RECENT STUDIES ON THE EFFECT OF SIGNAL FREQUENCY ON AUDITORY VERTICAL LOCALIZATION

Densil Cabrera
Sam Ferguson
Steven Tilley

Acoustics Research Laboratory
Faculty of Architecture
University of Sydney
Sydney NSW 2006, Australia
densil@arch.usyd.edu.au

Masayuki Morimoto

Environmental Acoustics Laboratory
Faculty of Engineering
Kobe University
Nada, Kobe, 657, Japan
mrmt@kobe-u.ac.jp

ABSTRACT

The authors have been involved in a series of studies showing that the fundamental frequency of a complex tone, the center frequency of a noise band, or the cut-off frequencies of simultaneously presented high- and low-passed noise bands, systematically affect the elevation of auditory images – such that high frequencies are associated with high elevations, and low frequencies with low elevations. These studies show this effect not just for median plane localization, but in some circumstances even for sources on the aural axis. This paper reviews the findings of these studies, and considers their implications and applications, including for auditory display.

1. INTRODUCTION

The use of the terms 'high' and 'low' to describe the pitch and frequency scales is common to most languages, and may reflect various associations between pitch and spatial elevation. These associations have been commented on previously [for example, in references 1, 2, 3]. The use of this metaphor is all the more intriguing in the light of experimental studies that appear to show a systematic perceptual correspondence between the frequency scale and vertical auditory localization in some circumstances. The first such study was conducted by Pratt [1] (published in 1930), and so this vertical localization effect is referred to as Pratt's effect.

Pratt [1] found that subjects localized five pure tones at octave intervals between 256 Hz and 4096 Hz to elevations from low to high in frequency order. For this experiment, Pratt used a small loudspeaker located behind the subject's head. Pratt concluded that the effect was not simply due to association, but was an inherent spatial quality of the tones, although there seems to have been little basis for this conclusion in the results. From these results, Pratt was intrigued by the possibility that the vertical localization of music could vary with the music's pitch. Although Dimmick and Gaylord [4] failed to duplicate Pratt's results, Trimble [5] found the Pratt's effect clearly evident in four out of five subjects for nine tones from 500 Hz to 3950 Hz. The sound for this experiment came from two loudspeakers on either side of the subject's head. Trimble also found a correlation between image elevation and pitch for tone sequences with pitch ascending or descending. However, he concluded that Pratt's effect was associative rather than innate.

More than thirty years later, Roffler and Butler [2] verified Pratt's effect, and examined the extent to which it is influenced by visual cues, body orientation and linguistic associations. They found the effect largely preserved for listener orientation (rather than the direction of gravity) when subjects lay on their backs or sides. Even though the proximity of the visual scale used to rate elevation affected the ratings, they also found Pratt's effect present in congenitally blind subjects. In an attempt to exclude linguistic associations, they tested 4- and 5-year old children, who appeared not to associate pitch with height in language – yet Pratt's effect was observed for those subjects.

These experiments were conducted in the absence of binaural difference cues for source localization (although it is possible that the non-anechoic conditions of some experiments may have introduced binaural differences). In natural listening this occurs on the median plane (the vertical plane that is orthogonal to the aural axis, and hence contains the directions front, above, behind and below for a listener). Further studies have shown that source-related cues for localization around this plane are conveyed by spectral information, particularly at high frequencies. In natural listening, these cues are generated by the interaction between sound and the listener's body – especially their pinnæ. The most important spectral cues for vertical localization are above 7 kHz [6]. However, much weaker cues may extend as low as 700 Hz (due to shoulder reflections) [7]. For spectral cues to be effective, the sound signal must have spectrally dense energy extending to the cue frequency range.

Cones of confusion are formed by constant angles around the aural axis, and are so named because the binaural difference cues are approximately constant on any cone of confusion – meaning that localization discrimination around the cone can be difficult [8]. The median plane could be considered to be the largest cone of confusion. Like the median plane, angular discrimination around any cone of confusion relies on spectral cues.

Morimoto and Aokata [9] examined how spectral cue manipulations could be used to move auditory images of white noise stimuli from the source position to other positions in the upper hemisphere of auditory space. By artificially filtering signals, the auditory images could be moved around the cones of confusion, since the binaural difference cues did not change. The effect of low-pass filtering the stimuli at 5 kHz is of particular interest in the present context – this caused the auditory image of an elevated source to be localized either behind or in front of the subject, on the same cone of confusion as the source. Similarly, Roffler and Butler [6] find that low-pass filtering noise at 2 kHz yields image elevations below the horizontal (using a vertical sound source array in front of the subject), as shown in Figure 1. Hence, like the studies of Pratt's effect that use tones, these results for filtered noise associate low image elevation with low frequency. The implication from Morimoto and Aokata is that beyond the median plane Pratt's effect follows the cones of confusion, and so becomes increasingly restricted for sound sources approaching the aural axis.

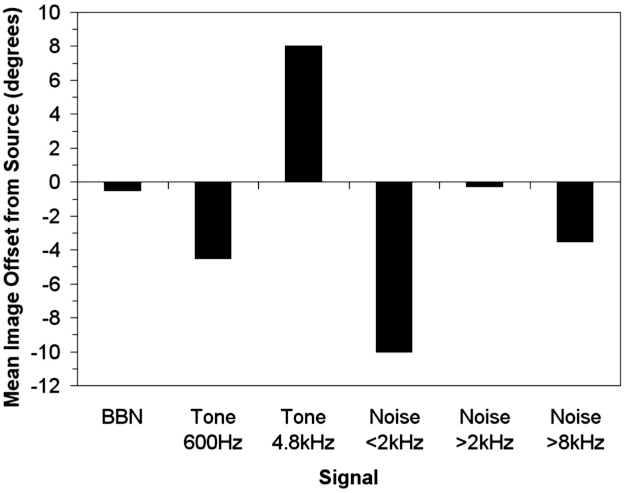

Figure 1. *Selected results of Roffler and Butler [6], showing the difference between image and source angle for broadband noise, two pure tones, and three limited noise bands. Results are for sound sources in the median plane in front of the subject (vertical source angles from -13° to +20°).*

Other studies also make observations relevant to Pratt's effect for noise stimuli. Blauert [10] observed that auditory images of high frequency noise bands are elevated. Gardner [11] and Noble [12] observe that low frequency noise bands tend to yield images with low elevations. Blothe and Elfner [13] and Gardner [11] note that the apparent 'pitch' of broadband noise appears to increase with source elevation.

