

# **Acoustical, psychoacoustical and subjective assessment of Alvin Lucier's *I am sitting in a room***

## **Keywords**

Alvin Lucier; psychoacoustics; subjective; analysis.

## **Introduction**

Composed in 1969-70, *I am Sitting in a Room* is one of Lucier's earliest mature works. It is highly distinctive musical composition, in that a room is used as the musical instrument. The two most accessible published recordings of this work made by Lucier date from 1970 and 1980 (Lucier 1990, 1997). This paper examines the sound of these recordings in acoustical, psychoacoustical and subjective terms. Thus a work that is most commonly discussed in terms of its production is analysed here in terms of its sound, with an emphasis on perceptual qualities.

The musical score of *I am Sitting in a Room* simply consists of a page or so of text (Lucier 1995). The performer reads aloud the following text, recording their voice in a room of their choice.

I am Sitting in a Room different to the one you are in now.

I am recording the sound of my speaking voice and I am going to play it back into the room again until the resonant frequencies of the room reinforce themselves so that any semblance of my speech, with perhaps the exception of rhythm, is destroyed.

What you will hear, then, are the natural resonant frequencies of the room articulated by speech.

I regard this activity not so much as a demonstration of a physical fact, but more as a way to smooth out any irregularities my speech might have.

The initial recording is played back into the room using a loudspeaker, and re-recorded with a microphone. As the sound is repeatedly re-recorded thus, the text is progressively overwhelmed by tones produced by the acoustical characteristics of the transducer-room combination. The final sentence of the text refers to Lucier's stutter.

The score is written quite flexibly, with the performer permitted to use multiple rooms, languages and microphone positions. However, the recordings made by Lucier were each made in a single room and in English. The 1970 recording presents 15 generations, while the 1980 recording presents 32 generations.

This paper compares the two recordings, and characterises the evolving sound quality of each. It does not consider the linguistic and cultural meanings of the work. Results are best described as indicative, and generalisations of the findings beyond these two recordings are difficult to make. The paper can be read as a case study of the subjective and psychoacoustical analysis techniques applied.

## **Acoustical Considerations**

If the process described by Lucier's score were executed with a stable and linear system, then we would expect a constant transfer function (and impulse response) between generations. Considering the room acoustics alone, these can be understood as the combined effect of interference between the direct sound and early reflections, and the excitation of room resonances (or modes). If the room is approximately rectangular, axial (one-dimensional) modes should dominate (Cremer and Müller

1978), with each dimension forming a harmonic series of spectral peaks. In typical rooms, the average absorption coefficient varies between surfaces, allowing one set of axial modes to dominate. Interference between the direct sound and early reflections is also characterised by a harmonic series for each pair combination, but multiple pairs contribute to a transfer function in which harmonic series may be obscured (especially when dispersion and absorption as a function of frequency are considered for each path). Whatever the transfer function, it should be reinforced at each generation, so that ultimately a single tone remains, corresponding to the frequency of the transfer function's highest peak.

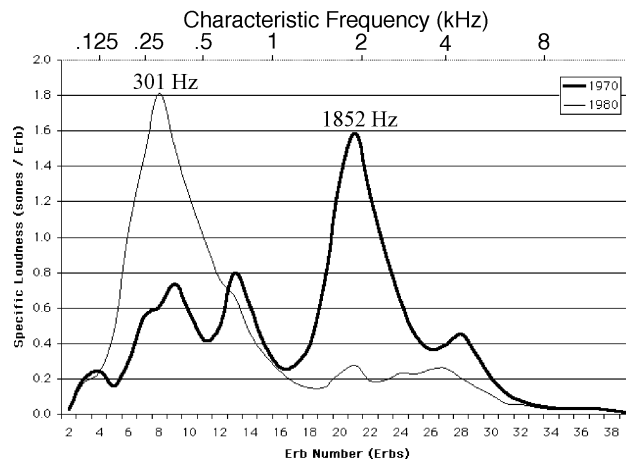
The reflections and resonances of the room add to the duration of each generation, because it takes time for the sound to build up and decay. This build-up and decay process also has the effect of smoothing the speech's temporal envelope.

Aspects of this idealised process are found in Lucier's recordings. However, the transfer functions between generations are far from constant. The presence of a (non-static) person in the room, movement of transducers, distortion in the electronic stage of the process, the use of variable electronic equalisation, and tape speed anomalies could reasonably contribute to the discrepancies. There is evidence for at least some of these factors in the recordings.

