Proceedings of the Stockholm Music Acoustics Conference, August 6-9, 2003 (SMAC 03), Stockholm, Sweden

TIMBRAL SEMANTICS AND THE PIPE ORGAN

Alastair C. Disley and David M. Howard

Department of Electronics
The University of York

http://aldisley.care4free.net/

ABSTRACT

Words used to describe timbre are usually difficult to define or relate to a measurable phenomenon. This paper attempts to establish if any such words are objective with a common understanding, or if they are all subjective. The pipe organ is used because it is both a complex timbral synthesizer and a non-electronic instrument from which repeatable samples can be taken.

A number of adjectives were gathered from English speaking subjects. The most common words were selected and used without specific definition as rating scales in comparative listening tests. Comparison with spectral analyses revealed possible cues for some words, and audio examples were synthesized to test these theories. Some adjectives have emerged as having degrees of common understanding across a majority of subjects.

1. INTRODUCTION

Adjectives are frequently used to describe sound, but little research has been done into how the words people use relate to measurable acoustic phenomena. It is not clear whether any such words are subjective, with meaning only to a given person, or objective across a number of people with common understanding. Previous research on adjectives describing aural and other sensations is summarized in Rioux [1]. Rioux also gathered descriptive words from an organ-builder [2], categorized them and conducted listener tests [3]. That research involved single pipes, and the words gathered mostly referred to pipe speech.

The use of pipe organs in both this and Rioux's experiments is because it is one of the few acoustic instruments to be an additive timbral synthesizer, where the player controls the timbre by combining pipes of different pitches and individual timbres into one chorus. Thus with its variety of sounds, different timbres can be synthesized without going outside the realms of a single acoustic musical instrument. It is also possible to take repeatable samples from a pipe organ because of the nature of its sound generation and control, something that is impossible on most other acoustic instruments.

Listeners do not usually hear pipes in isolation, so in order to gain meaningful results applicable to common situations, all audio samples of pipe organs used in this experiment were of complete choruses of pipes, played in chords as might be heard in a hymn. The words to be used in this test were chosen by frequency of occurrence in a wide scale listening experiment, to ensure that they were words that would be relevant to the listeners.

2. EXPERIMENTAL PROCEDURE

2.1 Adjective gathering

Forty-five listeners were played recordings from four different organs and asked to describe what they were hearing with no other prompting. The adjectives they used were then categorized and counted, including all derivatives, and the most frequently used were selected for further study. Listeners were players of or keen listeners to the pipe organ, gathered from a request on the internet mailing list piporg-l.

The most common words used were: "bright", "clear", "flutey", "reedy", "thin", "full", "balanced" and "warm". "Reedy" was not studied further as organ pipes are also divided into two types, flue and reed, which could cause confusion. In all experiments described in this paper, listeners were not given definitions of any terms to avoid the disturbance of any presence or lack of prior common understanding.

2.2 Initial experiments

As part of a seminar on this research, 75 people took part in a large scale listening test using the terms "bright" and "clear", which were the most common. All listeners were English speakers, most as their first language, and self-divided into three groups according to their musicality (none, some or lots). They were also asked if they were regular listeners to or players of the pipe organ. In this and all other experiments in this paper, testing took the form of comparison tests with listeners marking a scale from 1 to 11, such that an answer of 6 would indicate either "neither" or "both equal".

In the first case, the principal chorus of an organ in a reverberant acoustic was played with and without its mixture stop. A mixture is a stop that adds several higher harmonics to the

ensemble. Listeners on average thought that the mixture added some slight brightness, but made no difference to clarity.

In the second example, the 8' principal choruses of two different organs were played with their mixtures, one a 19th century instrument (the Schulze in Doncaster Parish Church) in a moderately reverberant environment, and the other a 20th century instrument (the Grant, Degens and Bradbeer at York University) in a drier acoustic. The latter instrument was thought to be both brighter and clearer, and Fourier analysis revealed that its higher harmonics were of greater amplitude, suggesting that as might be expected, the perception of brightness is directly related to the strength of higher harmonics. Clarity might have an inverse relationship with the amount of reverberation.