2. RECENT STUDIES

In recent years, the authors have investigated further aspects of Pratt's effect, partly with a view to applications. These studies are summarized in the following sections, with details of the study methods and results given in the cited papers.

2.1. Pratt's Effect for Octave Noise Bands

Cabrera and Tilley [14] investigated localization and image size effects for octave bands of noise (centered on 125 Hz, 500 Hz, 2 kHz and 8 kHz), broadband pink noise, and low-passed pink noise (at 3 kHz), at two loudness levels (64 phons and 84 phons). A visually hidden vertical array of five sound sources was in front of the subject (this experiment was in the median plane only), with source elevations of 0°, ±7.9° and ±15.6° with respect to ear height (the angle θ in Figure 2). Subjects reported the position of the top, bottom, left and right edges of the auditory image in relation to a vertical gridded screen in front of them. The image center and size was determined from these values.

In terms of Pratt's effect, this study addresses the question of whether un-pitched sound stimuli follow a similar localization pattern to the tones studied previously. Results show auditory image center elevations in frequency order (Figure 3). There is some effect of source position for the 8 kHz octave band and the low-passed noise. The sound sources for the broadband noise are localized accurately, and low-passed noise is localized less accurately. As the same trend is observed for both pitched and un-pitched stimuli, Pratt's effect appears to be related to spectral envelope rather than specifically to pitch.

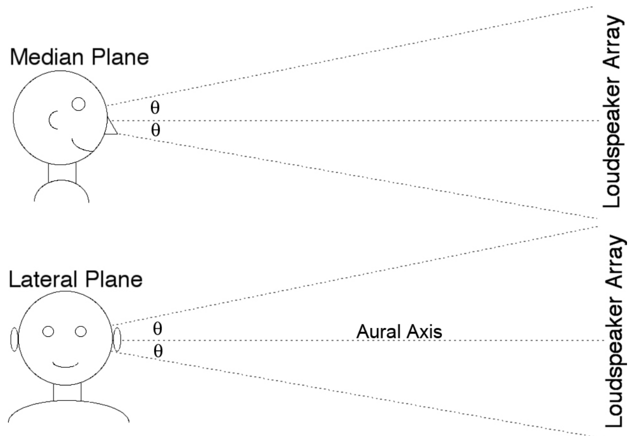

Figure 2. *General configuration for the recent vertical localization studies summarized in this paper.*

Although the results displayed in Figure 3 appear to show greater image elevation for low-passed noise than broadband noise, the mean differences are not significant in Scheffe tests. However, the fact that the low-passed noise result is not lower than the broadband noise result is not consistent with Roffler and Butler's result for noise low-passed at 2 kHz.

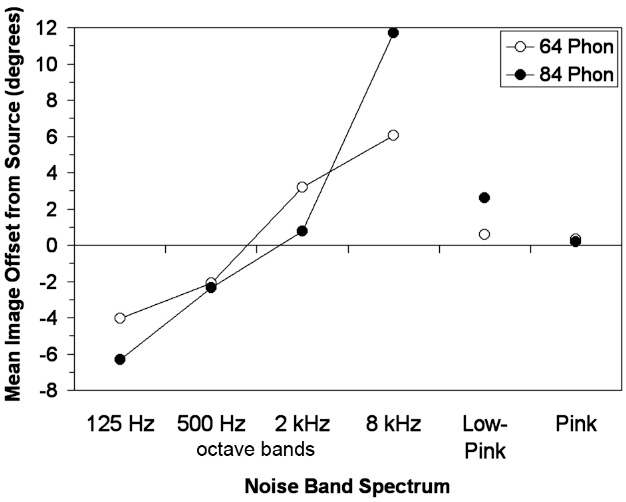

Figure 3. *Vertical localization results of Cabrera and Tilley [14] for octave bands of pink noise, low-passed pink noise, and broadband pink noise.*

Another question of interest is whether the loudness of stimuli interacts with Pratt's effect. Davis and Stephens [15] find that vertical localization accuracy increases with presentation level for both speech and white noise stimuli. However, the results of this study do not show a significant localization effect for loudness (based on analysis of variance), although there is a non-significant difference for the 8 kHz octave band. By contrast, a 20 phon loudness level change has a strong effect on auditory image size (both vertical and horizontal spread).

2.2. Pratt's Effect for Simultaneous Noise Bands

Based on the results of Cabrera and Tilley, an intriguing hypothesis was that Pratt's effect might apply not only to individual noise bands, but also to simultaneous bands. Furthermore, if the effect was based on gross spectral envelope, then it could find application in systems where arbitrary audio signals are filtered into discrete bands prior to reproduction by multiple sources. This is the situation with most loudspeaker systems with a wide frequency response.

Previous studies into simultaneous localization of pairs of noise bands indicated that it might be difficult to perceptually segregate two noise sources in the median plane. Best *et al.* [16] found that segregation did not occur for two broadband noise sources in the median plane. Morimoto *et al.* [17] found that when a pair of low-passed and high-passed noise bands (crossover of 4.8 kHz) were presented in the median plane, there tended to be no segregation, and that the angle of the high-passed source determined the auditory image angle.