There is little evidence of harmonic series contributing to the 1970 recording transfer functions. By contrast, the 1980 recording has a clear harmonic series based on 37 Hz, with at least 30 harmonics evident starting from the second harmonic. This series is most probably caused by axial modes between opposite walls 4.6 m apart.

Apparent impulse responses were obtained through multiple averaged cross-correlation functions between generations. Apparent reverberation times (following Schroeder (1965)) are of the order of 0.25 s and 0.4 s respectively for the 1970 and 1980 recordings. These reverberation times are reflected by the respective lengthening of the generation durations from Generation 1 to 15 by a factor of 1.02 and 1.05 for the 1970 and 1980 recordings. The contrast in reverberation time is consistent with the hypothesis that room modes have a stronger influence in the latter recording, while the former is more strongly influenced by interference between the direct sound and early reflections.

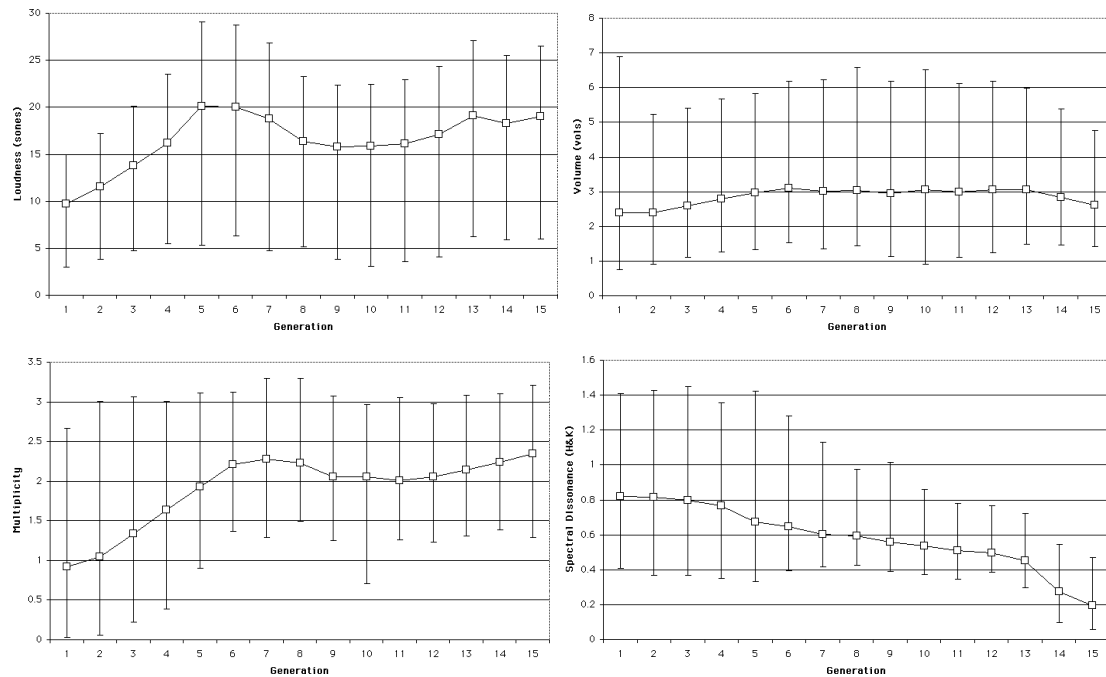
The later generations of the 1970 recording have a somewhat shrill timbre, caused principally by peaks between 1.5 kHz and 2 kHz. The 1980 recording is dominated by much lower frequencies. The subjective significance of the recordings' spectra is well-illustrated by Figure 1, where the specific loudness patterns (loudness attributable to specific auditory filters) is plotted against the psychoacoustical 'frequency' scale of Erb number, averaged over the entire duration of each recording. As the area under a specific loudness curve represents the total loudness, the very strong perceptual contributions of frequencies in the vicinity of 1900 Hz and 300 Hz to the 1970 and 1980 recordings respectively is evident.



**Figure 1.** Specific loudness patterns for the two recordings of *I am sitting in a room*, averaged over the entire recording duration.

