in the 4500-5000 Hz window but this may change after an intensive training process.

The changes above 2500 Hz will be in line with the threshold of hearing for humans. It may be due to the sensory awareness of the bone conduction on the hard palate and other facial bones. The awareness and active seeking of the bone conduction may lead to an involuntary shaping of the air resonator and, as Raphael and Scherer (1987) mention, it may have an effect on the stiffness of the pharyngeal walls. Another effect due to this awareness, that is framed by Verdolini (2002) as a recall and recognising schema (see Schmidt, 1992), may be the release of the larynx into a relatively low position with some supraglottal widening. These speculations definitely lead the way to possible further investigations.

In contradiction to these changes observed in the various training groups and reported on in Table 6, the second recordings of the Control Group primarily reflected a slight increase over the F0-5000 Hz spectrum, but no definite profile changes in the LTAS. Speculatively, it can be argued that these changes are due to increased breath pressure, as the aim of these untrained participants was mainly to "speak louder." No involuntary shaping of the resonator happened that could lead to an increased carrying power of the voice. These participants had no familiar event to model their sensation of projection on.

This study strongly indicates that the Tonal NRG of the Lessac Approach influences the acoustic profile of the female voice. It is necessary to draw from the literature discussed in Chapters 2 and 3 to compare whether the Lessac Approach positively reflects any of the existing theories about projection of the voice, and specifically of the female voice. The findings of this study will have to be compared to previous findings such as the "actor's formant" in general, and in particular to the outcomes of previous research addressing the effects of the Lessac Approach. For the comparison with existing literature, the three different means of voice analysis, as suggested by Laukkanen (1995:13), will be followed: The acoustical, perceptual and physiological.

5.3.2. Acoustic findings in comparison to existing literature

The singer's formant as defined by Sundberg (1974, 1987) is a cluster around 2,8 kHz observable in the LTAS of the western male opera singer. This phenomenon is also observed in the alto voice, and clearly higher partials in the 2-4 kHz window are observed in the soprano voice. The first tendency to cluster and increase energy in the female actor's voice in this study has been in the 2500-3000Hz (2,5-3kHz) window. Seeing that no attempt was made to divide the participants in this study into voice types, the fact that some had clusters in a slightly higher, and others in a slightly lower frequency window may be due to the fact that some of these participants were soprano and others alto speakers. This may open an area of further research. Although this study focuses on the female voice, the argument can be made that the Lessac Approach boosts an energy increase and clustering in the 2500-3000Hz window in the speaking voice seeing that the female speaking voice happens in the lower third of the possible female range and this can be comparable to the range of the male singing voice. To state this without doubt will only be possible after further research.

Laukkanen (1995:14) reminds that the increased ability to transfer sound can be visually observed in the LTAS as a spectral slope that is less steep. This is observed in the post-training recordings in this study for all the Lessac explorations and for the applications of the groups with 28 and more hours of training. This certainly points directly to curriculum building for voice, as the neuro-muscular repatterning needs to, in the holistic fashion, become the familiar event. A less steep slope can also be observed in the explorations modes for Test Group V that only had 14 contact hours.

The actor's formant, as defined by Leino (1993, 1995), is a cluster around 3,5 kHz observable in the LTAS of the Finnish male actor. The good voices, as categorized by a perception panel consisting of voice specialists,[3] in his study all had the clear cluster, as well as an energy increase. This is an important observation, as the study with Test Group V indicates that after only 14 hours of contact sessions the participants' post-training recordings often indicated a tendency to cluster but not yet an increase in energy. In comparison to this, the study with Test Group S indicated a tendency to cluster as well as an energy increase in the relevant regions. In comparison to Leino's findings with the male Finnish voice, this study indicates a tendency to cluster in the 3500-4000 Hz region. This cluster is however not as strong in dB as the one observed in the 2500-3000 Hz window.

Leino also commented on the fact that a peak can be too strong and, as indicated by Sundberg and Gauffin (1978), this can be a result of "pressed voice." This was never observed in this study. This may possibly be due to the fact that the Lessac Approach, as indicated by Verdolini (1998) and discussed in Chapter 2 of this study, leads to healthy vocal fold functioning. This will be discussed further on in this comparative analysis. In line with Leino's findings, a clear valley was always observed in the 3000-3500 Hz window, even if an increase in energy took place in that region.

Nawka et al. (1997) observed a cluster around the 3400 Hz peak in the voices of German males. This is a bit lower than what Leino observed and as such may support the argument of Weiss (1993) when he warns against setting up acoustic guidelines for preferred voice

---

[3] In this sense, the design of Leino's research is echoed in this study.

qualities. However, I want to argue that there is a difference between setting up guidelines that can act as ballpark windows for clustering, and narrowing it down to an exact 100 Hz window. In the same way as different voice types have clusters around a slightly higher or slightly lower frequency, different languages will have different resonance patterns. The higher singing voice has a tendency to cluster around slightly higher frequencies than the lower singing voice and the indication is there that this will also reflect in the speaking voice. The argument that the human voice has uniformity *as well as* uniqueness must hold true and be applied in any study of voice.

Direct acoustic analyses of the Lessac Approach have been done by Raphael & Scherer (1987) and Acker (1987). Acker (1987) specifically refers to a "ring formation"[4] with overtones clustered around 1300, 2000 and ~3600 Hz. This is very compatible to the findings in this study. Raphael and Scherer (1987) found that the F1 of the Call was tuned more sharply in speech mode as well as an increase at ~2000 –2500Hz and a decrease of energy at ~3000 Hz. This reflects the profile observed in this study to a certain extent, as Test Group V showed a decrease of energy in the Lessac applications from ~3000Hz onwards. This leaves the question that this may be a definite occurrence in the beginning of the training process where the increase over the higher spectrum happens later in the teaching process. Yet again, this is an area where further research can take place.

Linking to Leino's research, Munro et al. (1996) found that the Lessac Approach increased clustering in the Afrikaans male voice, in comparison to the untrained male voice, in the 3-4kHz region as well as a less steep spectral slope. This compares rather favourably with the findings in this study as it indicates a cluster in the 3500-4000 Hz window.

Leino (2001) did an in-depth study on the Finnish female voice and reports the actor's formant to be around 4,3kHz. This study reflects an interesting profile when taking the various groups into account. The group with the fewest contact hours indicated a decrease in energy in this region – especially when reflecting on the Lessac applications. Speculatively, this may be due to a shift in the use of the shape of the resonator but at a stage where the motor learning is still taking place. The groups with more training reflect

---

[4] In the light of Leino's research this can be read as "actor's formant."

an increase of energy in this area and sometimes a cluster. For this particular study it can be argued that the energy increase or tendency to cluster happens in the lower spectral regions first, and, in successive fashion, the adjustments in the resonator that effect the higher spectral regions follow. As suggested before, this is a possible area for further research. A speculative conclusion that can be mentioned is that the difference observed between the findings of this study and Leino's is, (drawing from Weiss' argument) that the difference is language specific. This is a difficult argument to make in this regard as the studies done with test groups V and S had various first language speakers participating. Again, this has to lead to further research as these findings can be coloured by the fact that the samples have been grouped together for analysis.

It can be concluded that although the findings of this study did not acoustically match the findings reported in current literature exactly, there is enough of a positive comparison to state that the Tonal NRG of the Lessac Approach does contribute to the projection of the female voice.

5.3.3. Perceptual findings and existing literature

Leino (1993,1995, 2001) relies on the perceptual analysis of his studies to find tendencies in the acoustical data and as such the importance of the perceptual analysis is fore grounded. It is after scores allotted, by a specialist perception panel, that he established the acoustic profile that lead to the concept of the "actor's formant." By interpretation then, it can be deduced from the work of Leino, that a specialist perception panel is capable of differentiating between very good, good, average and poor performers' voices and that the preference of the perception panel will be reflected in the acoustic findings. After all, it is still the human ear that has to determine the accepted quality to the voices of the actors, and specifically to this study, the female actor's voice.

Although only Test groups S and V had perceptual analyses as part of the research design, it is necessary to reflect in short on these findings. As discussed in Chapter 4, Sections 3 and 4 the perception panellists preferred the sound quality of the post-training recordings for all the various modes. All the raters for the Lessac explorations as well as Lessac applications, indicated significant improvement in the post-training recordings when tested against the pre-training recordings. In a paired t-test done on the raters for both Test Group S and Test Group V, raters 1 to 5 all had a p-value of <.0001 which indicates a very significant improvement of ratings in the post-training recordings.[5] This indicates that the group had a homogenous preference for the sound contained in the post-training recordings. For the application mode, four of the raters again had a p-value of <.0001 but one rater (rater 4) had a p-value of 0.0002. Although this difference between the four raters who all had the same p-value, and the rater who had the p-value of 0.0002, has to be observed, the p-value of all five raters individually, is still smaller than Alpha (α) and as such still statistically very significant. This indicates an inter-reliability in the rater population and as such the perception panel can be regarded as a homogenous group of experts in the field of voice.