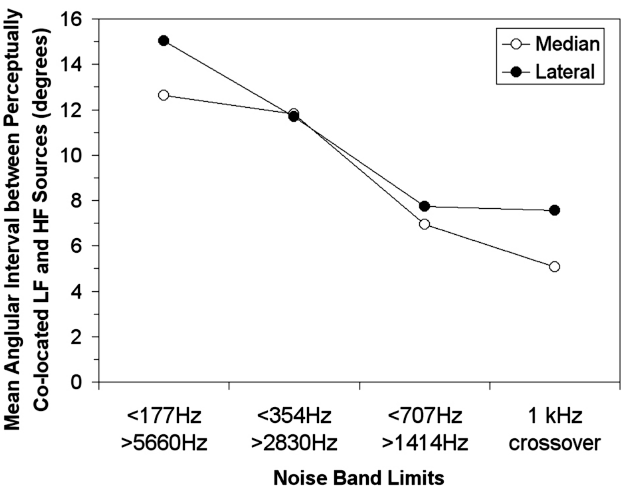

Figure 4. *Difference between low frequency and high frequency source angles for perceptual co-location (for both median and lateral planes), from the results of Cabrera and Tilley [18].*

A pilot study [18], using pairs of noise bands separated by four possible spectral gaps, gave an indication that the spectral interval between sources might correspond simply to vertical spatial interval. This pilot study was conducted with the experimenter moving two small hidden loudspeakers (a 'tweeter' and a 'woofer') between stimulus presentations, until the subject was satisfied that the sources were both at ear height. Spectral gaps of 5, 3 1 or 0 octaves were used. Both the median plane and lateral plane (with the loudspeakers to the subjects' side) were tested. Results showed that the woofer was consistently positioned above the tweeter for perceptual co-location (in both median and lateral planes). The spectral interval between noise bands was simply related to the spatial interval between the perceptually co-located sources, in accord with Pratt's effect. Figure 4 summarizes these results, showing that on average, the low frequency source was positioned above the high frequency source for perceptual co-location, and that this trend occurred in both median plane and lateral plane presentation. This experiment provided the impetus for a more careful series of experiments investigating this phenomenon, as reported by Ferguson and Cabrera [19].

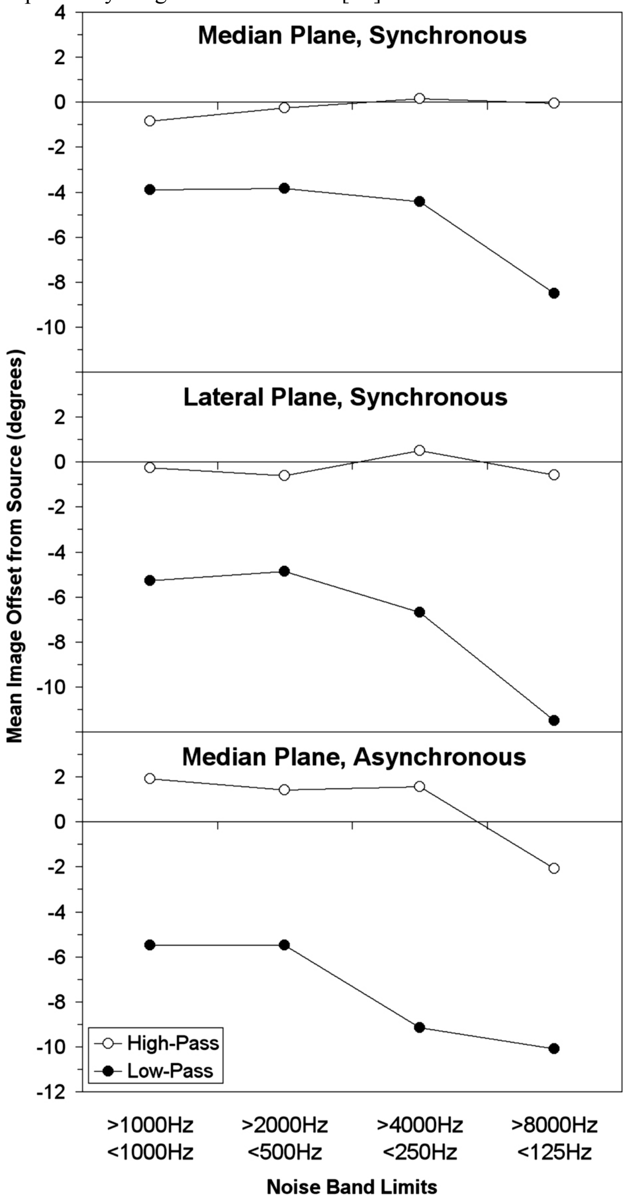

Figure 5. *Localization offset for simultaneous bands of high-passed and low-passed noise, from the experiments of Ferguson and Cabrera [19].*

Ferguson and Cabrera [19] conducted three main experiments. In the first, pairs of low-passed and high-passed noise bands were presented from a fixed loudspeaker array, very similar to that used by Cabrera and Tilley [14], with sources from -15° to +15°. The subjects' task was to localize the 'woofer' and the 'tweeter', using a vertical visual scale. There were four spectral intervals between noise bands (125 Hz – 8 kHz, 250 Hz – 4 kHz, 500 Hz – 2 kHz, and no interval with a crossover at 1 kHz). Every combination of woofer and tweeter position in the 5-loudspeaker array was tested. Subjects were tested with the array in either the median plane or the lateral plane. Results showed that tweeter images tended to be localized accurately, while the woofer images tended to be localized below the tweeter image. The woofer position had no effect on the woofer image elevation. The spatial interval between woofer and tweeter images increased with the spectral interval. There was no significant difference in the results for median and lateral planes (Figure 5).

In the second experiment, the noise bands were presented asynchronously – with the woofer signal delayed by 100 ms relative to the tweeter (each noise band was 200 ms long). Only median plane presentation was tested. The array of five loudspeakers spanned -20° to +20°. Results were similar to the first experiment, except that the woofer position did influence the woofer image position (for low-pass frequencies of 1 kHz and 500 Hz only).