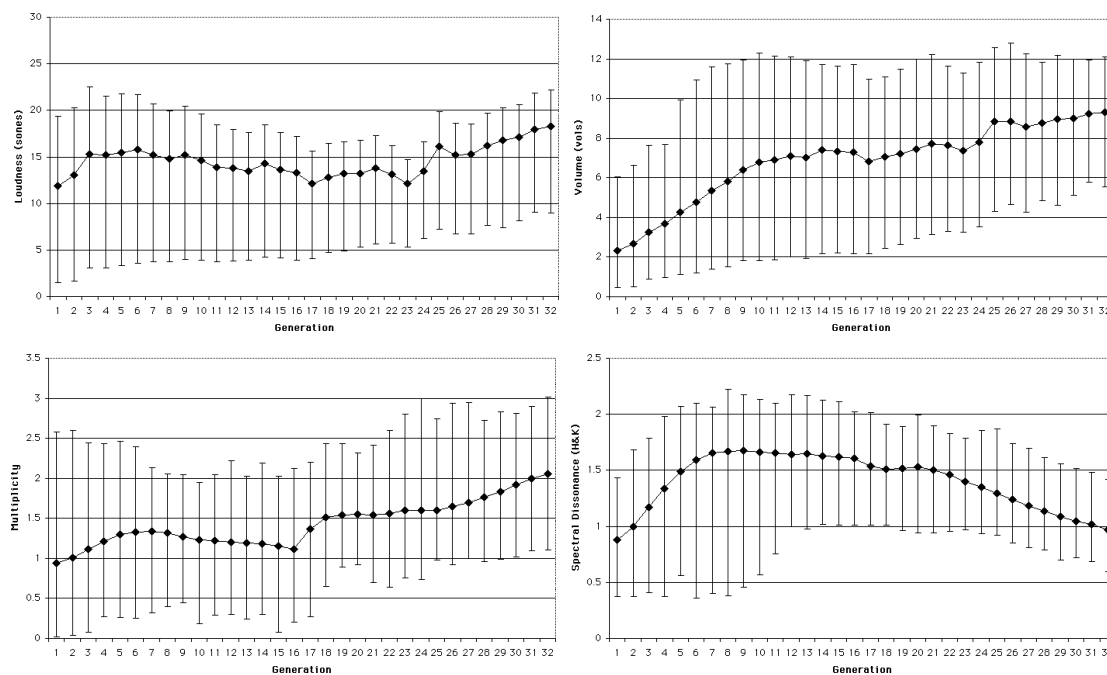
## Psychoacoustical Measures

The two recordings of *I am Sitting in a Room* were analysed using *PsySound 2.05* (Cabrera 2001), which applies numerous psychoacoustical models to sound recordings. The recordings were calibrated for an equivalent pressure level of 75 dB(A). Four psychoacoustical measures are illustrated: loudness (following Moore, Glasberg and Baer (1997)), volume (the apparent size of the sound, following Cabrera (2001)), pitch multiplicity (the number of distinct audible pitches, adapted from Parncutt (1989)) and spectral dissonance (the ‘noisiness’ of the sound, adapted from Hutchinson & Knopoff (1978)).



**Figure 2.** The mean, maximum and minimum loudness (top left), volume (top right) pitch multiplicity (bottom left) and spectral dissonance (bottom right) of the 1970 recording of *I am Sitting in a Room*.

The loudness and multiplicity series of the 1970 recording show a loose three-part structure. These measures increase significantly over Generations 1 to 5, stabilising and declining a little in the next five generations. The final five generations see a small recovery in these measures. This structure is accompanied by a continual decline in spectral dissonance, as the bandwidth of the sound narrows, and the noise components of speech are transformed into tones. The measures of sharpness (not illustrated, representing a scale from dull to sharp timbre) and volume yield little of interest.



**Figure 3.** The mean, maximum and minimum loudness (top left), volume (top right) pitch multiplicity (bottom left) and spectral dissonance (bottom right) of the 1980 recording of *I am Sitting in a Room*.

The 1980 recording measurement series reveal a change in the character of the recording after Generation 16. This appears to be due to the emergence of new sustained resonances above 500 Hz. It seems likely that the conditions of the room or of the electroacoustic system were altered at this point. Regardless of the reason for the change, its consequences are perceptually salient, and add much to the interest in this recording.