As mentioned before, it is interesting to note that although the applications (sound modes 4 and 5) of Test Group V (14 contact hours) did not improve acoustically to the same extent as those in Test Group S (28 contact hours), the raters indicated a clear difference in the pre- and post-training recordings of both groups. It can be speculated that the raters reacted to the changes in the lower range of the spectrum (F0-2500Hz), or that their ears are trained to such an extent that they could differentiate the smallest possible improvements. This will have to be addressed in a further research project. The fact that these two groups were trained by the same Lessac teacher probably indicates that teacher specificity does not play a vital role in the Approach but that time allotted for the re-patterning process of motor leaning is important. This will have to be addressed in the compilation of any voice building curriculum.

5.3.4. Physiological findings and existing literature

Although this study did not concern itself directly with a physiological investigation about the Tonal NRG of the Lessac Approach, certain observations can be made about possible changes in the physiological functioning of the human voice when applying the concepts of the Tonal NRG. A change in an acoustic profile implies a change in physiological

---

[5] Where $\alpha = 0.05$.

functioning. An increased energy in F0 indicates increased breath pressure. F0 stayed the same for most sound samples in the pre- and post-training recordings. When F0 dropped one can argue a relaxed larynx and when F0 rose it can possibly be attributed to the fact that the larynx was pushed down during the pre-training recording. But this is mere speculation and may lead to further research. The sometimes observed increase of energy in the F1 region implies better formant tuning as indicated by Raphael and Scherer (1987). When a lowering of F1 was sometimes observed, it implies the decrease of the mouth opening and an increase of the protrusion of the lips. This correlates with Lessac's idea of the "forward facial orientation" as well as the "inverted megaphone." This would reflect the orientation to the German style of classical singing.

A less steep spectral slope implies an increased ability of the vocal tract to transfer sound frequencies. This indicates the shaping of the resonator. The lowering of the peak frequencies of the clusters implies the use of a longer resonator – again reflecting on the forward facial orientation.

When referring to the research that Verdolini and her colleagues have done on the Lessac Approach and specifically the "Lessac-Madsen Resonant Voice Therapy" it is eminent that the Lessac Approach contributes to the vocal health in a physiologically sound manner where the voice is produced "...with a laryngeal configuration that maximizes voice output, keeps vocal fold impact stress small, and minimizes the subglottic pressure required for phonation" (Verdolini, 2002, personal e-mail communication). This supports the definition of good voice quality used in this study where Laukkanen (1995:18) posits that it is the "optimal use of the vocal organ in order to establish the maximum possible acoustic output by minimal muscular effort."

5.4. Conclusion

The question was asked in Chapter 1 whether the Tonal NRG of the Lessac Approach enhances the projection of the female voice, irrespective of

- Language or
- Teacher and training methodology specificity.

From the above reading there is clear indications that the Tonal NRG of the Lessac Approach does enhance the projection of the female voice and it is strongly compatible to existing literature about the acoustic profile of the well-projected voice.

The language aspects of the study did not come to full fruition, as the groups were too small for any statistically significant data. Yet, the interpretations made in Chapter 4, Sections 3 and 4 did indicate that the voices of all the participants, regardless of language, did improve with the training as analysed perceptually and acoustically. This bodes well for the use of the Lessac Approach in a multi-lingual training situation as experienced in South Africa.

The potential bias that may have arisen from a particular, idiosyncratic teacher and training methodology did not seem to have a major impact. What was clearly indicated was the necessary and required time that needed to be allotted for the training. There was a definite build acoustically in the profiles of the groups with more training. The argument can possibly be made that with enough time for training and thus time for repatterning of habitual motor responses, sustainability will occur. This will have to be kept in mind if this Approach is used in any curriculum for voice training for actors. The sustainability of the impact of the Lessac Approach opens the door, once again, to further research.

Further speculation may have it that (seeing that Test Group V, with only 14 contact hours demonstrated less carry over of the approach to the applications than Test Group S with 28 contact hours) the training that Test Group S experienced will lead to more sustainability than the training that Test Group V experienced. This again leads to another possible topic for further research. The concept of sustainability and economy is within the South African educational system of primary concern.

As long as the teacher is certified as a Lessac teacher in order to guarantee quality, the small differences in the approach that different teachers may have will not play a definite role in the building of the voice for theatre. This conclusion supports Hanson's view (1997:111) that the teacher needs to be well trained, but, whereas Hanson seems to see this as a stumbling block in training, it may in effect be seen as a vital and developmental attribute in the process. Enough evidence exists to support the views of Chabora (1994:114), Raphael (1997:209) and Tabish (1995:138) when they hail the Lessac Approach as a holistic and innovative teaching tool for building the performer's voice. This study indicates very strongly that the voice quality and carrying power of the voice, and in particular the female voice, improve through the use of the Tonal NRG of the Lessac Approach in a healthy manner.

5.5. Shortfalls of this study

This study has many shortfalls that should be improved on in future research about the Tonal NRG of the Lessac Approach.

- The study would have benefited from the possibility of having more countries included in this design. Unfortunately the Lessac teachers are still mostly Americans, and the German contingent pulled out of this project at a very late stage of the research design.
- The study would have benefited from the possibility of having the same technology used for pre- and post-training recordings on all the occasions. Direct comparisons would have been feasible. Financial constrains did not make this possible.
- The study would have benefited from the possibility that all the pre- and post-training recordings could have been done in the same location. Financial constrains do not make this possible.
- The study would have benefited from the possibility of having more Certified Lessac Teachers involved in the teaching of the South African classes. Unfortunately there are no other Certified Lessac teachers, at the time of writing here, in South Africa, yet.
- The study would have benefited from the possibility of having each language group in a separate class and instructed by a Certified Lessac Teacher who share the first language of that specific group. Other than the Americans and the one South African who are certified, there is only a Japanese Certified Lessac Teacher working in Japan.
- This study would have benefited from longer time samples of the different modes to use for the LTAS analysis. The reason for the short samples is that most of the participants in the groups have no previous experience of any voice training and seemed to be vocally so unfit that longer time samples in the pre-training recordings were not possible. The

subjective observation that the participants were capable of providing longer time samples during the post-training recordings demonstrates the positive dimension of the outcome of this study.

- A bigger control group and certainly a control group who had a different system of training would have been an advantage to this study. Reasons for the small control group have already been discussed in Chapter 4, section 2.

Despite all of the above-mentioned shortfalls of this study, it indicates a strong tendency about the positive influence of Lessac's Tonal NRG on the female voice. The female voices in Test Group S and V were perceived as of better quality and with improved projection capabilities in the post-training recordings. The acoustic profile of the female voices changed, when compared to existing literature, for the better. On a physiological level, strong evidence exists that the use of the Tonal NRG as a voice building tool leads to healthy and effective vocal fold functioning. Furthermore there are indications that the Tonal NRG is not bound to a specific language, nor to a teacher or training methodology.

The value of the Lessac Approach is uncontestable. The possible areas of research indicated in this study will hopefully pave the way to more research about this very valuable voice building approach.

BIBLIOGRAPHY.

ACKER, B. 1987. Vocal Tract Adjustments for the Projected Voice. In: Journal of Voice Vol. 1, No 1:77-82.

BARRICHELLO, O; **HEUER**, R.I; **DEAN**, C.M & **SATALOFF**, R.T. 2001. Comparison of Singer's Formant, Speaker's Ring, and LTA spectrum Among Classical Singers and Untrained Normal Speakers. In: Journal of Voice. Vol.15. No 3. pp. 344-350.

BARTON, R. & **DAL VERA**, R. 1995. *Voice: Onstage and Off.* Fort Worth: Harcourt Brace College Publishers.

BERRY, C. 1973. *Voice and the Actor.* London: Harrap.

BERRY, C. 1987. *The actor and his text.* NY: Charles Schribner's Sons.

BERRY, C. 1997. That Secret Voice. In: Vocal Vision. Eds. Hampton, M.,Acker, B. NY:Applause Books. pp. 25-36

BERRY, D.A., **VERDOLINI**, K., **MONTEQUIN**, D., **HESS**, M., **CHAN**, R.W. & **TITZE**, I.R. 2001. Indications of an optimum glottal width in vocal production. In: Journal of Speech and Hearing Research. No. 4, pp 29-37

BESTERBEURTJE, M.E & **SCHUTTE**, H.K. 2000. Resonance Strategies for Belting Style: Results of a Single Female Subject Study. In: Journal of Voice. Vol. 14. no. 2, pp. 194-204

BORDEN, G.J., **HARRIS**, K.S. & **RAPHAEL**, L.J. 1984. *Speech Science Primer: Physiology, Acoustics and Perception of Speech.* Philadelphia: Lippencott Williams & Wilkins.

