

# DOES REAL-TIME VISUAL FEEDBACK IMPROVE PITCH ACCURACY IN SINGING?

PAT WILSON

A THESIS SUBMITTED TO FULFIL THE REQUIREMENTS OF  
THE DEGREE OF MASTER OF APPLIED SCIENCE  
(COMMUNICATION SCIENCES AND DISORDERS)

SCHOOL OF COMMUNICATION SCIENCES AND DISORDERS  
THE UNIVERSITY OF SYDNEY  
SEPTEMBER 2006

**The soul never thinks without a mental image.**

Aristotle. (Hett, W.S., trans.) (1936). *On the soul; Parva naturalia; On breath*.  
Cambridge, MA, Harvard University Press, p. 428.

**One seeing is worth a hundred listenings.**

Chinese proverb quoted in Cole, Hugo. (1974). *Sounds and symbols: Aspects of musical notation*. London, Oxford University Press, p. 122.

# Abstract

The aim of this investigation was to investigate the effects of computer-based visual feedback in the teaching of singing. Pitch accuracy, a readily-measured parameter of the singing voice, was used in this study to gauge changes in singing for groups with and without visual feedback. The study investigated whether the style of feedback affects the amount of learning achieved, and whether the provision of concurrent visual feedback hampers the simultaneous performance of the singing task.

The investigation used a baseline–intervention–post-test between-groups design. Participants of all skill levels were randomly assigned to a control group or one of two experimental groups – with all participants given one hour of singing training. At intervention, the two experimental groups were offered one of two different displays of real-time visual feedback on their vocal pitch accuracy, while control participants had a non-interactive display. All sessions were recorded, and the vocal exercise patterns performed at baseline, intervention and post-test phases were acoustically analysed for pitch accuracy. Questionnaires assessed both general health and the amount of singing and music training of all participants; people in the two experimental groups were also given a further questionnaire about the visual feedback.

The results indicate that visual feedback improves pitch accuracy in singing. Cognitive load related to the decoding of visual information was a factor at intervention. At post-test, the two groups who had used real-time visual feedback demonstrated marked improvement on their initial pitch accuracy. There was no significant difference between the results of participants from the two experimental groups, although the participants with some background in singing training showed greater improvement using a simpler visual feedback design.

The findings suggest that a hybrid approach integrating standard singing teaching practices with real-time visual feedback of aspects of the singing voice may improve learning.

# Acknowledgements

If it takes a village to bring up a child, then it takes a university to write a thesis.

I offer my thanks and gratitude to all of these people.

To my supervisors:

Dr Kerrie **Lee**, Head of Department, School of Communication Sciences and Disorders, Faculty of Health Sciences, University of Sydney

Dr Jean **Callaghan**, voice researcher and specialist in singing pedagogy

Dr William **Thorpe**, Research Scientist, The Bioengineering Institute, University of Auckland

...and to everyone else:

(in alphabetical order)

Adrian **Barnes**

Liz **Barnes**, Biostatistician, NHMRC Clinical Trials Centre, University of Sydney - for help as a consultant statistician.

Jonathon **Crane** - for developing the computer displays used in this research.

Dr Rob **Heard**, Senior Lecturer, School of Behavioural and Community Health, University of Sydney - for assistance with statistical questions.

Neryla **Jolly**, Head of School, School of Applied Vision Sciences, Faculty of Health Sciences, University of Sydney - for useful advice.

Dr Patricia **McCabe**, Lecturer and Director of Clinical Education, School of Communication Sciences and Disorders, Faculty of Health Sciences, University of Sydney - for helpful advice, encouragement and keen insights.

Cate **Madill**, Lecturer, School of Communication Sciences and Disorders, Faculty of Health Sciences, University of Sydney - for allowing me to practice on her when developing the experimental sessions.

Bronwyn **Millar**, Research Assistant, School of Communication Sciences and Disorders, Faculty of Health Sciences, University of Sydney - for cheerfulness and helpful advice.

Evert-Jan **Mirck** - for valuable references related to cognitive load and the learning process.

Dr Kate **Peart-Reid** - for useful references about voice, learning and knowledge of results.