Results were also analysed against amount of musical training. Although there was little difference in average results between the groups, as musical training decreased, standard deviations increased, indicating a decrease in common understanding of words. Interestingly, the subset of listeners who indicated familiarity with the pipe organ had the largest standard deviation of all.

The next five most common adjectives, "flutey", "thin", "full", "balanced" and "warm", were used in a listener test conducted over the internet with 23 subjects, all of whom were familiar with the pipe organ as regular performers or enthusiastic listeners. All were native English speakers: thirteen were American, eight were English and two were Canadian. All but one were male, and age and geographical data were collected, but these appeared to have no significant bearing on results. Samples were presented in uncompressed 16-bit stereo WAV format, sampled at 22kHz, as a compromise between download time and maximum resolution. Following comments from previous such tests, reverberation was not cut off early.

Four recordings of different organs were used, with in each case the same hymn-type extract played on the 8' principal chorus. Acoustic analyses of the four organs are given below, with steady-state Fourier analyses (4096 point Hamming window) simplified to harmonics only, and shown at ten points across the keyboard range. The principal chorus of all four organs includes 8', 4' and 2' stops. Heslington (figure 1) includes a Twelfth, all other organs include a Mixture. St. Chad's (figure 3) is in a very reverberant acoustic, and like Heslington is a 19th century instrument. Port Sunlight (figure 2) is an early 20th century instrument and has a more muted chorus typical of its time. The University of York organ (figure 4) is also typical of its later 20th century date, by which time the fashion was for stronger mixtures and weaker 8' stops.