In the final experiment, music recordings were presented from every combination of woofer and tweeter positions, using a crossover of 1 kHz. Excerpts from a chamber orchestra recording of Shostakovitch's 8th String Quartet (third movement), and James Brown's *The Big Payback* were used as the music samples. Subjects were required to localize the woofer and tweeter. The ±15° five-loudspeaker array was used. Results (Figure 6) showed that the woofer was localized, on average, below the tweeter. Both the woofer and tweeter positions had some influence on the woofer and tweeter image elevations. In general terms, the results are similar to those of the second experiment (asynchronous noise bands) except that tweeter localization was less accurate with music stimuli.

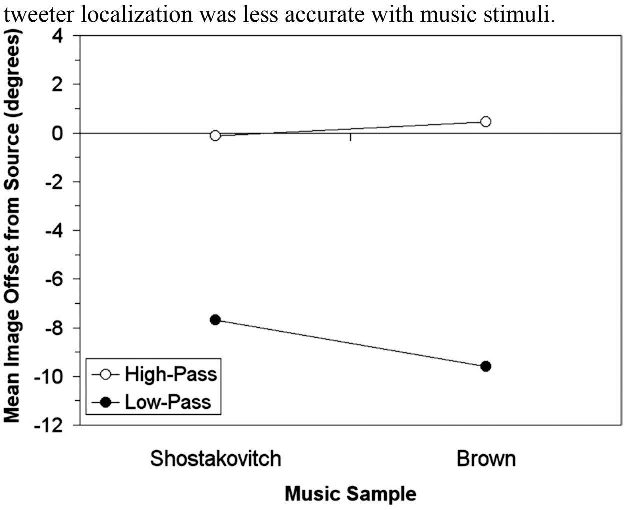

Figure 6. *Localization offset for music samples with a cross-over frequency of 1 kHz, from Ferguson and Cabrera [19].*

2.3. Pratt's Effect for Harmonic Tones and Tone Pairs

One concern with Cabrera and Tilley's and Ferguson and Cabrera's results was that they displayed little or no difference between results for median plane and lateral plane localization. In the lateral plane, binaural difference cues are available, and on the aural axis they are unambiguous. If these cues are effective, they should reduce or even eliminate Pratt's effect on the aural axis, and this is implied in Morimoto and Aokata's study of filtered white noise stimuli [9]. Cabrera and Morimoto [20] examined this apparent discrepancy between their previous studies through an experiment using harmonic tones from sources in either the median or lateral planes.

Cabrera and Morimoto [20] investigated Pratt's effect for single tones (monads) and simultaneous tone pairs (dyads). They used harmonic tones of two types: ones consisting of only five low order harmonics; and square wave tones (with odd harmonics extending to 16 kHz). The loudspeakers were at -15°, 0°, +15° and +30° from the horizontal plane at ear level. The subject either faced the loudspeaker array (and reported the image angle in the median plane), or had the array directly to their left (and reported the image angle in the lateral plane). Tone fundamental frequencies were 55 Hz, 110 Hz, 220 Hz, 311 Hz, 622 Hz and 1244 Hz. When presented as dyads, 55 Hz was paired with 1244 Hz, 110 Hz with 622 Hz, and 220 Hz with 311 Hz. Subjects localized the auditory images in darkness, with just a small spot of light illuminating the response sheets. Subjects with some musical training were selected, because these were more likely to be able to discriminate between one or two tones.

Having the subjects report the auditory image angle, rather than elevation in relation to a visual vertical scale, gave the subjects greater freedom, and so may have ameliorated an artifact of previous experiments. In the summarized results presented here, front-back reversal responses are mirrored to the front.

For the monads, Pratt's effect was observed for the 5-harmonic tone in both the median and lateral planes (with little apparent difference between the mean results of two planes), and for the square wave tone in the median plane only (based on analysis of variance, the positive slope in Fig 7 for the lateral plane square wave is not significant). The source position did not significantly affect image elevation for the 5-harmonic tones, but did for the square wave tones (both planes).

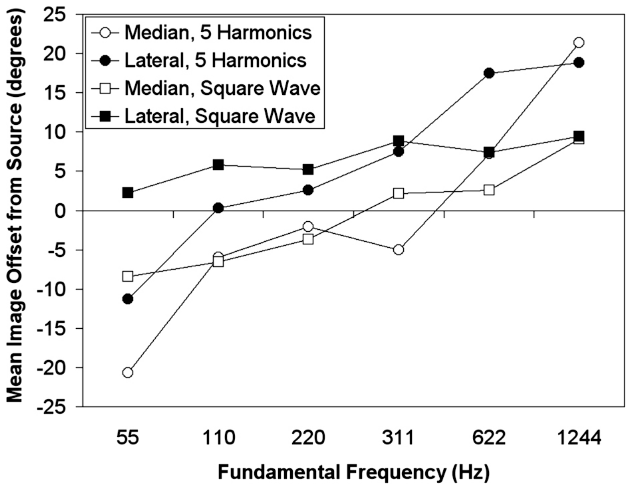

Figure 7. *Pratt's effect for single complex tones, in the median and lateral planes, as observed by Cabrera and Morimoto [20].*

Results of this study confirm that Pratt's effect is not restricted to the median plane, and for some stimuli, can be found in the lateral plane, including on the aural axis. However, closer examination indicated that Pratt's effect is smallest for stimuli on the aural axis, and grows with the cone of confusion. Results imply that the effect is expressed partly through rotation of the image around a cone of confusion, and partly through a simple vertical displacement.

For the 5-harmonic tone dyads, Pratt's effect was observed in the mean localization of high frequency tones above their low frequency counterpart. This occurred in all conditions, but to a smaller extent for lateral presentation and for square wave stimuli than for median plane presentation and 5-harmonic stimuli (Figure 8).

Pratt's effect is also observed, for the 5-harmonic tones, by a simple relationship between frequency interval and image spatial interval. This effect was not evident for the square wave. While this result for 5-harmonic dyads concurs with Ferguson and Cabrera's result for low- and high-passed noise bands, the result is simpler and stronger for these dyads.