Professor Vicki A. **Reed**, Head and Professor, Communication Sciences and Disorders, James Madison University, Virginia, U.S.A. - for early advice and guidance.

Mary **Routledge**, Office Manager, School of Communication Sciences and Disorders, University of Sydney - for consistent helpfulness and encouragement.

Ashley **Scott**, Information and Communications Technology, University of Sydney - for foundational work which enabled this research to go forward.

Associate Professor Kate **Stevens**, Deputy Director, MARCS Auditory Laboratories, University of Western Sydney - for early observations, encouragement and inspiration.

Dr Leanne **Togher**, Lecturer, School of Communication Sciences and Disorders, Faculty of Health Sciences, University of Sydney - for allowing me to practice on her when developing the experimental sessions.

David **Trembath**, Lecturer, School of Communication Sciences and Disorders, Faculty of Health Sciences, University of Sydney - for invaluable practical assistance.

Professor Jan **van Doorn**, Department of Clinical Sciences, Division of Logopedics, Umeå University, Sweden - for research which formed the foundation of my current work.

Dr Kittie **Verdolini**, Associate Professor, Communication Sciences and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh - for pointing me in a useful direction.

Prof. Graham **Welch**, Chair of Music Education, School of Arts and Humanities, Institute of Education, University of London - for permission to use materials from his recent paper.

Dr Mark **Wiggins**, Senior Lecturer, School of Psychology and MARCS Auditory Laboratories, University of Western Sydney - for constructive insights offered in the early stages of this research.

# Submission Statement

The work contained in this thesis has not been submitted for a degree to any other university or institution.

The conduct of this research was approved by the University of Sydney Human research Ethics Committee.

Pat H. Wilson

Date:

This thesis entitled 'Does real-time visual feedback improve pitch accuracy in singing?' by Pat Wilson is in a form that is acceptable for submission for the degree of Master of Applied Science, Communication Sciences and Disorders.

# Presentations arising from this thesis

## Conference presentations with refereed abstracts in published conference proceedings

**Wilson, P.**, Callaghan, J. and Thorpe, C.W. (2005). 'The singer, the song, the voice and its visuals: Can real-time visual feedback help singing students?' In *Deliver the Voice. 7<sup>th</sup> Voice Symposium of Australia*. Sydney: Australian Voice Association.

**Wilson, P.**, Thorpe, C. W., and Callaghan, J. (2005). 'Looking at singing: Does real-time visual feedback improve the way we learn to sing?' In *2<sup>nd</sup> APSCOM Conference*. Seoul, South Korea: Asia-Pacific Society for the Cognitive Sciences of Music. [Book ISBN 89-5708-088-0; CD-ROM ISBN 89-5708-090-2]

## Conference presentations

**Wilson, P.**, Lee, K., Callaghan, J. and Thorpe, C.W. (2005). Seeing the singing voice: The effect of different visual displays in real-time visual feedback for singing training. In *Research for a Healthy Future – Colloquium 2005*. Faculty of Health Sciences, University of Sydney, 1<sup>st</sup> – 2<sup>nd</sup> December.

**Wilson, P.** (2004). Seeing, singing and computing. *One-day Showcase, Information and Communications Technology*, College of Sciences and Technology, University of Sydney, 13<sup>th</sup> August. (Poster presentation)

**Wilson, P.** (2003). How to be a good-looking singer: The effect of the visual displays in real-time visual feedback for singing training. *Postgraduate Research Day*, Australian Centre for Applied Research in Music Performance, Conservatorium of Music, University of Sydney, 8<sup>th</sup> October.

**Wilson, P.** (2002). Looking and singing: The effect of the visual display of elements in real-time visual feedback on learning in singing. *Research Day*, School of Communications Sciences and Disorders, Faculty of Health Sciences, University of Sydney, November.

## **Awards for papers based on this research**

First Prize: Australian Music & Psychology Society (AMPS) Inc. Graduate Student International Conference Travel Assistance Scheme Award (April 2005). The award was based on the quality of the abstract of my paper for the *Second International Conference of the Asia-Pacific Society for the Cognitive Science of Music*, Seoul, South Korea, August 2005.

