Latent, Imperfect and Practically Perfect Pitch
# Table of Contents

1. Overview p 4

2. What is Perfect Pitch? p 5-9
   2.1 Timbre and Pitch Register p 10
   2.2 Other physiological phenomena p 11

3. Why is absolute pitch of interest? p 12
   3.1 Neurological pathways p 12-14
      3.1.1 The planum temporale p 14-18

4. Genetics of Absolute Pitch p 19

5. Cognitive representation of absolute pitch p 21
   5.1 Pitch Perception p 21-23
   5.2 Pitch memory p 23-29
   5.3 Good pitch memory is widespread p 29-30
   5.4 Pitch associations p 30-33

6. Intermodal effects p 34-35

7. Why is AP unresolved? p 36-37

8. Summary & Conclusions p 38-42
Abstract

Absolute pitch is the ability to name or produce a note without an external reference, it is believed to be a rare ability with a commonly cited figure of 1 in 10,000 in the general population. Whilst good pitch memory is widespread, it is only a few who are able to put labels to pitches. A critical or sensitive period of musical education before the age of 5 or 6 has been suggested, and the type of musical education may also play a part. The roles of the “fixed do” system and the possible genetic contribution to the acquisition of AP are discussed. Anatomical differences of an enlarged left Planum temporale in AP possessors have been noted on MRI, the aetiology of which is still unclear.
1. Overview

Absolute Pitch (AP), or the ability to name a note without an external reference seems to be of fascination to musicians, psychologists, neurologists and especially the wider public. Perhaps it is because famous musicians and composers such as Bach, Beethoven, Rubenstein, Toscanini and Heifetz were noted to have AP, that it has been glamourised as a mysterious and desirable musical endowment (Deutsch et al, 2004). The commonly cited figure is that only 1 in 10,000 people claim to have perfect pitch (Profita & Bidder, 1988). Persons with AP tend to identify tones of their main instrument more reliably than other timbres (Lockhead & Byrd, 1981), however, the degree of accuracy of AP varies, with most AP possessors being able to name between 70 to 100% of middle range piano tones (Miyazaki, 1988). Accuracy has also been known to vary with the menstrual cycle (Wynn 1970)!

“How do they do that? “ seems to be the most common question asked. From perspective of psychology and neurology, understanding AP may shed some light on the organization of the brain and perceptual mechanisms (Zatorre 1998) as it involves a complex interplay of several neural processes, namely, pitch perception, classification and naming, long term pitch memory and retrieval (Levitin 1999).

Why would anyone want to have AP? Musicians with AP are not necessarily more sensitive or imaginative interpreters, they are not necessarily better composers nor are they technically more adept at their instruments (Levitin 1999). Nevertheless, the advantage of AP is a greater change of “picking up” a tune in the right key when it is heard, or ”picking up” when playing in a group or accompanying when one gets lost, and atonal music is easier to “pitch”.
The disadvantages include poor ability to quickly transpose a piece of music when playing, an unease when listening to a piece transposed into another key and great discomfort or inability to play “scordatura” in stringed instruments where the strings are tuned to different pitches.

This aim of this paper is to examine studies with an emphasis on anatomical, functional and genetic components to understand the nature of the relationship between genetic predisposition and musical training in the incidence of AP.
2. What is perfect pitch?

There are many definitions of Absolute Pitch, e.g. “the ability to identify the frequency or musical name of a specific tone” (page 265, Ward, 1999), “the ability to name or produce a note of a particular pitch in the absence of a reference note” (page 200, Deutsch 2001) or “the ability to identify the pitch of a musical tone or to produce a musical tone at a given pitch without the use of an external reference pitch” (page 345, Takeuchi and Hulse, 1999). However, one of the most widely quoted descriptions is that of Levitin (page 255, 1999) who states that Absolute (AP) or “Perfect’ Pitch is the ability to identify the chroma (pitch class, e.g., C, C#, D) of any isolated tone and/or to reproduce any given tone without an external reference.

Deutsch (1999) may argue that musical pitch should also take into account pitch height, which places a note within a particular octave. (This said, it has been argued that AP possessors differ from no AP possessors in their ability to identify pitch class but simply estimate pitch height in the same way that non AP possessors do (Lockhead and Byrd 1981). On this basis, and because of the difficulty in denoting octaves it is felt that the definition of AP should exclude octave designation.

There are 2 dimensions to pitch: the octave and in Western Music, the cycle of 12 semitones that make up the octave; in perception, this is referred to as pitch chroma and pitch height respectively. The notes are plotted on the helix with one complete octave per circuit (see Figure 1).
Perfect Pitch a misnomer, musicians with AP are not necessarily better at discriminating tones than non AP musicians, nor are they able to recognise or reproduce “perfect intonation”, be it mean or equal temperament. AP musicians are able to recognise/reproduce a range of pitches that are very close to the target frequency, but not necessarily exactly centered on the target frequency eg A440 (Levitin 1999). Correspondingly, AP musicians are not more accurate at identifying pitch height than non AP musicians.
Despite the general consensus of researchers of the approximate parameters that define Absolute Pitch, the criteria by which research teams define a subject as having Absolute Pitch varies widely. Most researchers test pitch recognition rather than pitch production and would exclude octave errors and errors of an average of 1 semitone either flat or sharp are accepted as being consistent with Absolute Pitch (Takeuchi and Hulse, 1993). Research studies are not identical in their methodology, with some utilising sine waves and others using piano or synthesized piano tones as stimuli, furthermore, there is a wide variation in the range of frequencies tested (see Table 1). In addition, the degree of accuracy of AP varies, however, most can name between 70 to 100% of middle range piano tones (Miyazaki, 1988). For example, this is reflected in several studies that accept a mean accuracy of between 74% and 99% (Takeuchi and Hulse, 1993), a rather large range of means, although it is still significantly better than chance.
Table 1 (from Takeuchi and Hulse, 1993)