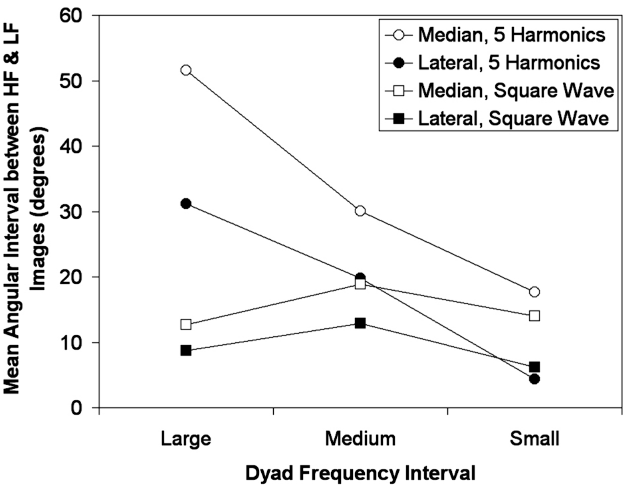

Figure 8. *Mean difference between the high frequency and low frequency image angles for tone dyads in the median and lateral planes, as observed by Cabrera and Morimoto [20].*

3. IMPLICATIONS AND APPLICATIONS

3.1. Generality of Pratt's effect

Pratt's effect allows the gross control of image elevation without using pinna-related spectral cues. In headphone applications, individualization is often required for spectral localization cues to be effective. By contrast, notwithstanding some issues discussed below, Pratt's effect provides a vertical localization cue with no need for individualization. While all studies of Pratt's effect are with loudspeakers, the loudspeaker arrangement used by Trimble [5] is similar to diotic headphone presentation, suggesting that Pratt's effect will be present in headphone presentation. Control of perceived elevation from a single fixed loudspeaker is also simple to implement, using Pratt's effect.

Although Pratt's effect is seen in a majority of individual subject results, a minority of subjects do not appear to be sensitive to Pratt's effect, although this observation is confounded with a host of factors that could affect subject performance. The recent studies referred to in this paper were not designed for individual subject analysis, but tentatively indicate that results of 10%–20% of subjects do not show any sign of Pratt's effect (considering both Australian and Japanese subject groups). Noble [12] finds that 25% of people exhibit poor sensitivity for vertical source discrimination (due to the form of their pinnae), and so are particularly prone to Pratt's effect.

The main experimental studies cited here were conducted in universities in the USA, Australia and Japan. Considering that each study investigated different stimulus sets, there do not appear to be major differences between the responses of these subject groups in terms of Pratt's effect. Nonetheless, in a cross-cultural study of visual associations with sound stimuli, Walker found less of an association between pitch and height in visual patterns in cultures with little exposure to Western culture [21] (Walker's study is of associations, not an auditory localization experiment, and so is not directly concerned with Pratt's effect).

There is potential for competing effects to undermine Pratt's effect. Apart from competition from binaural difference localization cues, the spectral effects variously known as 'boosted bands' [8], 'spatial referent maps' [22] or 'covert peak areas' [23] can defeat Pratt's effect for narrow bands of noise. For such stimuli, subjects tend to localize the auditory image to the direction for which the stimulus frequency is received with maximum efficiency, and this occurs in both monaural and binaural listening. Individual differences reflect the individuality of head-related transfer functions. The implication is that an auditory display that aims to take advantage of Pratt's effect should have its audio signals tested in a localization experiment prior to final implementation.

3.2. Influence of room acoustics

Early studies of Pratt's effect were not conducted in anechoic conditions, although presumably they were mostly conducted in rooms with substantial sound absorption. Interference from discrete reflections and standing waves in rooms can make pure tone localization no better than chance [24], and directional localization accuracy for many other types of sound signals is also degraded compared to free field conditions. Onset transients, spectral density, and spectral width are important in maintaining localization accuracy [25, 26]. When reliable localization cues are ambiguous or absent, localization 'illusions' can be strengthened [27]. There are no studies examining the effect of normal room acoustical conditions on Pratt's effect, but these observations suggest that the effect remains in such conditions, and may even be facilitated by the weakening of localization cues related to the source direction.

The influence of the floor reflection is of particular interest when considering Pratt's effect. It is more common for sound sources to be in the lower half of a room than in the upper half (near the ceiling), and in outdoor situations there are no ceiling reflections. In human activity, a floor (or ground) is almost omnipresent, and its distance from the ears depends on posture. A single reflection produces interference between the direct and reflected sound, which in spectral terms results in a comb filter. The delay between direct and reflected sound determines the density of the filter peaks and notches (which are regularly spaced on the linear frequency scale), and a short delay restricts the notches to the high frequency range. The strength of this comb filter effect from a floor reflection depends on the relative intensity of direct and reflected sound at the receiver position, which is determined by the path length difference and surface absorption and scattering. In realistic environments, this floor reflection effect is mixed with the effects of many other reflections.

As an example, we have considered the floor-related transfer function for a sound source with a 2.4 m horizontal separation from a receiving point. Like the more familiar head related transfer function, this is determined from the ratio of energy transfer with the intervention (the floor) versus the free field condition (in Fig 9 it is expressed as a simple ratio of squared sound pressures, not in decibels, to avoid excessive detail in the notches). The receiving point has an elevation of 1.2 m (about the ear height for a person seated in a chair) and the source elevation was varied. This situation was modeled for a point source in hemi-anechoic conditions (with a reflective floor only, having an absorption coefficient of 0.5), and also measured in a normally furnished room with either carpet or a hard floor. In all cases, the floor-related transfer function sees a progressive increase in squared sound pressure below 250 Hz, as the source approaches the floor from an elevation of 1.05 m