First Prize: Presentation at *Research for a Health Future – Colloquium 2005*. Faculty of Health Sciences, University of Sydney (December 2005).

# Table of Contents

Abstract.....	iii
Acknowledgements.....	iv
Submission Statement.....	vi
Presentations arising from this thesis.....	vii
Awards for papers based on this research.....	viii
Figures.....	xii
Tables.....	xiii
Glossary.....	xiv
Preface.....	xv
Pitch designation.....	xvii
Chapter 1 Singing and learning.....	1
1.1 Overview.....	1
1.2 The beginning of singing.....	2
1.3 Early views on learning singing.....	4
1.3.1 Traditional singing teaching.....	5
1.4 Learning.....	7
1.4.1 Introduction.....	7
1.4.2 Learning to sing.....	8
1.4.3 Cognitive load theory.....	9
1.4.4 Learning through the senses.....	10
1.4.5 Learning styles.....	11
1.5 Psychomotor learning.....	13
1.5.1 Introduction.....	13
1.5.2 Neural mechanisms of learning and retention.....	14
1.5.3 Practice and motor programs.....	17
1.5.4 Mental imagery as an enhancer of motor skills learning.....	19
1.6 Learning and feedback.....	19
1.6.1 Introduction.....	19
1.7 Visual input as part of learning.....	22
1.7.1 Auditory input as part of learning – Perception.....	24
1.7.2 Feedback and self-directed learning.....	26
1.7.2.1 Feedback on demand.....	26
1.8 Knowledge of results.....	27
1.8.1 Knowledge of results and the singer.....	29
1.9 Feedback technologies.....	33
1.9.1 Sound recording.....	34
1.9.2 Visual records of performance.....	34
1.9.3 Medical technologies.....	35
1.10 Technology and learning.....	36
1.10.1 Introduction.....	36
1.10.2 Human-computer interaction.....	37
1.10.3 Design factors in a visual display.....	37
1.10.3.1 Optimisation of graphics.....	39
1.10.3.2 Design of interactive visual displays for computers.....	40
1.10.3.3 Speed of access of visual elements.....	43
1.10.4 Technology and music training.....	44
1.10.5 Technology and singing training.....	46

1.10.6	Auditory, visual and kinaesthetic inputs in teaching singing .....	50
1.10.7	The acoustic analysis of sound .....	50
1.10.8	The perceptual analysis of sung sound .....	52
1.10.9	Technology and the singer .....	52
1.11	Research questions .....	53
Chapter 2	Method .....	54
2.1	Research design .....	54
2.1.1	Pitch matching .....	54
2.2	Collection of singing data .....	56
2.2.1	Baseline recording – All groups .....	56
2.2.2	Intervention recordings – Experimental groups .....	57
2.2.3	Intervention recordings – Control group .....	57
2.2.4	Post-test recordings – All groups .....	58
2.3	The singing exercise patterns .....	58
2.4	Visual feedback displays .....	59
2.4.1	Experimental group 1 – Dynamic pitch grid display .....	60
2.4.2	Experimental group 2 – Interactive keyboard display .....	61
2.4.3	Control Group Visual Display .....	64
2.4.4	Comparison of the three visual feedback conditions .....	65
2.5	Equipment .....	66
2.6	Data collection procedures .....	66
2.6.1	Pitch accuracy differences – Baseline versus post-test .....	67
2.7	Participants .....	68
2.7.1	Sample size .....	69
2.7.2	Exclusion of participants – Data fidelity .....	69
2.7.3	Participant summary .....	70
2.8	Questionnaires .....	70
2.8.1	Pre-session questionnaire on music/singing training .....	70
2.8.2	Post-session questionnaire .....	71
Chapter 3	Results .....	75
3.1	Calculation of pitch error .....	75
3.1.1	Incidence of ‘flat’ notes .....	75
3.2	Effect of feedback on performance .....	78
3.3	Interference effect of intervention .....	79
3.4	Summary - Effect of feedback .....	82
3.5	Novice versus experienced singers .....	82
3.6	Visually-biased participants .....	83
3.7	Results from post-session questionnaire .....	85
3.7.1	Post-session Questionnaire .....	85
3.7.1.1	Learning style preferences .....	85
3.7.1.2	Handedness and colour-blindness .....	86
3.7.1.3	Vocal problems .....	86
3.7.1.4	Ear/hearing problems .....	87
3.7.1.5	Smoking .....	87
3.7.1.6	Medication .....	88
3.7.1.7	Written responses to VFB displays .....	88
Chapter 4	Discussion .....	94
4.1	Introduction .....	94
4.2	Does pitch accuracy performance improve from baseline to post-test as a result of visual feedback at intervention? .....	94