BLOOTHOOFT, G & **PLOMP**, R. 1986. The sound level of the singer's formant in professional singing. In: Journal of the Acoustical Society of America. 1986:79(6):2029-2033)

CHABORA, P.D. 1994. *A Descriptive Study of the Application of Research in Neurophysiology to Self-use Training for Actors.* Unpublished dissertation. UMI Dissertation Services.

COFFIN, B. 1976. *The Sounds of Singing.* Boulder, Colorado: Pruett Publishing Co.

COFFIN, B. 1980. *Overtones of Bel Canto.* Boulder, Colorado: Pruett Publishing Co.

COHEN, L. & **MANION**, L. 1997 (4th ed. reprint). *Research Methods in Education.* London: Routledge.

DAL VERA, R. (ed). 2000. *Standard Speech and other Contemporary Issues in Professional Voice and Speech Training.* New York: Applause.

DMITRIEV, L & **KISELEV**, A. 1979. Relationship between the Formant Structure of Different Voice Types of Singing Voices and the Dimensions of Supraglottic Cavities. In: Folia Phoniatrica. 31:238-241.

ESTILL, J. 1980. Observations about the Quality called "Belting". In: Transcripts of the 9th Symposium: Care of the Professional Voice, Part II: 82-88 (1980)

ESTILL, J. 1984. Excerpt from a letter to Donal Henahan, Music Critic, New York Times.

ESTILL, J., **Baer**, T., **Honda**, K. & **Harris**, K.S. 1985. The Control of Pitch and Quality, Part II: An EMG study of infrahyoid muscle. In: V. Lawrence (ed), Transcripts of the 13th Symposium, 1985. The Voice Foundation. New York.

ESTILL, J. 1986. An EMG Study of Two Voice Qualities: Nasal Twang and Opera with "Squillo". In: Proceedings of International Conference on Voice. A satellite meeting of XXth Congress of IALP, July 28029, 1986. Kurume, Japan.

FANT, G. 1970. *Acoustic Theory of Speech Production.* Paris, France: Mouton.

GARBERS, J. (red). 1996. *Doeltreffende geesteswetenskaplike navorsing.* Pretoria: J.L van Schaik.

GAUFFIN, J & **SUNDBERG**, J. 1989. Spectral Correlates of Glottal Voice Source Waveform Characteristics. In: Journal of Speech and Hearing Research. Vol. 32, 556-565, September 1989.

HAMPTON, M & **ACKER**, B. (eds.). 1997. *The Vocal Vision: Views on Voice.* New York: Applause.

HANSON, A.M. 1997. *An analysis of the Physiological Assumptions in Vocal Instructional Systems for actors.* Unpublished PhD. Dissertaion. University of Kansas. UMI Number: 9827481

INTERNATIONAL PHONETIC ASSOCIATION. 1999. *Handbook of the International Phonetic Association.* New York, Cambridge: University Press.

KENT, R.D. 1997. *The Speech Sciences.* San Diego: Singular Press.

KENT, R.D. & **READ**, C. 1992. *The Acoustic Analysis of Speech.* San Diego: Singular Press.

KIUKAANNIEMI, H; **SIPONEN**, P. & **MATTILLA**, P. 1982. Individual Differences in the Long-Term Speech Spectrum. In: Folia Phoniatrica. 34: 21-28.

LAUKKANEN, A-M. 1995. *On Speaking Voice Exercises.* Academic dissertation. ACTA Universitatis Tamperensis. Ser. A vol. 445. Tampere: University of Tampere.

LAUKKANEN, A-M. 2002. On the Goals and Bases of Vocal Exercises. (Submitted for publication) Voice and Speech Review.

LAUKKANEN, A-M. 1999-2002. Personal e-mails. In possession of the researcher.

LEEDY, P & **ORMROD**, J.E. 2001, 7th edition. *Practical research: Planning and design.* New Jersey: Prentice Hall.

LEEDY, P. 1981. *How to read research and understand it.* New York: Macmillan Publishing Co., Inc.

LEINO, T. 1993. Long-term average spectrum study on speaking voice quality in male actors. In: Proceedings of the Stockholm Music Acoustics Conference, Jul.28 - Aug.1, 1993 (Eds: A. Friberg, J.Iwarsson, E.Jansson & J. Sundberg).

LEINO, T. & **KÄRKKÄINEN**, P. 1995. On the effects of vocal training on the speaking voice quality of male student actors. In: Proceedings of the XIIIth International Congress of Phonetic Sciences. Stockholm, Sweden, 13 - 19 Aug. 1995. (Eds: K. Elenius & P. Branderud).

LEINO, T. 2001. Voice Quality Of Finnish Female Actors. Poster presented at PEVOC, Stockholm 2001.

LESSAC, A. 1960. 1st edition. *The Use and Training of the Human Voice.* New York: Drama Book Publishers.

LESSAC, A. 1967. 2nd edition. *The Use and Training of the Human Voice.* New York: Drama Book Publishers.

LESSAC, A. 1997a. 3rd Edition. *The Use and Training of the Human Voice: a Bio-Dynamic Approach to Vocal Life.* California: Mayfield.

LESSAC, A. 1990. 2nd Edition. *Body Wisdom: The Use and Training of the Human Body.* California: L.I.P. CO.

LESSAC, A. 1997b. From Beyond Wildness to Body Wisdom, Vocal Life, and Healthful Functioning: A Joyous Struggle for Our Discipline's Future. In: The Vocal Vision: Views on Voice. 1997:13-24. Eds: Hampton & Acker.

LESSAC, A. 1998. A Short Colloquy of How I Approach Pathological and Functional Voice and Speech Disorders through my Kinesensic Training and Research. In: Special Interest Division 3, Voice and Voice Disorders Newsletter, Vol. 8, No. 1 April 1998 (np.).

LESSAC, A. 1999-2002. Personal e-mails. In possession of the researcher.

LESSAC, A. 2000. Lessac Kinesensics: A Short Treatise of some of its Vocal Life aspects. In: *Standard Speech and other Contemporary Issues in Professional Voice and Speech Training.* Dal Vera, R (ed). New York: Applause. pp 191-192.

LESSAC, A. 2000-2001. Personal discussions.

LINKLATER, K. 1976. Freeing the Natural Voice. New York: Drama Book Publishers.

LINKLATER, K. 1997. Thoughts on Theatre, Therapy and the Art of Voice. In: Vocal Vision. Eds. Hampton, M. & Acker, B. NY: Applause Books. pp. 3-12.

LUNDY, D.S., **ROY**, S., **ROY**, C.R., **XUE**, J.W & **EVANS**, J. 2000. Acoustic Analysis of the Singing and Speaking Voice in Singing Students. In: Journal of Voice. Vol. 14. No. 4. pp. 490-493 (2000).

MARTIN, J.1991. *Voice in Modern Theatre.* London: Routledge.

MULHOLLAND, M. 1984. *Speech Training for Actors in Twentieth Century America: change and continuity.* Unpublished PhD dissertation: University Microfilms International.

MCKINNEY, J. 1982. *The Diagnosis and Correction of Vocal Faults.* Nashville, Tennessee: Broadman Press.

MENDOZA, E., **VALENCIA**, N., **MUNOZ**, J., & **TRUJILLO**, H. 1996. Differences in Voice Quality Between Men and Women: Use of the Long-term Average Spectrum (LTAS). In: Journal of Voice. Vol. 10. No. 1. pp. 59-66 (1996).

MILLER, D. 1999. The Use of the electroglottograph in the Voice Studio. In: Voice-tradition and Technology. Nair, G. pp. 211-226.

MILLER, D. 2000. *Registers in Singing: Empirical Studies and Systematic Studies in the Theory of the Singing Voice.* Groningen: Rijksuniversiteit – Proefschrift.

MILLER, R. 1986. *The Structure of Singing: System and Art in Vocal Technique.* New York: Schrimer Books.

MOUTON, J. 2001. *How to succeed in your Master's & Doctoral studies. A South African guide and resource book.* Pretoria: Van Schaik.

MUNRO, M & **LAUKKANEN**, A-M. 2001. On The Effects Of Lessac Method On Female Voices: Preliminary observations. Poster presented at PEVOC, Stockholm, 2001.

MUNRO, M., **LEINO**, T., **WISSING**, D. 1996. Lessac's Y-buzz as a pedagogical tool in the teaching of the Projection of an Actor's Voice. In: South African Journal of Linguistics. Suppl.34, Dec. 1996. pp. 25-36.