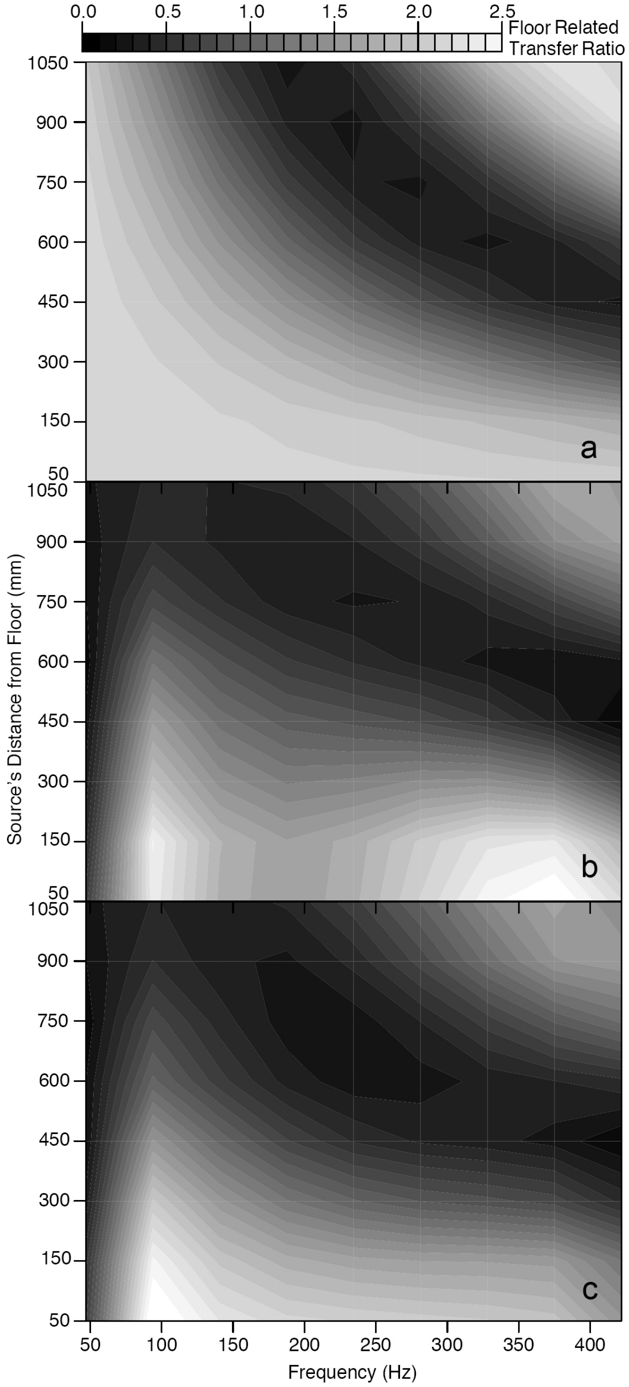

Figure 9. *Floor related transfer functions for a sound source horizontally separated from a receiver by 2.4 m, the receiver at 1.2 m elevation, and the source elevation varied from 50 mm to 1050 mm. Chart 'a' is the theoretical result for a point source in a hemi-anechoic environment (floor absorption coefficient of 0.5); 'b' is a measurement in a furnished room with a hard floor, and 'c' is a measurement in a furnished room with a carpeted floor.*

(higher positions were not measured). This is due to the first notch of the comb filter increasing in frequency – which means that the energy transfer ratio for frequencies below this notch increases. The greater loudness exponent at low frequencies [28] should make this increase in low frequency sound pressure more audible than the more complex transformations at higher frequencies. The measured results have a low frequency limit evident in Fig 9 (due to the characteristics of the small loudspeaker used). Note that, following reciprocity theory, the same effects would occur if the source and receiver were exchanged (notwithstanding the different directional characteristics of the transducers).

3.3. Music

Pratt believed that the effect he observed applies to music perception, stating:

> *We usually think of movement as the change in spatial location of an object which during the change is recognized as the same object or quality. But in musical movement it appeared as though the qualities changed while the spatial location was constant. From the present results, however, one would have to argue that musical movement resembles any other kind of movement to the extent that when different pitches are presented successively they change their spatial location with respect to one another.* [1]

While music theorists have tended to emphasize the metaphoric nature of the relationship between pitch and elevation (e.g. Scruton [3]), the studies reviewed in this paper imply that Pratt could have been right, at least in some circumstances. Musical tones that lack strong spectral localization cues could be subject to Pratt's effect, notwithstanding visual interactions. The relatively diffuse sound fields that are favored for un-amplified music produce imprecise localization, so that Pratt's effect could be unimpeded. Informal evidence for Pratt's effect in musical melodies is in Ferguson and Cabrera's experiment using music stimuli, in which some subjects commented that as a violin melody moved, so too did the image elevation.

Cabrera and Morimoto's study suggests that Pratt's effect could apply both to melodic movement and harmonic intervals, but again, limited to situations where source-related vertical localization cues are weak.

These are only implications – formal research on the time-varying localization of musical sound images remains to be done.

3.4. Spatial Imagery in Audio

Ferguson and Cabrera [19] examined the question of whether Pratt's effect could be applied to loudspeaker design. The most common driver arrangement for a multi-way loudspeaker has low frequency drivers (often referred to as 'woofers') below high frequency drivers ('tweeters'), and subwoofers are usually designed to sit on the floor. There are practical reasons for this: (i) having a low center of gravity gives the loudspeaker cabinet mechanical stability; (ii) a floor reflection (from a loudspeaker close to the floor) will enhance low frequencies, but create comb filtering at high frequencies due to their shorter wavelengths; (iii) typical room furnishings have obstructions in the low part of the room – which cause no problem for low frequencies (due to diffraction), but could block the path of high frequency sound; and (iv) the high frequencies dominate the spectral vertical localization cues, and so should be produced by a loudspeaker at ear level if vertical image displacement is not desired. Taking these factors into account, there could still be some benefit in positioning a woofer above a tweeter, if perceptual co-location of auditory images is taken as a design goal. For a 1 kHz crossover frequency, Ferguson and Cabrera's results indicate that the sound image produced by a woofer can be moved up at least to ear level, by elevating the woofer – although it may not be possible to raise the image any higher than that. Nevertheless, there is little or no evidence that perceptual co-location is desirable in music reproduction, and tentative results from a preference pilot study by Ferguson and Cabrera did not show any preference for a particular vertical configuration of woofer and tweeter. Since an expansion of auditory image size is often considered to be desirable in music, the conventional arrangement (with woofer below tweeter) – which exaggerates Pratt's effect – may indeed be the optimum.