4.2.1	Why did the control group show no improvement with practice?.....	94
4.3	Does the type of visual information available in the feedback produce differences in pitch accuracy performance? .....	95
4.4	Informational role of KR .....	96
4.4.1	Accessibility of KR.....	96
4.4.2	Amount of KR.....	98
4.4.3	Frequency of KR.....	98
4.4.4	Precision of KR.....	99
4.5	Motivational role of KR.....	99
4.6	Associational role of KR.....	100
4.7	What is the effect of visual feedback during task performance? .....	101
4.7.1	Why did the two groups which received visual feedback show a decrement in performance at intervention?.....	102
4.7.2	Why did the keyboard group perform better than the grid group at intervention? .....	103
4.8	Exceptions to the rule: Individuals who improved during intervention ...	105
4.9	Impact of skill level .....	105
4.9.1	Contrasting the results of beginner and advanced participants.....	106
4.10	Did visual-spatial learners do better with visual feedback?.....	108
4.10.1	Verbal versus visual KR .....	108
4.11	Post-session questionnaire responses.....	109
4.12	Pitch measurement – ‘Flat’ or ‘Sharp’? .....	110
4.13	Future directions .....	111
References	.....	114
Appendices	.....	128
Appendix A:	An annotated list of voice analysis software.....	129
Appendix B:	Session scripts for experimental groups.....	133
Appendix B:	Session script for control group.....	137
Appendix C:	Pre-session Questionnaire .....	141
Appendix C:	Post Session Questionnaires.....	143
Appendix D:	Participants’ Consent Sheet.....	146
Appendix E:	Subject Information Sheet - Research Studies .....	147
Appendix F:	Post-session questionnaire - Written responses .....	150
Appendix G:	Post-session questionnaire – Text analysis .....	157
Appendix H:	Advertisement - Call for participants.....	163

# Figures

Figure 1-1:	A traditional model of teacher feedback (knowledge of results [KR]) during the course of a singing lesson	29
Figure 1-2:	A real-time model of feedback (knowledge of results [KR]) during the course of a singing lesson	30
Figure 2-1:	Singing exercise pattern for ‘Beginner’ skill level comprising a 1 – 3 – 1 pattern	58
Figure 2-2:	Singing exercise pattern for ‘Advanced-1’ skill level comprising a 1 – 4 – 6 – 4 – 1 pattern	59
Figure 2-3:	Singing exercise pattern for ‘Advanced-2’ skill level comprising a 1 – flat7 – 2 – 3 – 4 pattern	59
Figure 2-4:	Experimental Screen Design – Dynamic Pitch Grid Display	60
Figure 2-5:	Experimental Screen Design - Interactive Keyboard Display showing the target notes in green	62
Figure 2-6:	Experimental Screen Design – Interactive Keyboard Display – third note being sung. The target turns red when the note is sung.	63
Figure 2-7:	Experimental Screen Design – Non-interactive Control Display	64
Figure 3-1:	Mean pitch error per note (N1 to N5) for each group at baseline, intervention and post-test. The results for the Beginner singers who sang a three-note pattern and the Advanced singers who sang a five-note pattern are shown, and the mean for N1 N2 and N3 for all participants combined. The mean for all notes for Beginners and Advanced combined is also shown.	77
Figure 3-2:	Mean $F_0$ error at baseline and at post-test for each group: Outcomes of performance after feedback	78
Figure 3-3:	Mean pitch error for each group at baseline, intervention and post-test	80
Figure 3-4:	Difference score for $F_0$ error between performance at baseline and at intervention for individual participants in each group	81
Figure 3-5:	Pitch accuracy differences between baseline and post-test of beginner (Skill Level 1) and advanced (Skill Levels 2 and 3) participants in the two feedback conditions (grid and keyboard) and the control condition	82
Figure 3-6:	Difference in pitch accuracy between baseline and post-test for beginner and advanced singers in all groups	83
Figure 3-7:	Distribution of $F_0$ difference error of pitch accuracy change between baseline and post-test, showing the placement of the eight participants (shaded green) who identified themselves as visual-spatial learners on a questionnaire	84