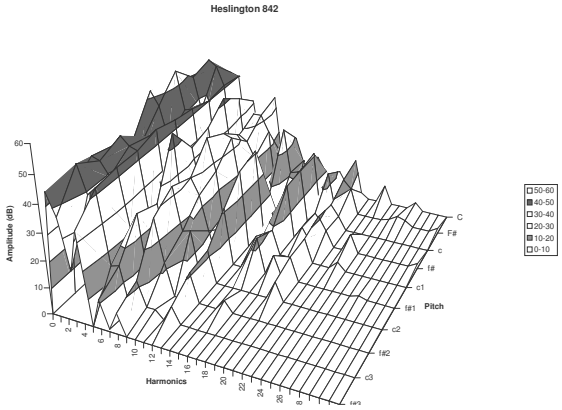

Figure 1: *Principal chorus of Heslington*

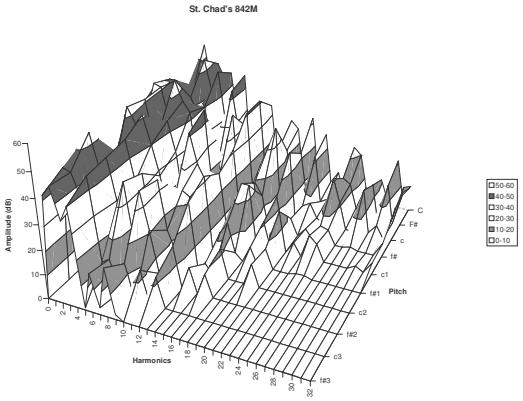

Figure 3: *Principal chorus of St. Chad's*

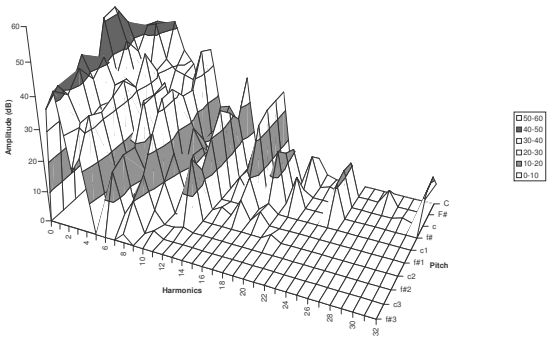

Figure 2*: Principal chorus of Port Sunlight*

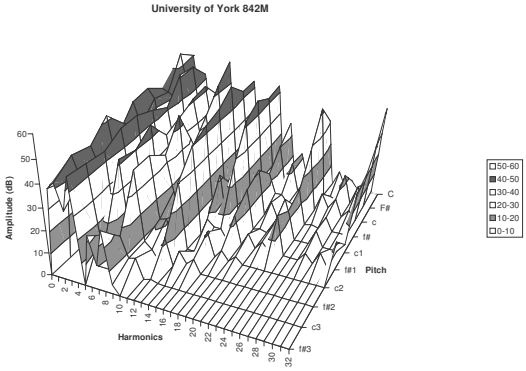

Figure 4: *Principal chorus of University of York*

2.3 Results and analysis

For some of the adjectives tested, the mean in many comparisons was near the central value of 6. For the term 'balanced", the mean was always within 0.91 of that central value, and standard deviations varied from 2.14 to 2.61, indicating some difference of opinion on what 'balanced" meant, but also that no organs in the test were perceived to be unbalanced. Full again had a mean always within 0.96 of 6, and standard deviations varied from 2.23 to 3.01. Full as an adjective seems to have less common understanding than any other word used.

The word 'thin" produced its strongest results (means biased 1.83 and 1.52 from the centre, standard deviations 2.1 and 2.04 respectively) when it was applied to the University of York instrument compared with Port Sunlight and St. Chad's. Possible explanations of this from the acoustic analyses are the stronger high and weaker low harmonic components of the University of York's ensemble, which each require separate testing.

The word "flutey" appeared to follow the instrument with weaker higher harmonics. The strongest bias from the centre was 2.09, with a standard deviation of 2, when the Jack Lyons organ was compared with the Port Sunlight instrument. Standard deviations varied from 1.88 to 2.15, suggesting a degree of uncertainty among participants.

The word 'warm" produced the greatest agreement among subjects, with standard deviations varying from 1.34 to 1.5 in tests involving the Jack Lyons instrument, which was clearly perceived to be not warm in comparison with other instruments. The other three instruments were perceived to be warmer by between 2.26 and 2.65 from the central value. However it appears that warmth is not solely a results of less strong higher harmonics, as St. Chad's was perceived to be slightly warmer than Heslington (by 1.52, with a standard deviation of 1.65). This may be because of St. Chad's greater reverberation.

These results suggest that an organ with stronger higher harmonics will be perceived as less flutey and less warm. Perception of a sound as 'thin" may be related to weaker lower harmonics, and/or stronger higher harmonics. Warmth also seems to be affected by the strength of higher harmonics, but may also be affected by the amount of reverberation present. In cases where the amount of reverberation is similar, the results for warmth and flutey are also similar.

2.4 Verification of theories

To develop these theories, a further set of comparative listener tests was developed for the same subject group, of whom 14 took part. The samples were developed to establish whether adding stronger higher harmonics made people describe the sound as less 'warm", less " flutey", and more 'thin". The samples also needed to test whether reducing the strength of lower harmonics produced a description of more 'thin". Thirdly they needed to test whether adding reverberation produced an increased perception of 'warmth". The words 'balanced" and 'full" were rejected at this stage, and replaced by the terms 'bright" and 'clear" that had been used in the first listening test.

In order to place an identical sound in two acoustic environments, the four samples for this test were produced on an electronic organ, a Wyvern B235 using Bradford synthesis technology [4]. The two reverberant environments were simulated on a Yamaha S-REV 1 convolution-based device. The more reverberant acoustic used its preset program 'St. John the Divine, no. 3", with the other three samples using preset 'Warm Wooden Church no. 1". Timbral manipulation of the samples needed to be within the range of sounds that a pipe organ might produce, so synthesis of timbres was achieved by addition and removal of stops. Stonger higher harmonics were produced by adding a mixture to a basic chorus of 8', 4' and 2' principals. Replacing the 8' Principal with an 8' Stopped Flute for the fourth sample produced a reduction in lower harmonics.