MUNRO, M. & **WISSING**, D. 1999. Teaching Voice in a Multi-Lingual Set-up: Some ideas for the Future. In: Voice and Speech Trainer Association, Inc. Newsletter. Spring/Summer 1999. Vol. 13, no.2.

NAIR, G. 1999. *Voice Tradition and Technology.* New York: Singular Press.

NAWKA, T., **ANDERS**, L.C., **CEBULLA**, M., **ZURAKOWSKI**, D. 1997. The Speaker's Formant in Male Voices. In: Journal of Voice. Vol 11. No. 4, pp. 422-428. 1997.

PERKINS, W. & **KENT**, R. 1986. *Functional Anatomy of Speech, Hearing and Language.* Boston: Allyn & Bacon.

PETERSON, K.L., **VERDOLONI-MARSTON**, K., **BARKMEIER**, J.M & **HOFFMAN**, H.T. 1994. Comparison of Aerodynamic and Electroglottographic Parameters in Evaluating Clinically Relevant Voicing Patterns. In: Ann Otol Rhinol Loaryngol 103: 1994 pp. 335-346.

RAPHAEL, B. & **SCHERER**, R. 1987. Voice modification for stage actors: Acoustic analysis. In: Journal of Voice, Vol. 1, No. 183-87.

RAPHAEL, B. 1997. A Consumer's Guide to Voice and Speech Training. In: Vocal Vision. Eds. Hampton, M., Acker, B. NY: Applause Books. pp. 203-214.

RODENBURG, P. 1992. *The Right to Speak: Working with the Voice.* London: Methuen Drama.

RODENBURG, P. 1993. *The Need for Words.* London: Methuen Drama.

ROSE, A. 1978. *The Singer and the Voice: Vocal Physiology and Technique for Singers.* London: Scholar Press.

ROTHENBERG, M. 1981. An Interactive Model for the Voice Source. STL-QPSR 4.

ROTHENBURG, M. 1988. Acoustic Reinforcement of Vocal Fold Vibratory behaviour in Singing. In: Vocal Fold Physiology. Vol.2 pp. 379-398. Fujimura, O. (ed.) New York: Raven Press.

SATALOFF, R.T. 1991. *Professional Voice. The Science and Art of Clinical Care.* New

York: Raven Press.

SCHUTTE, H & **MILLER**, D. 1993. Belting and Pop, Nonclassical Approaches to the Female Middle Voice: Some Preliminary Considerations. In: Journal of Voice. Vol. 7, No. 2, pp. 142-150.

SCHMIDT, R. 1991. *Motor Learning and Performance: From Principle to Practice.* Champagne, ILL: Human Kinetics Books.

STANISLAVSKI, K. 1980 (paperback ed.). *An Actor Prepares.* London: Eyre Methuen.

STEYN, A.G.W., **SMIT**, C.F., **DU TOIT**, S.H.C & **STRASSHEIM**, C. 2000, 4th Impression. *Modern Statistics in Practice.* Pretoria: J.L van Schaik, Academic.

STONE, R.E. (ed); **CLEVELAND**, T.F & **SUNDBERG**, J. 1999. Formant Frequencies in Country Singers' Speech and Singing. In: Journal of Voice. Vol. 13. No. 2, pp. 161-167.

SUNDBERG, J. & **GAUFFIN**, J. 1978. Waveform and Spectrum of the Glottal Voice Source. In: Speech Transmission Laboratory, Quarterly status and progress report, 2-3/1978, 35-50. Stockholm: Royal Institute of Technology.

SUNDBERG, J. 1974. Articulatory Interpretation of the "Singing Formant". In: JASA. Vol. 55, No. 4. April 1974.

SUNDBERG, J. 1987. *The Science of the Singing Voice.* Dekalb: Northern Illinois Univ. Press.

SUNDBERG, J. 1988. Vocal Tract Resonance in Singing. In: The NATS Journal March/April 1988:11-31.

TABISH, D. R. 1995. *Kinesthetic Engagement Technique: Theories and Practices for Training the Actor.* Unpublished dissertation. UMI Dissertation services.

TITZE, I. 1994a. *Workshop on Acoustic Voice Analysis: Summary Statement.* National Center for Voice and Speech.

TITZE, I. 1994b. *Principles on Voice Production.* New Jersey: Prentice Hall.

TITZE, I. 1999. The use of low first formant vowels and nasals to train the lighter mechanism. In: Journal of Singing. March/April 1999. Vol. 55. No. 4. pp. 41-43.

THURMAN, L. & **WELCH**, G (eds). 1997. *Bodymind and Voice: Foundations of Voice education.* The VoiceCare Network; National Center of Voice & Speech; The Voice Center of Fairview, Fairview Arts Medicine Center.

VAN ROOY, B. & **VAN HUYSSTEEN**, G.B. 2000. The vowels of BSAE: current knowledge and future prospects. In: Supplement 38, SA Journal of Linguistics (pp. 15-43).

VENNARD, W. 1967. *The Mechanism and Technique*. New York: Carl Fischer.

VERDOLINI-MARSTON, K., **BURKE**, M.K., **LESSAC**, A., **GLAZE**, L. & **CALDWELL**, E. 1995. A Preliminary Study on two methods of treatment for Laryngeal Nodes. In: Journal of Voice. Vol. 9: 74-85.

VERDOLINI, K. 1997. Principles of skill acquisition: Implications for Voice Training. In: The Vocal Vision: Views on Voice. 1997:65-80. Eds: Hampton & Acker.

VERDOLINI, K., **DRUKER**, D.G., **PALMER**, P.M. & **SAMAWI**, H. 1998. Laryngeal adduction in resonant Voice. In: Journal of Voice.

VERDOLINI, K. 1998. Professional Voice Corner: Introductory Comments. In: Special Interest Division 3, Voice and Voice Disorders Newsletter. Vol. 8, No. 1 April 1998 (np.).

VERDOLINI, K. 1999-2002. Personal e-mails. In possession of the researcher.

VERDOLINI, K. 2002 draft (in preparation). Motor Learning Principles. In: I.R. Titze & K. Verdolini, Vocology. Iowa City: National Center for Voice and Speech.

WEDIN, S., **LEANDERSON**, R. & **WEDIN**, L. 1978. Evaluation of Voice Training: Spectral analysis Compared with Listeners' Judgements. In: Folia phoniat. 30: 103-112

WEISS, W. 1993. Comparison of Spectral and Voice Quality Parameters of Voice Trained Anglophones and Francophones in Continuous Speech. In: Canadian Acoustics 21(1) 3-8 (1993).

WULF, G., & **PRINZ**, W. 2001. Directing attention to movement effects enhances learning: A review. In: Psychonomic Bulletin & Review. 8 (4), 648-660 (2001).

http://www.ncvs.org/vpt/tutorial/filter.html

Addendum

Statistically significant improvements of the mean for the various 100Hz windows as indicated on the Bar graphs in Chapter 4, Section 3 and Section 4.

For each 100Hz window that indicates a statistically significant improvement the Shapiro Wilk test for normality is provided. Data is accepted as normally distributed when the p-value is > 0.05. The t-test for independence is done when the data is normally distributed. The null-hypothesis for this t-test is Ho:$\mu=0$. This is rejected when the p-value of the t-test is ≤0.1 and the t-value is positive. The alternative of $\mu>0$ is then accepted (p-value <0.1) For the p-value to be compared with alpha (*α*) the p-value has to be divided by 2 seeing that this study focuses on the one sided testing. Alpha (α) in this case is thus 0.05.