Another observation from these studies is that in the absence of visual cues, for predominantly low frequency sound images to be elevated, there needs to be associated high frequency sound capable of conveying pinna-related elevation cues. Without this, the experiment results indicate that low frequency images will almost certainly be below the horizontal plane, even if the loudspeakers are elevated. This observation has application in audio production. It is not known whether visual cues of elevated sources (for example, in motion pictures or computer games) can defeat Pratt's effect, but it seems likely that the inclusion of pinna-related spectral cues of elevation (for elevated predominantly low frequency images) should improve audio-visual integration.

3.5. Auditory Display

3.5.1. Representing Vertical Space

As previously discussed by Cabrera and Tilley, Pratt's effect, and related associations, could provide an intuitive coding for auditory display applications in which a spatial scale from low to high is being represented. An example of this might be auditory alert signals corresponding to parts of a visual display: an alert for a high part of the display could have higher frequency content than an alert for a low part of the display.

3.5.2. Representing Analogous High-low Scales

Pitch and frequency are not alone in using an elevation analogy – other concepts such as temperature and price – also use the same spatial analogy. In these two examples, the terms 'high' or 'low' are adopted because of the use of one-dimensional numerical scales used to measure quantities. Visual displays (such as an analog thermometer, or stock price chart) tend to use a vertical value scale, in accord with the high-low metaphor.

3.5.3. Representing Emotion

The terms 'high' and 'low' have some application in emotion. Using the two-dimensional model of emotion of Russell [29] and Schubert [30] illustrated in Figure 10, 'high' refers to positive arousal and positive valence, while 'low' refers to negative arousal and negative valence. This spatial metaphor may be related to posture, but also to emotion as it is represented in speech and music. There are associations between high pitch and high valence [31, 32], and between high pitch and high arousal [33, 34]. However, harsh or sharp timbres, which are associated with strong high frequency harmonic content [28], are associated with negative valence [34, 35].

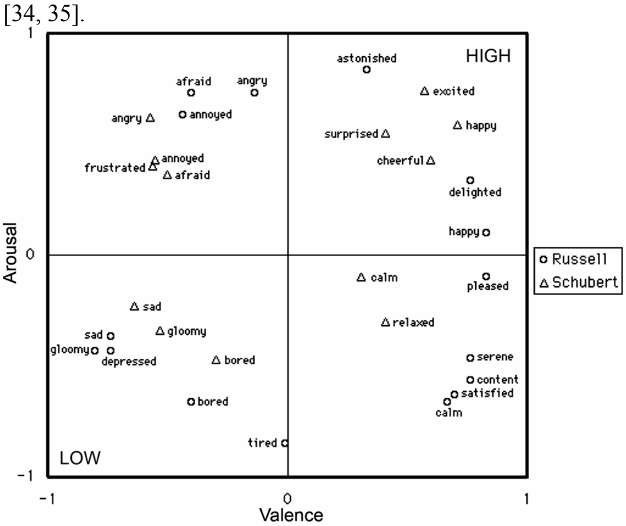

Figure 10. *Ratings of emotion words in a two-dimensional emotion space, according to Russell [29] and Schubert [30]. The words 'high' and 'low' are in the quadrants containing emotion words associated with them.*

There are other sound parameters that are effective in consistently representing emotion, including loudness and temporal attributes. In both music and speech, loudness is closely related to arousal [30, 35, 36]. Similarly, the rapidity of auditory events (e.g. expressed through syllabic rate or musical tempo) is closely related to arousal [30, 37, 38]. However, studies tend to show little relationship between valence and loudness nor valence and the rapidity of events.

In summary, there can be a concordance between Pratt's effect and the scale of emotions ranging between 'high' and 'low', as expressed through the frequency of auditory stimuli

4. CONCLUSIONS

The studies summarized in this paper indicate that Pratt's effect is not restricted to median plane localization, and exists even in the presence of unambiguous binaural difference cues. This result highlights the importance of spectral cues in accurate auditory localization, even in situations where binaural difference cues are available. Pratt's effect is controlled by spectral envelope – although this does not preclude the possibility that a pitch change, without a gross spectral envelope change, could produce a similar effect. Although the effect may be subject to cultural influences, studies in various countries show the effect. There are multiple associations that are in harmony with Pratt's effect, including the spectral changes that occur with elevation in relation to a reflective floor, and the authors have suggested that the widespread availability of this floor-related elevation cue could be an influence on Pratt's effect. While some previous studies treat Pratt's effect as a deficit in auditory localization, the authors suggest that the effect could be useful in various applications, including auditory display. Further studies are warranted into the application of Pratt's effect in music, and the presence of Pratt's effect in a range of room acoustical conditions.

5. REFERENCES

[1] C. C. Pratt, "The spatial character of high and low tones," *J. Exp. Psych.*, vol. 13, pp. 278-285, 1930.

[2] S. K. Roffler and R. A. Butler, "Localization of tonal stimuli in the vertical plane," *J. Acoust. Soc. Am.*, vol. 43, no. 6, pp. 1260-1266, 1968.

[3] R. Scruton, "Understanding music," *Ratio*, vol. 25, no. 2, pp.97-120, 1983.

[4] F. L. Dimmick and E. Gaylord, "The dependence of auditory localization upon pitch," *J. Exp. Psych.*, vol.17, pp. 593-599, 1934.

[5] O. C. Trimble, "Localization of sound in the anterior-posterior and vertical dimensions of 'auditory' space," *Brit. J. Psych.*, vol. 24, pp. 320-334, 1934.