## Tables

Table 2-1:	Comparison of visual feedback characteristics of the visual display for each group	65
Table 2-2:	Gender and age details for participants in the study	68
Table 2-3:	Mean and standard deviation for all groups (dynamic pitch grid, interactive keyboard and control) for each test (baseline, intervention and post-test), for all participants, and for group after some participants were excluded	69
Table 2-4:	Consolidated table: Males/Females, Skill Level and Feedback Mode	70
Table 2-5:	Questionnaire and Multiple Intelligences Theory categories	72
Table 3-1:	Incidence of ‘flat’ notes in sample	76
Table 3-2:	Results of One-Way Analysis of Variance and planned contrasts using difference in Hz between pitch error at baseline and at post-test	79
Table 3-3:	Results of repeated measure ANOVA based on comparing the grid, keyboard and control groups on their F <sub>0</sub> error performance at baseline and at intervention	80
Table 3-4:	Results of Dunnett <i>t</i> test: Comparison of control group to the two treatment groups	81
Table 3-5:	Mean and standard deviation for pitch error between baseline and intervention for each group	82
Table 3-6:	Visual/spatial learning preferences	83
Table 3-7:	Learning style preferences	86
Table 3-8:	History of vocal problems	87
Table 3-9:	Caffeine consumption	88
Table 3-10:	Comparison of frequency of usage of types of words in written responses of participants using Dynamic Pitch Grid screen display and Interactive Keyboard screen display; word frequencies in responses to Question 1	89
Table 3-11:	Comparison of frequency of usage of types of words in written responses of participants using Dynamic Pitch Grid screen display and Interactive Keyboard screen display; word frequencies in responses to Question 2	90
Table 3-12:	Comparison of frequency of usage of types of words in written responses of participants using Dynamic Pitch Grid screen display and Interactive Keyboard screen display; word frequencies in responses to Question 3	91
Table 3-13:	Comparison of frequency of usage of types of words in written responses of participants using Dynamic Pitch Grid screen display and Interactive Keyboard screen display; word frequencies in responses to Question 4	92

# Glossary

## Acronyms and technical terms used in scientific fields related to the areas of this research

ADC	Analog-to-digital convertor
ANOVA	Analysis of variance
BFP	Bio-feedback programmes
CAL	Computer-assisted learning
CQ	Closed quotient
CLT	Cognitive load theory
EGG	Electroglottogram
FFT	Fast Fourier Transform
F <sub>0</sub>	Fundamental frequency
GUI	Graphic user interface
HCI	Human-computer interaction
HE	Heuristic evaluation
Hz	Hertz
ICT	Information and communications technologies
ILT	Information and learning technologies
IV	Information visualisation
KP	Knowledge of performance
KR	Knowledge of results
LEMG	Laryngeal electromyography
LTM	Long-term memory
PA	Public Address system
PC	Personal computer
SPL	Sound pressure level
SPR	Singing power ratio
UIMS	User-interface management systems
VDT	Visual display terminal
VEP	Visual evoked potential
VSL	Videostrobolaryngoscopy

# Preface

How am I going? This is a question which arises frequently when a person sets about the acquisition of a new set of neuromuscular skills. Learning to sing draws together a complex web of inter-related tasks which has always challenged tyro singers. Both singers and their teachers have traditionally sought improved means of defining and assessing these tasks. Anything that assists the understanding of the neuromuscular, emotional, aesthetic and intellectual work involved in singing has been welcome. This investigation examines a new means of offering immediate feedback to a singer. It is software which displays information about a singer's voice in (near) real-time on a computer screen.