Test Group S

Test Group S: Y-buzz

F0.
Mean: 8.279

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.931933	**Pr < W**	0.2916

Test	Statistic		p Value	
Student's t	**t**	4.484744	**Pr > |t|**	0.0005

200-300 Hz
Mean: 8.0979

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.922726	**Pr < W**	0.2406

Test	Statistic		p Value	
Student's t	**t**	3.816765	**Pr > |t|**	0.0021

1800-1900 Hz
Mean: 4.587

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.959903	**Pr < W**	0.6907

Test	Statistic		p Value	
Student's t	**t**	1.971647	**Pr > |t|**	0.0687

1900-2000 Hz
Mean: 8.210

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.976291	**Pr < W**	0.9379

Test	Statistic		p Value	
Student's t	**t**	2.39443	**Pr > |t|**	0.0312

2000-2100 Hz
Mean: 12.091

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.980837	**Pr < W**	0.9749

Test	Statistic		p Value	
Student's t	**t**	3.836732	**Pr > |t|**	0.0018

2100-2200 Hz
Mean: 7.097

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.983327	**Pr < W**	0.9872

Test	**Statistic**		**p Value**	
Student's t	**t**	2.231415	**Pr > |t|**	0.0425

2200-2300 Hz
Mean: 13.085

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.879632	**Pr < W**	0.0468

Test	**Statistic**		**p Value**	
Student's t	**t**	3.153928	**Pr > |t|**	0.0070

2300-2400 Hz
Mean: 6.077

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.901203	**Pr < W**	0.0993

Test	**Statistic**		**p Value**	
Student's t	**t**	1.903284	**Pr > |t|**	0.0778

2500-2600 Hz
Mean: 4.485

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.980984	**Pr < W**	0.9758

Test	Statistic		p Value	
Student's t	**t**	1.776257	**Pr > |t|**	0.0974

3600-3700 Hz
Mean: 5.167

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.946024	**Pr < W**	0.4641

Test	Statistic		p Value	
Student's t	**t**	1.855091	**Pr > |t|**	0.0848

3700-3800 Hz
Mean: 5.875

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.895115	**Pr < W**	0.0802

Test	Statistic		p Value	
Student's t	**t**	2.20688	**Pr > |t|**	0.0445

Test Group S: +Y-buzz

F0
Mean: 7.808

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.936594	Pr < W	0.3415

Test	Statistic		p Value	
Student's t	t	4.732244	Pr > |t|	0.0003

200-300 Hz
Mean: 6.156

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.965797	Pr < W	0.7917

Test	Statistic		p Value	
Student's t	t	3.147568	Pr > |t|	0.0071

1200-1300Hz
Mean: 4.549

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.973751	Pr < W	0.9092

Test	Statistic		p Value	
Student's t	t	2.009892	Pr > |t|	0.0641

1400-1500Hz
Mean: 5.725

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.924814	Pr < W	0.2280

Test	Statistic		p Value	
Student's t	t	3.276764	Pr > |t|	0.0055

1500-1600Hz
Mean: 1.991

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.88861	Pr < W	0.0639

Test	Statistic		p Value	
Student's t	t	0.81384	Pr > |t|	0.4294

1600-1700Hz
Mean: 7.025

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.948691	Pr < W	0.5041

Test	Statistic		p Value	
Student's t	t	3.93852	Pr > |t|	0.0015

1700-1800Hz
Mean: 9.315

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.811438	**Pr < W**	0.0052

Test	**Statistic**		**p Value**	
Student's t	**t**	2.983039	**Pr > |t|**	0.0099

1800-1900Hz
Mean: 8.433

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.904861	**Pr < W**	0.1129

Test	**Statistic**		**p Value**	
Student's t	**t**	4.626446	**Pr > |t|**	0.0004

1900-2000Hz
Mean: 9.039

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.930816	**Pr < W**	0.2806

Test	**Statistic**		**p Value**	
Student's t	**t**	3.007317	**Pr > |t|**	0.0094

2000-2100Hz
Mean: 6.774

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.96574	**Pr < W**	0.7907

Test	**Statistic**		**p Value**	
Student's t	**t**	2.4644	**Pr > |t|**	0.0273

2100-2200Hz
Mean: 7.097

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.983327	**Pr < W**	0.9872

Test	**Statistic**		**p Value**	
Student's t	**t**	2.231415	**Pr > |t|**	0.0425

2300-2400Hz
Mean: 6.077

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.901203	**Pr < W**	0.0993

Test	**Statistic**		**p Value**	
Student's t	**t**	1.903284	**Pr > |t|**	0.0778

2500-2600Hz
Mean: 4.485

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.980984	**Pr < W**	0.9758

Test	**Statistic**		**p Value**	
Student's t	**t**	1.776257	**Pr > |t|**	0.0974

2600-2700Hz
Mean: 7.774

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.945278	**Pr < W**	0.4534

Test	**Statistic**		**p Value**	
Student's t	**t**	2.933158	**Pr > |t|**	0.0109

2700-2800Hz
Mean: 5.639

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.930594	**Pr < W**	0.2785

Test	**Statistic**		**p Value**	
Student's t	**t**	3.178662	**Pr > |t|**	0.0067

2800-2900Hz
Mean: 5.081

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.957388	Pr < W	0.6471

Test	Statistic		p Value	
Student's t	t	2.254829	Pr > |t|	0.0407

3400-3500Hz
Mean: 4.811

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.969818	Pr < W	0.8553

Test	Statistic		p Value	
Student's t	t	2.277679	Pr > |t|	0.0390

3500-3600Hz
Mean: 4.701

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.952039	Pr < W	0.5571

Test	Statistic		p Value	
Student's t	t	2.334919	Pr > |t|	0.0350

3600-3700Hz
Mean: 8.703

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.982409	**Pr < W**	0.9833

Test	**Statistic**		**p Value**	
Student's t	**t**	3.524569	**Pr > |t|**	0.0034

3700-3800Hz
Mean: 8.196

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.986393	**Pr < W**	0.9958

Test	**Statistic**		**p Value**	
Student's t	**t**	3.889726	**Pr > |t|**	0.0016

3800-3900Hz
Mean: 5.078

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.91333	**Pr < W**	0.1523

Test	**Statistic**		**p Value**	
Student's t	**t**	1.814886	**Pr > |t|**	0.0910

4500-4600Hz
Mean: 5.960

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.886631	**Pr < W**	0.0596

Test	**Statistic**		**p Value**	
Student's t	**t**	2.369031	**Pr > |t|**	0.0328

4600-4700Hz
Mean: 5.851

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.946777	**Pr < W**	0.4752

Test	**Statistic**		**p Value**	
Student's t	**t**	2.147528	**Pr > |t|**	0.0497

Test Group S: Calls

500-600 Hz
Mean: 9.147

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.939194	**Pr < W**	0.4081

Test	**Statistic**		**p Value**	
Student's t	**t**	2.012111	**Pr > |t|**	0.0654

900-1000 Hz
Mean: -0.376

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.984995	**Pr < W**	0.9927

Test	**Statistic**		**p Value**	
Student's t	**t**	-0.1168	**Pr > |t|**	0.9087

1000-1100 Hz
Mean: 11.588

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.93619	**Pr < W**	0.3717

Test	**Statistic**		**p Value**	
Student's t	**t**	5.616096	**Pr > |t|**	<.0001

1100-1200 Hz
Mean: 7.756

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.879988	**Pr < W**	0.0581

Test	**Statistic**		**p Value**	
Student's t	**t**	2.834536	**Pr > |t|**	0.0141

1200-1300 Hz
Mean: 10.855

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.984263	**Pr < W**	0.9924

Test	Statistic		p Value	
Student's t	t	4.545813	Pr > |t|	0.0005

1300-1400 Hz
Mean: 7.246

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.941935	Pr < W	0.4437

Test	Statistic		p Value	
Student's t	t	2.36487	Pr > |t|	0.0343

2100-2200 Hz
Mean: 6.359

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.851067	Pr < W	0.0230

Test	Statistic		p Value	
Student's t	t	2.5423	Pr > |t|	0.0245

2200-2300 Hz
Mean: 5.064

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.908109	Pr < W	0.1479

Test	Statistic		p Value	
Student's t	t	2.845826	Pr > |t|	0.0138

2300-2400 Hz
Mean: 7.986

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.926682	Pr < W	0.2740

Test	Statistic		p Value	
Student's t	t	2.966277	Pr > |t|	0.0109

2400-2500 Hz
Mean: 6.061

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.904952	Pr < W	0.1331

Test	Statistic		p Value	
Student's t	t	2.922499	Pr > |t|	0.0119

2500-2600 Hz
Mean: 8.094

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.929152	Pr < W	0.2969

Test	Statistic		p Value	
Student's t	t	2.95223	Pr > |t|	0.0112

2600-2700 Hz