[6] S. K. Roffler and R. A. Butler, "Factors that influence the localization of sound in the vertical plane," *J. Acoust. Soc. Am.*, vol. 43, pp. 1255-1259, 1968.

[7] V. R. Algazi, C. Avendano, and R. O. Duda, "Elevation localization and head-related transfer function analysis at low-frequencies," *J. Acoust. Soc. Am.*, vol. 109, no. 3, pp. 1110-1122, 2001.

[8] J. Blauert, *Spatial Hearing : The Psychophysics of Human Sound Localization*, rev. ed., MIT Press, Cambridge MA, 1997.

[9] M. Morimoto and H. Aokata, "Localization cues of sound sources in the upper hemisphere," *J. Acoust. Soc. Jpn. (E)*, vol. 5, no. 3, pp 165-173, 1984.

[10] J. Blauert, "Sound localization in the median plane," *Acustica*, vol.22, pp. 205-213, 1969/70.

[11] M. B. Gardner, "Some monaural and binaural facets of median plane localization," *J. Acoust. Soc. Am.*, vol. 54, pp. 1489-1495, 1973.

[12] W. Noble, "Auditory localization in the vertical plane: Accuracy and constraint on bodily movement," *J. Acoust. Soc. Am.*, vol. 82, no. 5, pp. 1631-1636, 1987.

[13] S. J. Bothe and L. F. Elfner, "Monaural vs binaural auditory localization for noise bursts in the median vertical plane," *J. Aud. Res.*, vol. 12, pp. 291-296, 1972.

[14] D. Cabrera and S. Tilley, "Parameters for the auditory display of height and size," Proceedings of the 2003 International Conference on Auditory Display, Boston MA, pp29-32, 2003.

[15] R. J. Davis and S. D. G. Stephens, "The effect of intensity on the localization of different acoustical stimuli in the vertical plane," *J. Sound Vib.*, vol. 35, 223-229.

[16] V. Best, A. Van Schaik, and S. Carlile, "Spatial effects on the segregation of sounds in virtual auditory space," in Proc. 8th Western Pacific Acoustics Conf., paper TE33 (Melbourne, Australia, 2003).

[17] M. Morimoto, M. Yairi, K. Iida, and M. Itoh, "The role of low frequency components in median plane localization," *Acoustical Science and Technology*, vol. 24, pp. 76-82, 2003.

[18] D. Cabrera and S. Tilley, "Vertical localization and image size effects in loudspeaker reproduction," in Proc. AES 24th Int. Conf. on Multichannel Audio, paper no. 46 (Banff, Canada, 2003).

[19] S. Ferguson and D. Cabrera, "Vertical localization of sound from multi-way loudspeakers," *J. Audio Eng. Soc.*, vol. 53, pp.163-173, 2005.

[20] D. Cabrera and M. Morimoto, "Auditory vertical localization of complex tones and tone pairs," in preparation, 2005.

[21] R. Walker, "The effects of culture, environment, age, and musical training on choices of visual metaphors for sound," *Perception and Psychophysics*, vol. 42, no. 5, pp. 491-502, 1987.

[22] A. D. Musicant and R. A. Butler, "The psychophysical basis of monaural localization," *Hearing Research*, vol. 14, pp. 185-190, 1984.

[23] M. E. Rogers and R. A. Butler, "The linkage between stimulus frequency and covert peak areas as it relates to monaural localization," *Perception and Psychophysics*, vol. 52, no. 5, pp. 536-546, 1992.

[24] W. M. Hartmann, "Localization of sound in rooms," *J. Acoust. Soc. Am.*, vol. 74, no. 5, pp. 1380-1391, 1983.

[25] B. Rakerd and W. M. Hartmann, "Localization of sound in rooms, II: the effects of a single reflecting surface," *J. Acoust. Soc. Am.*, vol. 78, no. 2, pp. 524-533, 1985.

[26] B. Rakerd and W. M. Hartmann, "Localization of sound in rooms, III: Onset and duration effects," *J. Acoust. Soc. Am.*, vol. 80, no. 6, pp. 1695-1706, 1986.

[27] W. M. Hartmann and B. Rakerd, "Localization of sound in rooms, IV: The Franssen effect," *J. Acoust. Soc. Am.*, vol. 83, no. 4, pp. 1366-1373, 1989.

[28] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*. Springer-Verlag, Heidelberg, Germany, 1990.

[29] J. A. Russell, "Measures of Emotion," in *Emotion: Theory Research and Experience*. R. Plutchik and H. Kellerman, eds., Academic Press, New York, vol. 4: pp. 81-111, 1989.

[30] E. Schubert, *Measurement and Time Series Analysis of Emotion in Music*. Thesis (Ph.D.), University of New South Wales, 1999.

[31] J. F. Thayer, *Multiple Indicators of Affective Response to Music*, Thesis (Ph.D.), New York University, 1986.

[32] K. Hevner, "The Affective Value of Pitch and Tempo in Music," *Am. J. Psych.*, vol. 49, pp. 621-630, 1937.

[33] S. C. Collins, *Subjective and Autonomic Responses to Western Classical Music*. Thesis (Ph.D.), University of Manchester, 1989.

[34] K. R. Scherer. and J. S. Oshinsky, "Cue Utilization in Emotion Attribution from Auditory Stimuli," *Motivation and Emotion*, vol. 1: 331-346, 1977.

[35] A. Gabrielsson and P. N. Juslin, "Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience." *Psych. of Music*, vol. 24, pp. 68-91, 1996.

[36] S. Namba, S. Kuwano, T. Hatoh and M. Kato, "Assessment of Musical Performance using the Method of Continuous Judgement by Selected Description," *Music Perception*, vol. 8, pp. 251-276, 1991.

[37] K. R. Scherer, "Speech and emotional states," in *Speech Evaluation in Psychiatry*, J. K. Darby, ed., Grune and Stratton, New York, 1981.

[38] G. L. Collier, "Affective synaesthesia: extracting emotion space from simple perceptual stimuli," *Motivation and Emotion*, vol. 20, pp. 1-32, 1996.