Another loose structural feature can be seen in the way the measures change rapidly initially, and then become more stable. Loudness stabilises at Generation 3, spectral dissonance forms a plateau at Generation 8, and sharpness and volume become more stable at around Generation 10. There is a sense that the process has an initial settling in period, followed by a relatively stable period, which is broken at Generation 16. The few generations following Generation 16 might be characterised as a second settling in period. The final few generations are characterised by increasing loudness, volume, tonalness and multiplicity, and decreasing spectral dissonance.

The variation in mean loudness is much smaller than for the 1970 recording, but the timbral measures of volume and sharpness have large variations with strong trends (sharpness decreases, while volume increases).

When measures are averaged over the entire recording durations, the two recordings have nearly the same loudness (due to the calibration). The 1970 recording is 1.6 times sharper than, and has less than half the volume of, the 1980 version. The 1970 recording has somewhat stronger pitch. The 1980 recording has more spectral dissonance (noisiness), but less tonal dissonance (dissonance between discrete tones).

## Subjective Assessments

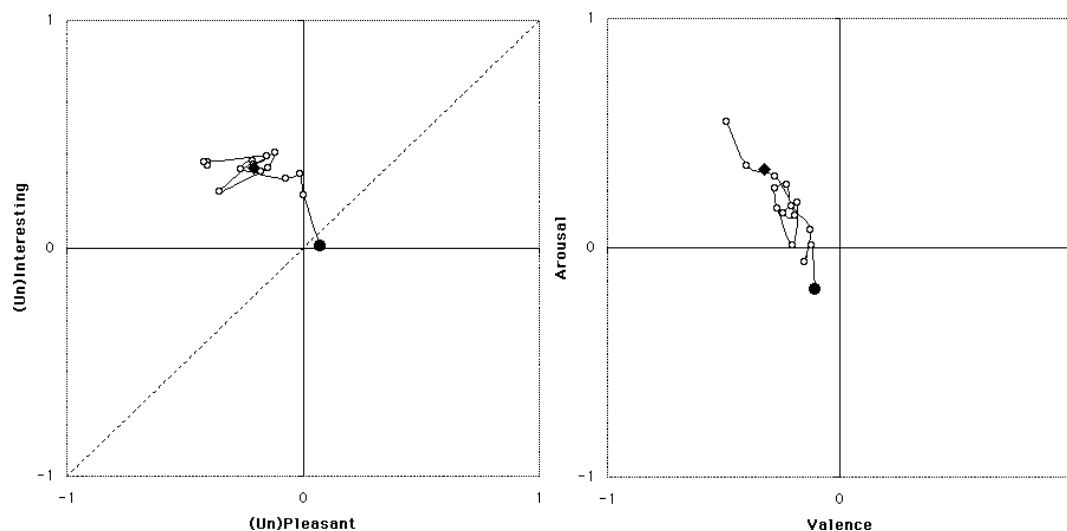
Thirteen subjects (all with substantial experience in music or audio) rated stimuli consisting of the first sentence of Lucier's text for each generation of the two recordings. The rating scales were bipolar, consisting of (Dis)Like, (Un)Interesting, (Un)Pleasant, Arousal and Valence. The first three are general, intellectual and sensory evaluation scales. The final two are emotion scales, with scale extremes represented by face icons (following Schubert (1999)).



**Figure 4.** Emotion face icons representing, from left to right, negative arousal, positive arousal, negative valence and positive valence.

The stimuli were presented in the order of the generations, with half the subjects assessing the 1970 recording first. The process of the piece, and the fact that it was a musical composition, were explained to the subjects before the experiment.

The subjective assessments of the sequences of stimuli can be visualised as trajectories in evaluation and emotion spaces. To reduce evaluation space to two dimensions, we draw on the fact that (Dis)Like ratings are approximately equal to the average of the other two evaluation scales.