While it is no new thing for singers to use visual inputs whilst singing, nor is it unusual for singing teachers to use visual information as an element in the teaching of singing, the nature of the information provided by this software begs a more thorough exploration of the way in which singers learn to sing, and, by extension, whether real-time visual feedback of aspects of a singer's voice is beneficial in their learning process. Two different modes of presentation of visual information about the singer's voice were used in this study, as well as a non-interactive control mode.

Research questions addressed in this study include:

- Does visual feedback improve performance?
- How do singers process visual information whilst singing?
- Is there an interference effect associated with simultaneous processing of aural, visual and other sensory information by singers?
- Do singers benefit from knowledge of results (KR)?
- If so, what sort, how much, and when?

Faced with visually-based technological assistance in the area of singing pedagogy, the task of practical research is to derive a coherent visual language for a process which has until now been predominantly auditory, linguistic and kinaesthetic.

As technological advances continue to offer us new ways of seeing and understanding our world, so must pedagogical processes adjust in response to these developments. It is necessary that the worth of such advances is critically examined in the light of their likely pedagogical usefulness.

P.H.W.

# Pitch designation

In this thesis, pitch notation follows the Scientific Pitch Notation convention. Here is a comparative list of this, and other, pitch labelling conventions.

## Comparative international pitch labelling conventions

Frequency (Hz)	Note name	Scientific Pitch Notation	Helmholz Pitch Notation	Another common pitch notation system
783.9	G	G <sub>5</sub>	g''	g <sup>2</sup>
740.0	F sharp	F# <sub>5</sub>	f#''	f# <sup>2</sup>
698.4	F	F <sub>5</sub>	f''	f <sup>2</sup>
659.3	E	E <sub>5</sub>	e''	e <sup>2</sup>
622.3	D sharp	D# <sub>5</sub>	d#''	d# <sup>2</sup>
587.3	D	D <sub>5</sub>	d''	d <sup>2</sup>
554.4	C sharp	C# <sub>5</sub>	c#''	c# <sup>2</sup>
523.2	C	C <sub>5</sub>	c''	c <sup>2</sup>
493.9	B	B <sub>4</sub>	b'	h <sup>1</sup>
466.2	A sharp	A# <sub>4</sub>	a#'	a# <sup>1</sup>
440.0	A	A <sub>4</sub>	a'	a <sup>1</sup>
415.3	G sharp	G# <sub>4</sub>	g#'	g# <sup>1</sup>
392.0	G	G <sub>4</sub>	g'	g <sup>1</sup>
370.0	F sharp	F# <sub>4</sub>	f#'	f# <sup>1</sup>
349.2	F	F <sub>4</sub>	f'	f <sup>1</sup>
329.6	E	E <sub>4</sub>	e'	e <sup>1</sup>
311.1	D sharp	D# <sub>4</sub>	d#'	d# <sup>1</sup>
293.7	D	D <sub>4</sub>	d'	d <sup>1</sup>
277.2	C sharp	C# <sub>4</sub>	c#'	c# <sup>1</sup>
261.6	<b>Middle C</b>	C <sub>4</sub>	c'	c <sup>1</sup>
246.9	B	B <sub>3</sub>	b	h
233.1	A sharp	A# <sub>3</sub>	a#	a#
220.0	A	A <sub>3</sub>	a	a
207.7	G sharp	G# <sub>3</sub>	g#	g#
196.0	G	G <sub>3</sub>	g	g
185.0	F sharp	F# <sub>3</sub>	f#	f#
174.6	F	F <sub>3</sub>	f	f
164.8	E	E <sub>3</sub>	e	e
155.6	D sharp	D# <sub>3</sub>	d#	d#
146.8	D	D <sub>3</sub>	d	d
138.6	C#	C# <sub>3</sub>	c#	c#
130.8	C	C <sub>3</sub>	c	c
123.4	B	B <sub>2</sub>	B	H

{Hirano, 1980 #167, p. 16}