Mean: 7.655

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.95921	Pr < W	0.7101

Test	Statistic		p Value	
Student's t	t	2.64408	Pr > |t|	0.0202

2800-2900 Hz
Mean: 6.322

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.920704	Pr < W	0.2251

Test	Statistic		p Value	
Student's t	t	2.287821	Pr > |t|	0.0395

2900-3000 Hz
Mean: 6.502

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.982866	Pr < W	0.9881

Test	Statistic		p Value	
Student's t	t	2.219289	Pr > |t|	0.0449

3000-3100 Hz
Mean: 5.697

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.959651	Pr < W	0.7173

Test	Statistic		p Value	
Student's t	t	2.103379	Pr > |t|	0.0555

3100-3200 Hz
Mean: 5.867

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.952768	Pr < W	0.6044

Test	Statistic		p Value	
Student's t	t	2.18452	Pr > |t|	0.0478

3200-3300 Hz
Mean: 6.428

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.964315	Pr < W	0.7930

Test	Statistic		p Value	
Student's t	t	2.71633	Pr > |t|	0.0176

3300-3400 Hz
Mean: 8.401

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.984764	Pr < W	0.9937

Test	Statistic		p Value	
Student's t	t	4.064245	Pr > |t|	0.0013

3400-3500 Hz
Mean: 10.501

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.959318	Pr < W	0.7118

Test	Statistic		p Value	
Student's t	**t**	4.864192	**Pr > |t|**	0.0003

3500-3600 Hz
Mean: 11.354

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.924822	**Pr < W**	0.2578

Test	Statistic		p Value	
Student's t	**t**	6.563072	**Pr > |t|**	<.0001

3600-3700 Hz
Mean: 11.593

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.959024	**Pr < W**	0.7070

Test	Statistic		p Value	
Student's t	**t**	4.007157	**Pr > |t|**	0.0015

3700-3800 Hz
Mean: 9.639

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.950946	**Pr < W**	0.5755

Test	Statistic		p Value	
Student's t	**t**	3.983618	**Pr > |t|**	0.0016

3800-3900 Hz
Mean: 7.372

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.972057	**Pr < W**	0.9029

Test	**Statistic**		**p Value**	
Student's t	**t**	2.941933	**Pr > |t|**	0.0114

3900-4000 HZ
Mean: 8.652

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.897362	**Pr < W**	0.1033

Test	**Statistic**		**p Value**	
Student's t	**t**	3.021102	**Pr > |t|**	0.0098

4000-4100 Hz
Mean: 6.011

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.968079	**Pr < W**	0.8500

Test	**Statistic**		**p Value**	
Student's t	**t**	2.192652	**Pr > |t|**	0.0471

4100-4200 Hz
Mean: 6.804

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.881352	**Pr < W**	0.0608

Test	Statistic		p Value	
Student's t	t	2.638726	Pr > |t|	0.0204

4200-4300 Hz
Mean: 7.433

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.916224	Pr < W	0.1940

Test	Statistic		p Value	
Student's t	t	2.655919	Pr > |t|	0.0198

4300-4400 Hz
Mean: 5.986

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.936643	Pr < W	0.3770

Test	Statistic		p Value	
Student's t	t	2.15999	Pr > |t|	0.0500

4400-4500 Hz
Mean: 5.562

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.975502	Pr < W	0.9403

Test	Statistic		p Value	
Student's t	t	1.805917	Pr > |t|	0.0941

4500-4600 Hz
Mean: 5.994

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.972091	**Pr < W**	0.9033

Test	**Statistic**		**p Value**	
Student's t	**t**	1.908995	**Pr > |t|**	0.0786

4600-4700 Hz
Mean: 5.26

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.969305	**Pr < W**	0.8672

Test	**Statistic**		**p Value**	
Student's t	**t**	1.791676	**Pr > |t|**	0.0965

4700-4800 Hz
Mean: 6.17

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.947332	**Pr < W**	0.5201

Test	**Statistic**		**p Value**	
Student's t	**t**	2.19497	**Pr > |t|**	0.0469

4900-5000 Hz
Mean: 5.088

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.963683	**Pr < W**	0.7830

Test	Statistic		p Value	
Student's t	t	2.107165	Pr > |t|	0.0551

Test Group S: English phrases and Call words

1000-1100 Hz
Mean: 1.886

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.915636	Pr < W	0.2189

Test	Statistic		p Value	
Student's t	t	2.125831	Pr > |t|	0.0550

1200-1300 Hz
Mean: 2.626

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.903606	Pr < W	0.1499

Test	Statistic		p Value	
Student's t	t	2.376506	Pr > |t|	0.0350

1300-1400 Hz
Mean: 2.421

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.92767	Pr < W	0.3176

Test	Statistic		p Value	
Student's t	t	2.640086	Pr > |t|	0.0216

1500-1600 Hz
Mean: 2.135

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.943775	**Pr < W**	0.5076

Test	**Statistic**		**p Value**	
Student's t	**t**	2.104278	**Pr > |t|**	0.0571

2200-2300 Hz
Mean: 2.26

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.980797	**Pr < W**	0.9831

Test	**Statistic**		**p Value**	
Student's t	**t**	1.923495	**Pr > |t|**	0.0785

2300-2400 Hz
Mean: 2.391

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.971412	**Pr < W**	0.9105

Test	**Statistic**		**p Value**	
Student's t	**t**	2.1176	**Pr > |t|**	0.0558

2400-2500 Hz
Mean: 2.591

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.954208	**Pr < W**	0.6632

Test	Statistic		p Value	
Student's t	t	2.478641	Pr > |t|	0.0290

2500-2600 Hz
Mean: 3.504

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.946282	Pr < W	0.5432

Test	Statistic		p Value	
Student's t	t	2.889938	Pr > |t|	0.0136

2600-2700 Hz
Mean: 3.658

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.959944	Pr < W	0.7528

Test	Statistic		p Value	
Student's t	t	3.860879	Pr > |t|	0.0023

2700-2800 HZ
Mean: 4.085

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.933086	Pr < W	0.3738

Test	Statistic		p Value	
Student's t	t	3.798935	Pr > |t|	0.0025

2800-2900 Hz
Mean: 4.042

Tests for Normality

Test	Statistic		p Value	
Shapiro-Wilk	W	0.950085	Pr < W	0.5996

Test	Statistic		p Value	
Student's t	t	4.561633	Pr > |t|	0.0007

2900-3000 Hz
Mean: 4.416

Tests for Normality

Test	Statistic		p Value	
Shapiro-Wilk	W	0.974718	Pr < W	0.9435

Test	Statistic		p Value	
Student's t	t	5.03029	Pr > |t|	0.0003

3000-3100 Hz
Mean: 3.319

Tests for Normality

Test	Statistic		p Value	
Shapiro-Wilk	W	0.892117	Pr < W	0.1043

Test	Statistic		p Value	
Student's t	t	3.795536	Pr > |t|	0.0026

3100-3200 Hz
Mean: 2.138

Tests for Normality

Test	Statistic		p Value	
Shapiro-Wilk	W	0.887113	Pr < W	0.0891

Test	Statistic		p Value	
Student's t	t	2.365803	Pr > |t|	0.0357

3200-3300 Hz
Mean: 2.146

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.898835	Pr < W	0.1289

Test	Statistic		p Value	
Student's t	t	2.056198	Pr > |t|	0.0622

3300-3400 Hz
Mean: 2.421

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.873458	Pr < W	0.0582

Test	Statistic		p Value	
Student's t	t	2.749856	Pr > |t|	0.0176

3400-3500 Hz
Mean: 2.824

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.906789	Pr < W	0.1657

Test	Statistic		p Value	
Student's t	t	3.13542	Pr > |t|	0.0086

3500-3600 Hz
Mean: 2.917

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.895719	**Pr < W**	0.1168

Test	**Statistic**		**p Value**	
Student's t	**t**	3.240553	**Pr > |t|**	0.0071

3600-3700 Hz
Mean: 3.732

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.962863	**Pr < W**	0.7972

Test	**Statistic**		**p Value**	
Student's t	**t**	3.485244	**Pr > |t|**	0.0045

3700-3800 Hz
Mean: 4.514

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.971254	**Pr < W**	0.9088

Test	**Statistic**		**p Value**	
Student's t	**t**	4.260432	**Pr > |t|**	0.0011

3800-3900 Hz
Mean: 4.604

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.965182	**Pr < W**	0.8309

Test	Statistic		p Value	
Student's t	t	4.323474	Pr > |t|	0.0010

3900-4000 Hz
Mean: 4.387

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.933092	Pr < W	0.3738

Test	Statistic		p Value	
Student's t	t	4.935653	Pr > |t|	0.0003

4100-4200 Hz
Mean: 2.866

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.957915	Pr < W	0.7212

Test	Statistic		p Value	
Student's t	t	3.337019	Pr > |t|	0.0059

4200-4300 Hz
Mean: 2.683

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.991677	Pr < W	0.9999

Test	Statistic		p Value	
Student's t	t	2.650129	Pr > |t|	0.0212

4300-4400 Hz
Mean: 4.27