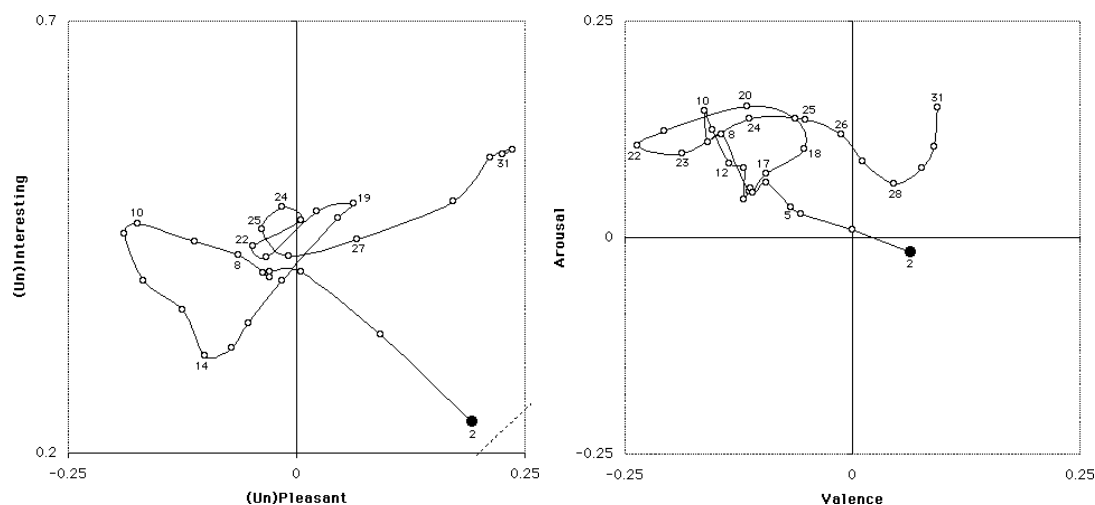
**Figure 5.** Subjective ratings of the 1970 recording of *I am Sitting in a Room*. The black circle represents Generation 1, and the diamond represents an overall rating made after all generations had been assessed.

Over the first few generations of the 1970 recording, the stimuli become more interesting and a little less pleasant (see Figure 5). Then the bulk of the stimuli are rated as moderately interesting and a little unpleasant. The final three stimuli are the least pleasant and among the most interesting. The trajectory in the emotion space is strong and simple, with Arousal tending to increase over the course of the generations while Valence decreases a little.

Trends are more obscure for the 1980 recording ratings: apart from the beginning, the ratings tend to form a dense cluster instead. The evaluation and emotion trajectories for the 1980 sequence become discernible by plotting a running average, and by enlarging the scale of the charts (Figure 6).

There is a general decrease in pleasantness to Generation 10, accompanied by a smaller increase in interest. Over the remainder of the piece, the generations become increasingly pleasant. Note that the loops in the early twenties are not significant. Interest declines a little to Generation 14, and then increases a little over the remaining generations. There is more variation in pleasantness than interest, and while interest ratings are positive, pleasantness ratings are almost equally positive and negative.

Emotion ratings are very close to neutral, with a bias towards positive arousal and negative valence. Most of the variation in the Arousal ratings is not significant. Valence initially declines, but then increases in the final few generations.



**Figure 6.** Running average of subjective ratings of the 1980 recording of *I am Sitting in a Room*. Three consecutive generations were used in the running average.

### Psychoacoustical measures related to subjective ratings

The problems of serial correlation, low significance and confounded independent variables preclude a detailed analysis of relationships. Nevertheless, the results suffice for broad observations.

As there are simple trends in the 1970 recording ratings (especially for (Un)Pleasant, Arousal and Valence), any psychoacoustical measure that also follows a trend shows a significant correlation with these measures. Loudness, spectral dissonance and multiplicity all have such correlations. The remaining two evaluation scales do not

have strong correlations with any psychoacoustical measure (although there is weak evidence of inverted-U relationships with loudness and multiplicity). Correlation coefficients are shown below:

1970 Stimuli	(Dis)Like	(Un)Interesting	(Un)Pleasant	Arousal	Valence
(Un)Interesting	-0.06				
(Un)Pleasant	<b>0.74</b>	-0.52			
Arousal	-0.70	0.61	<b>-0.91</b>		
Valence	<b>0.82</b>	-0.43	<b>0.88</b>	<b>-0.89</b>	
Loudness	-0.39	<b>0.71</b>	<b>-0.85</b>	<b>0.81</b>	-0.63
S. Dissonance	<b>0.75</b>	-0.57	<b>0.82</b>	<b>-0.84</b>	<b>0.91</b>
Multiplicity	-0.48	<b>0.74</b>	<b>-0.84</b>	<b>0.82</b>	<b>-0.74</b>

That Arousal is not related to loudness in the 1980 series almost certainly reflects the much more modest range of loudness therein. However, the ratio of the maximum to minimum loudness for each generation (which can be called loudness contrast), along with multiplicity, is weakly related to the (Un)Interesting ratings. The relationship between spectral dissonance and the (Un)Pleasant and Valence scales is reversed from the 1970 recording.