Results from this final test were more consistent, with an average standard deviation of 1.74, compared with 2.18 for the previous test. When the sample with the 8' stopped flute was compared with that with the 8' Principal, it was described as more 'thin" (by 2.08 from the mean in both more and less reverberant examples). "Flutey" gave significant results when a pair of samples included that with the mixture, and means varied between 1.46 and 2.38 towards the samples without the mixture. When the mixture was added, perception of brightness clearly increased, by 3.54 and 3.62 from the mean both with and without reverb (standard deviations 1.13 and 0.96 respectively).

Subjects gave more varied answers when the sample including the mixture was compared with the sample containing the stopped flute – although the average perception of brightness was 2.00 towards the mixture sample, the standard deviation was 2.68. The perception of 'warm", like " flutey", was significant in those samples compared with the mixture. The stopped flute sample was perceived to be slightly less warm than those other samples that lacked a mixture. In all cases clarity appeared to follow brightness, although to a lesser extent, and with a slight bias away from the more reverberant samples and towards both the mixture and stopped flute samples.

3. CONCLUSIONS

It was suggested in section 2.2 that the addition of a mixture would increase the perception of brightness, and that clarity was related to less reverberation. The tests in section 2.4 appear to support these theories. Listeners appear to have a common understanding of brightness as being an increase in strength and/or number of higher harmonics. The perception of the sample with reduced lower harmonic strength suggests that this increase is measured relative to the strength of lower harmonics. It is more difficult to make such a declaration for clarity, as reverberation is not a simple linear phenomenon, but there is at least a distinct degree of common understanding as to its meaning.

'Full" appears to have no common understanding among subjects, and 'balanced" was also less well understood, although

both may have greater understanding outside the limitations of this experiment. "Flutey" was used to describe instruments with weaker higher harmonics when they were compared with a mixture, but was less consistently used when neither sample had a mixture.

In the final test, "Thin" did not appear to be related to an increase in strength of higher harmonics, but to weaker lower harmonics in the sample including the Stopped Flute. "Warm" was also related slightly to greater strength of lower harmonics, but was most strongly associated with the sample in a pair exhibiting weaker high harmonics. Perception of a sound as "warm" also increased with the amount of reverberation.

In summary, most of the common words provided by the initial gathering of adjectives have proved to have degrees of common understanding for certain circumstances. Some words, particularly "bright" and "warm", have good common understanding across all circumstances used in this report, within a set of listeners for whom English is a first language and who are familiar with the pipe organ.

Although reverberation has been mentioned as having an effect for several words studied, the complex nature of reverberation means that its precise effect in any circumstance is hard to quantify without further study. The first large-scale experiment indicated that as musical training decreased common understanding of terms also decreased, which suggests that these terms might be introduced subconsciously as part of the growth of a person's musical understanding. Future research could also investigate how without specific formal training, the common understanding identified in this paper been gained.

4. REFERENCES

[1] Rioux, V., Sound Quality of Flue Organ Pipes, PhD thesis, Chalmers University of Technology, Gothenburg, 2001, Chapter II.

[2] Rioux, V. and Västfäll, D., "Verbal descriptions of flue organ pipe sounds (I): a qualitative analysis", in [1], pp. 99-115.

[3] Rioux, V. "Methods for an objective and subjective description of starting transients of some flue organ pipes – integrating the view of an organ-builder", Acustica, Vol. 86, No. 4, pp. 634-641, 2000.

[4] Comerford, P. J., "Simulating and organ with additive synthesis", Computer Music Journal, Vol. 17, No. 2, pp. 55-65, 1993.