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.960527	**Pr < W**	0.7618

Test	Statistic		p Value	
Student's t	**t**	2.250093	**Pr > |t|**	0.0440

4400-4500 Hz
Mean: 2.431

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.981671	**Pr < W**	0.9866

Test	Statistic		p Value	
Student's t	**t**	2.45377	**Pr > |t|**	0.0304

4599-4600 Hz
Mean: 2.563

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.958087	**Pr < W**	0.7239

Test	Statistic		p Value	
Student's t	**t**	1.922295	**Pr > |t|**	0.0786

Test Group S: First Language texts

F0
Means: 1.226

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.954326	Pr < W	0.5950

Test	Statistic		p Value	
Student's t	t	1.807366	Pr > |t|	0.0922

1200-1300 Hz
Means: 2.447

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.9385	Pr < W	0.3639

Test	Statistic		p Value	
Student's t	t	4.665306	Pr > |t|	0.0004

1300-1400 Hz
Means: 2.001

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.939076	Pr < W	0.3709

Test	Statistic		p Value	
Student's t	t	3.858825	Pr > |t|	0.0017

1400-1500 Hz
Mean: 1.273

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.946479	Pr < W	0.4708

Test	Statistic		p Value	
Student's t	t	2.182032	Pr > |t|	0.0466

1500-1600 Hz
Mean: 3.013

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.93962	Pr < W	0.3776

Test	Statistic		p Value	
Student's t	t	5.146263	Pr > |t|	0.0001

1600-1700 Hz
Mean: 2.955

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.898729	Pr < W	0.0910

Test	Statistic		p Value	
Student's t	t	4.974268	Pr > |t|	0.0002

1700-1800 Hz
Mean: 2.912

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.964806	Pr < W	0.7751

Test	Statistic		p Value	
Student's t	t	4.997377	Pr > |t|	0.0002

1800-1900 Hz
Mean: 3.353

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.934901	**Pr < W**	0.3226

Test	**Statistic**		**p Value**	
Student's t	**t**	4.446506	**Pr > |t|**	0.0006

1900-2000 Hz
Mean: 3.586

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.944329	**Pr < W**	0.4399

Test	**Statistic**		**p Value**	
Student's t	**t**	4.555289	**Pr > |t|**	0.0004

2000-2100 Hz
Mean: 3588

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.916307	**Pr < W**	0.1692

Test	**Statistic**		**p Value**	
Student's t	**t**	4.786415	**Pr > |t|**	0.0003

2100-2200 Hz
Mean: 3.994

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.9132	**Pr < W**	0.1516

Test	Statistic		p Value	
Student's t	t	5.644769	Pr > |t|	<.0001

2200-2300 Hz
Mean: 3.688

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.934423	Pr < W	0.3174

Test	Statistic		p Value	
Student's t	t	4.80322	Pr > |t|	0.0003

2300-2400 Hz
Mean: 3.531

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.925848	Pr < W	0.2363

Test	Statistic		p Value	
Student's t	t	5.506405	Pr > |t|	<.0001

2400-2500 Hz
Mean: 3.743

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.91895	Pr < W	0.1857

Test	Statistic		p Value	
Student's t	t	4.548332	Pr > |t|	0.0005

2500-2600 Hz
Mean: 4.14

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.969143	**Pr < W**	0.8452

Test	**Statistic**		**p Value**	
Student's t	**t**	4.38691	**Pr > |t|**	0.0006

2600-2700 Hz
Mean: 3.933

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.967835	**Pr < W**	0.8248

Test	**Statistic**		**p Value**	
Student's t	**t**	4.494484	**Pr > |t|**	0.0005

2700-2800 Hz
Mean: 3.733

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.969127	**Pr < W**	0.8449

Test	**Statistic**		**p Value**	
Student's t	**t**	4.014262	**Pr > |t|**	0.0013

2800-2900 Hz
Mean: 3.991

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.92328	**Pr < W**	0.2161

Test	Statistic		p Value	
Student's t	t	5.004458	Pr > |t|	0.0002

2900-3000 Hz
Mean: 4.423

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.923905	Pr < W	0.2209

Test	Statistic		p Value	
Student's t	t	5.074005	Pr > |t|	0.0002

3000-3100 Hz
Mean: 3.899

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.949955	Pr < W	0.5237

Test	Statistic		p Value	
Student's t	t	5.092367	Pr > |t|	0.0002

3100-3200 Hz
Mean: 3.589

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.960192	Pr < W	0.6957

Test	Statistic		p Value	
Student's t	t	4.886319	Pr > |t|	0.0002

3200-3300 Hz
Mean: 3.353

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.950428	Pr < W	0.5312

Test	Statistic		p Value	
Student's t	t	4.768125	Pr > |t|	0.0003

3300-3400 Hz
Mean: 3.103

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.941204	Pr < W	0.3978

Test	Statistic		p Value	
Student's t	t	4.648327	Pr > |t|	0.0004

3400-3500 Hz
Mean: 2.807

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.934013	Pr < W	0.3130

Test	Statistic		p Value	
Student's t	t	5.271403	Pr > |t|	0.0001

3500-3600 Hz
Mean: 2.941

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.984437	Pr < W	0.9911

Test	Statistic		p Value	
Student's t	t	4.336459	Pr > |t|	0.0007

3600-3700 Hz
Mean: 2.607

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.966058	Pr < W	0.7960

Test	Statistic		p Value	
Student's t	t	3.864966	Pr > |t|	0.0017

3700-3800 Hz
Mean: 2.747

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.889337	Pr < W	0.0655

Test	Statistic		p Value	
Student's t	t	3.71648	Pr > |t|	0.0023

3800-3900 Hz
Mean: 3.493

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.890811	Pr < W	0.0690

Test	Statistic		p Value	
Student's t	t	5.015805	Pr > |t|	0.0002

3900-4000 Hz
Mean: 2.741

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.869738	**Pr < W**	0.0334

Test	**Statistic**		**p Value**	
Student's t	**t**	4.100334	**Pr > |t|**	0.0011

4200-4300 Hz
Mean: 1.911

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.936224	**Pr < W**	0.3373

Test	**Statistic**		**p Value**	
Student's t	**t**	3.154416	**Pr > |t|**	0.0070

4700-4800 Hz
Mean: 1.459

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.912531	**Pr < W**	0.1481

Test	**Statistic**		**p Value**	
Student's t	**t**	1.922123	**Pr > |t|**	0.0752

Test Group V

Test Group V: Y-buzz

200-300 Hz
Mean: 9.383

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.93295	**Pr < W**	0.3723

Test	Statistic		p Value	
Student's t	**t**	9.146084	**Pr > |t|**	<.0001

600-700 Hz
Mean: 5.007

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.920983	**Pr < W**	0.2586

Test	Statistic		p Value	
Student's t	**t**	2.618548	**Pr > |t|**	0.0224

1300-1400 Hz
Mean: 3.506

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.959396	**Pr < W**	0.7443

Test	Statistic		p Value	
Student's t	**t**	1.708419	**Pr > |t|**	0.1133

1500-1600 Hz
Mean: 6.103

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.928177	Pr < W	0.3225

Test	Statistic		p Value	
Student's t	t	3.17574	Pr > |t|	0.0080

1600-1700 Hz
Mean: 7.557

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.938583	Pr < W	0.4388

Test	Statistic		p Value	
Student's t	t	3.668022	Pr > |t|	0.0032

1700-1800 Hz
Mean: 7.315

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.95022	Pr < W	0.6016

Test	Statistic		p Value	
Student's t	t	3.06237	Pr > |t|	0.0099

1800-1900 Hz
Mean: 5.229

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.872278	Pr < W	0.0561

Test	Statistic		p Value	
Student's t	t	1.820275	Pr > |t|	0.0937

1900-2000 Hz
Mean: 7.322

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.976128	Pr < W	0.9552

Test	Statistic		p Value	
Student's t	t	2.816192	Pr > |t|	0.0156

2000-2100 Hz
Mean: 9.32

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.961139	Pr < W	0.7712

Test	Statistic		p Value	
Student's t	t	2.987526	Pr > |t|	0.0113

2200-2300 Hz
Mean: 8.947

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.963066	Pr < W	0.8002

Test	Statistic		p Value	
Student's t	t	2.245119	Pr > |t|	0.0444

2300-2400 Hz
Mean: 5.107

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.954572	**Pr < W**	0.6689

Test	**Statistic**		**p Value**	
Student's t	**t**	1.864257	**Pr > |t|**	0.0869

2400-2500 Hz
Mean: 6.452

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.977937	**Pr < W**	0.9680

Test	**Statistic**		**p Value**	
Student's t	**t**	3.00635	**Pr > |t|**	0.0109

2600-2700 Hz
Mean: 3.711