1980 Stimuli	(Dis)Like	(Un)Interesting	(Un)Pleasant	Arousal	Valence
(Un)Interesting	0.23				
(Un)Pleasant	<b>0.83</b>	0.06			
Arousal	-0.12	0.67	-0.19		
Valence	0.61	-0.02	<b>0.83</b>	-0.18	
Loud. Contrast	0.12	<b>-0.77</b>	0.15	-0.59	0.24
S. Dissonance	-0.69	0.00	<b>-0.85</b>	0.14	<b>-0.82</b>
Multiplicity	0.38	<b>0.76</b>	0.44	0.51	0.31

## Conclusions

We can conclude, on the basis of several findings, that the 1980 recording is much more subtle (from a listener's perspective) than the 1970 recording of *I am sitting in a room*. This is borne out by the more complex patterns found in the psychoacoustical measures and by the much more chaotic but limited subjective assessments. The near-constant Arousal responses reflect the tendency for subjects to favour the less obvious subjective responses when assessing the 1980 sequence. The 1980 sequence involves markedly lower frequencies (even in Generation 1), is twice as long and involves significant changes in course (whereas the 1970 sequence more or less follows a single course).

Neither recording can be described as emotional. The emotion space trajectory of the 1970 recording is easily understood as Arousal being affected by loudness (which is generally the pattern for studies of this type). In the absence of any positive valence cues, the louder sounds become more annoying, and hence receive more negative Valence ratings. Thus, the 1970 recording is emotionally simple. The 1980 recording emotion ratings are all clustered near the centre of the emotion space, indicating an emotionally neutral (or weak) character. In contrast to the 1970 recording, most variation is in the Valence dimension.

With only two recordings, we cannot generalise about the significance of overall comparisons, including those concerned with the room acoustics. However the more positive evaluations of the 1980 recording, which relies more on room resonance, is consistent with other experiments concerned with the subjective evaluation of sustained room resonance (Cabrera 1997).

Although this analysis of *I am Sitting in a Room* has not addressed any of the symbolic or cultural issues raised by the work, it has shown that the sound itself is rich for analysis, especially when approached from a listener's perspective. The simple progressive structure of the work makes this type of analysis particularly interesting. Clearly similar approaches to analysis could be applied to almost any type of music, as long as the music involved sound. The relevance of such a method depends on the importance of the non-symbolic aspects of sound, and the presence of structural features to give the analysis direction.

## References

- Cabrera, D. 1997. "Stimulation and Simulation of Rectangular Room Resonant Modes." *Australian Computer Music Association conference*, Auckland: 14-23.
- Cabrera, D. 2001. "'PsySound': A Computer Program for the Psychoacoustical Analysis of Music." *MikroPolyphonie* 5, <http://farben.latrobe.edu.au/mikropol/>.
- Cabrera, D. 2001. "The size of sound: auditory volume reassessed." *MikroPolyphonie* 5, <http://farben.latrobe.edu.au/mikropol/>.
- Cremer, L. and H. Müller. 1978. *Principles and Applications of Room Acoustics*. London, Applied Science Publishers.
- Hutchinson, W. and L. Knopoff. 1978. "The Acoustical Component of Western Consonance." *Interface* 7:1-29.
- Lucier, A. 1990. *I am Sitting in a Room* (1980 sound recording). New York, Lovely Music.
- Lucier, A. 1995. *Reflections / Reflexionen*. Köln, MusikTexte.
- Lucier, A. 1997. "I am Sitting in a Room" (1970 sound recording). *With a minimum of means*. Målsryd, Content.
- Moore, B., B. Glasberg and T. Baer. 1997. "A Model for the Prediction of Thresholds, Loudness and Partial Loudness". *Journal of the Audio Engineering Society* 45: 224-40.
- Parncutt, R. 1989. *Harmony: a Psychoacoustical Approach*. Berlin, Springer.
- Schroeder, M. 1965. "New Method of Measuring Reverberation Time." *Journal of the Acoustical Society of America* 37:409-12.
- Schubert, E. 1999. *Measurement and Time Series Analysis of Emotion in Music*. Thesis (PhD), University of New South Wales.