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.960034	**Pr < W**	0.7542

Test	**Statistic**		**p Value**	
Student's t	**t**	2.029524	**Pr > |t|**	0.0652

3800-3900 Hz
Mean: 2.96

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.930515	**Pr < W**	0.3461

Test	Statistic		p Value	
Student's t	t	1.706168	Pr > |t|	0.1137

Test Group V: +Y-buzz

200-300 Hz
Mean: 4.891

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.948902	Pr < W	0.5818

Test	Statistic		p Value	
Student's t	t	1.944482	Pr > |t|	0.0757

1200-1300 Hz
Mean: 7.769

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.958039	Pr < W	0.7232

Test	Statistic		p Value	
Student's t	t	3.156944	Pr > |t|	0.0083

1400-1500 Hz
Mean: 9.618

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.870313	Pr < W	0.0528

Test	Statistic		p Value	
Student's t	t	3.934115	Pr > |t|	0.0020

1600-1700 Hz
Mean: 10.928

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.978586	**Pr < W**	0.9719

Test	**Statistic**		**p Value**	
Student's t	**t**	4.077464	**Pr > |t|**	0.0015

1800-1900 Hz
Mean: 8.77

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.924884	**Pr < W**	0.2917

Test	**Statistic**		**p Value**	
Student's t	**t**	2.744419	**Pr > |t|**	0.0178

1900-2000 Hz
Mean: 6.449

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.93778	**Pr < W**	0.4288

Test	**Statistic**		**p Value**	
Student's t	**t**	2.906005	**Pr > |t|**	0.0132

2000-2100 Hz
Mean: 7.625

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.961071	**Pr < W**	0.7701

Test	Statistic		p Value	
Student's t	t	2.712544	Pr > |t|	0.0189

2100-2200 Hz
Mean: 8.686

<table>
<tr><th colspan="5">Tests for Normality</th></tr>
<tr><th>Test</th><th colspan="2">Statistic</th><th colspan="2">p Value</th></tr>
<tr><td>Shapiro-Wilk</td><td>W</td><td>0.908759</td><td>Pr < W</td><td>0.1763</td></tr>
</table>

Test	Statistic		p Value	
Student's t	t	3.371656	Pr > |t|	0.0056

2300-2400 Hz
Mean: 6.811

<table>
<tr><th colspan="5">Tests for Normality</th></tr>
<tr><th>Test</th><th colspan="2">Statistic</th><th colspan="2">p Value</th></tr>
<tr><td>Shapiro-Wilk</td><td>W</td><td>0.880649</td><td>Pr < W</td><td>0.0728</td></tr>
</table>

Test	Statistic		p Value	
Student's t	t	4.198774	Pr > |t|	0.0012

2500-2600 Hz
Mean: 6.595

<table>
<tr><th colspan="5">Tests for Normality</th></tr>
<tr><th>Test</th><th colspan="2">Statistic</th><th colspan="2">p Value</th></tr>
<tr><td>Shapiro-Wilk</td><td>W</td><td>0.951195</td><td>Pr < W</td><td>0.6165</td></tr>
</table>

Test	Statistic		p Value	
Student's t	t	3.323856	Pr > |t|	0.0061

2700-2800 Hz
Mean: 6.992

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.951954	**Pr < W**	0.6282

Test	Statistic		p Value	
Student's t	**t**	3.162434	**Pr > |t|**	0.0082

2800-2900 Hz
Mean: 4.878

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.878356	**Pr < W**	0.0677

Test	Statistic		p Value	
Student's t	**t**	2.011041	**Pr > |t|**	0.0673

3400-3500 Hz
Mean: 4.204

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.920212	**Pr < W**	0.2525

Test	Statistic		p Value	
Student's t	**t**	2.295588	**Pr > |t|**	0.0405

3500-3600 Hz
Mean: 1.255

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.920151	**Pr < W**	0.2520

Test	Statistic		p Value	
Student's t	t	0.98446	Pr > |t|	0.3443

3600-3700 Hz
Mean: 6.419

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.915944	Pr < W	0.2210

Test	Statistic		p Value	
Student's t	t	3.420892	Pr > |t|	0.0051

3700-3800 Hz
Mean: 6.88

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.89409	Pr < W	0.1110

Test	Statistic		p Value	
Student's t	t	3.330773	Pr > |t|	0.0060

3800-3900 Hz
Mean: 4.977

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.922586	Pr < W	0.2718

Test	Statistic		p Value	
Student's t	t	2.43481	Pr > |t|	0.0315

Test Group V: Calls

1200-1300 Hz
Mean: 6.837

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.955035	**Pr < W**	0.6761

Test	**Statistic**		**p Value**	
Student's t	**t**	3.285842	**Pr > |t|**	0.0065

1600-1700 Hz
Mean: 4.066

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.981241	**Pr < W**	0.9850

Test	**Statistic**		**p Value**	
Student's t	**t**	2.568834	**Pr > |t|**	0.0246

2300-2400 Hz
Mean: 3.491

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.969896	**Pr < W**	0.8930

Test	**Statistic**		**p Value**	
Student's t	**t**	2.545203	**Pr > |t|**	0.0257

3200-3300 Hz
Mean: 3.501

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.888085	**Pr < W**	0.0918

Test	**Statistic**		**p Value**	
Student's t	**t**	1.968161	**Pr > |t|**	0.0726

3400-3500 Hz
Mean: 5.703

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.891887	**Pr < W**	0.1035

Test	**Statistic**		**p Value**	
Student's t	**t**	3.532524	**Pr > |t|**	0.0041

3500-3600 Hz
Mean: 4.907

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.892196	**Pr < W**	0.1045

Test	**Statistic**		**p Value**	
Student's t	**t**	3.219972	**Pr > |t|**	0.0074

3600-3700 Hz
Mean: 6.713

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.94693	**Pr < W**	0.5526

Test	Statistic		p Value	
Student's t	t	4.275361	Pr > |t|	0.0011

3700-3800 Hz
Mean: 3.52

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.962472	Pr < W	0.7913

Test	Statistic		p Value	
Student's t	t	1.980419	Pr > |t|	0.0711

Test Group V: English phrases and Call words

200-300 Hz
Mean: 5.35

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.957325	Pr < W	0.7451

Test	Statistic		p Value	
Student's t	t	3.205184	Pr > |t|	0.0084

1500-1600 Hz
Mean: 1.588

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.956255	Pr < W	0.7293

Test	Statistic		p Value	
Student's t	t	2.368895	Pr > |t|	0.0372

1600-1700 Hz
Mean: 2.619

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.859441	**Pr < W**	0.0481

Test	Statistic		p Value	
Student's t	**t**	3.441777	**Pr > |t|**	0.0055

Test Group V: First language texts

700-800 Hz
Mean: 1.2

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.953278	**Pr < W**	0.6853

Test	Statistic		p Value	
Student's t	**t**	1.836214	**Pr > |t|**	0.0935

800-900 Hz
Mean: 3.698

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.938763	**Pr < W**	0.4822

Test	Statistic		p Value	
Student's t	**t**	3.98764	**Pr > |t|**	0.0021

900-1000 Hz
Mean: 3.525

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.947105	**Pr < W**	0.5951

Test	**Statistic**		**p Value**	
Student's t	**t**	3.518732	**Pr > |t|**	0.0048

1000-1100 Hz
Mean: 2.275

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.938092	**Pr < W**	0.4738

Test	**Statistic**		**p Value**	
Student's t	**t**	2.927368	**Pr > |t|**	0.0138

1100-1200 Hz
Mean: 2.368

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.937452	**Pr < W**	0.4658

Test	**Statistic**		**p Value**	
Student's t	**t**	2.693795	**Pr > |t|**	0.0209

1200-1300 Hz
Mean: 2.583

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.95105	**Pr < W**	0.6524

Test	Statistic		p Value	
Student's t	t	3.211552	Pr > |t|	0.0083

1300-1400 Hz
Mean: 2.19

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.9516	Pr < W	0.6605

Test	Statistic		p Value	
Student's t	t	4.185759	Pr > |t|	0.0015

1500-1600 Hz
Mean: 5.029

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.97081	Pr < W	0.9191

Test	Statistic		p Value	
Student's t	t	5.880458	Pr > |t|	0.0001

1600-1700 Hz
Mean: 6.751

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.965246	Pr < W	0.8552

Test	Statistic		p Value	
Student's t	t	7.390818	Pr > |t|	<.0001

1700-1800 Hz
Mean: 3.957

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.962596	Pr < W	0.8201

Test	Statistic		p Value	
Student's t	t	5.243638	Pr > |t|	0.0003

1800-1900 Hz
Mean: 2.048

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.938326	Pr < W	0.4767

Test	Statistic		p Value	
Student's t	t	3.159691	Pr > |t|	0.0091

1900-2000 Hz
Mean: 1.965

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.925147	Pr < W	0.3315

Test	Statistic		p Value	
Student's t	t	2.750029	Pr > |t|	0.0189

2500-2600 Hz
Mean: 2.173

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.887322	Pr < W	0.1088

Test	**Statistic**		**p Value**	
Student's t	**t**	3.346474	**Pr > |t|**	0.0065