<table>
<thead>
<tr>
<th>Study</th>
<th>No. AP tested</th>
<th>Range</th>
<th>%correct Mean</th>
<th>Timbre</th>
<th>Pitch Register</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balzano (1984)</td>
<td>3</td>
<td>84.3 –91.2</td>
<td>87.7</td>
<td>Sine</td>
<td>A2-G#5</td>
</tr>
<tr>
<td>Benguerel &amp; Westdal (1991)</td>
<td>10</td>
<td>62-100</td>
<td>88</td>
<td>Sine</td>
<td>C3-C5</td>
</tr>
<tr>
<td>Carroll (1975)</td>
<td>5</td>
<td>79.7-95.3</td>
<td>86.6</td>
<td>Piano</td>
<td>A3-C5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48.0-92.8</td>
<td>71.0</td>
<td>Piano</td>
<td>A1-C7</td>
</tr>
<tr>
<td>Lockhead &amp; Byrne</td>
<td>4</td>
<td>99</td>
<td>58</td>
<td>Piano, Sine</td>
<td>A1-C7, C2-A8</td>
</tr>
<tr>
<td>Miyazaki (1998a)</td>
<td>7</td>
<td>90.2-99.4</td>
<td>94.1</td>
<td>Sawtooth</td>
<td>C3-C6</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>45-99.7</td>
<td>82.7</td>
<td>Synth. Piano</td>
<td>C3-C6</td>
</tr>
<tr>
<td>Miyazaki (1989)</td>
<td>7</td>
<td>91.6</td>
<td>74.4</td>
<td>Piano</td>
<td>C1-B7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sine</td>
<td>C2-B4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Synth Piano</td>
<td>C1-B7</td>
</tr>
<tr>
<td>Miyazaki (1990)</td>
<td>10</td>
<td>53.3-99</td>
<td></td>
<td>Synth Piano</td>
<td>C2-B6</td>
</tr>
<tr>
<td>Takeuchi &amp; Hulse (1991)</td>
<td>19</td>
<td>58.4 – 100</td>
<td>83.1</td>
<td>Synth Piano</td>
<td>E3-D#5</td>
</tr>
<tr>
<td>Zatorre &amp; Beckett (1989)</td>
<td>21</td>
<td>27.1-100</td>
<td>74.8</td>
<td>Piano</td>
<td>C2-C5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>37.3-100</td>
<td>89.5</td>
<td>Piano</td>
<td>C4-C5</td>
</tr>
</tbody>
</table>
2.1 Timbre and pitch register

Timbre may also play a role in the accuracy of AP; persons with AP tend to identify tones of their main instrument more reliably than other timbres (Lockhead & Byrd, 1981); this may be due in part to a familiarity with the timbre of their primary instrument. Most instruments have pitch range limitations, for example the violin tends to concentrate on the higher registers and the double bass, in contrast, wallows in the depths of the lower frequencies. This has confounding effect on studies as a wider range of pitches can be tested on some instruments compared with others, hence, difficulties in labelling pitches played on certain instruments (eg double bass or piccolo) may be due in part to the pitch register rather than the timbre of the instrument. Notwithstanding this, it is of interest to note that the pitches played on a piano are apparently easier to identify than other musical instruments despite its large range (Takeuchi and Hulse, 1993), suggesting factors other than pitch range or register may be accountable for the higher accuracy. Most of the older studies on AP have been performed on pianos partly for their accessibility in the community and also for their large range of register that encompasses most of the instruments of the orchestra (Levitin 1999, Ward, 1999).

Most music is written in the middle registers and it has been observed that AP possessors are generally more accurate and faster in identification of pitch class that fall within the middle registers (Takeuchi and Hulse, 1993). This may be a result of familiarity and exposure; possible exceptions would include instrumentalists whose instruments whose range fall in the extremes, for example violinists, piccolo players, double bass and tuba players.
Miyazaki (1988) noted that pitch identification accuracy and speed were superior for piano white notes compared to black notes, a finding that concurs with those of Takeuchi and Hulse (1991, cited in Takeuchi & Hulse, 1993). Possible explanations include the early learning theory that piano beginners almost always start using white notes in the key of C, that white key pitches are used more frequently in music than that of black keys, at the same time, the problem may be one of labelling, with more decisions required to label a black note (a two step process), eg, C# or B flat as opposed to simply “C”. It has been suggested (Takeuchi and Hulse 1993) that one may get around the problem of labelling by using the term “C Natural” or even using enharmonic labelling such as “B sharp”. The latter suggestion is extremely complicated, (certainly not part of a beginner music student’s vocabulary) and makes for much more complex labelling; perhaps this is an attempt to compensate for the presumed early learning theory of white note dominance in beginner piano students.

2.2 Other physiological phenomena

Accuracy has also been known to vary with the menstrual cycle (Wynn 1970) and age. Whilst the aetiology of the variation in accuracy during the menstrual cycle is not clear, it has been postulated that oestrogen may have an effect on the sodium-potassium ratio causing changes in the electrical activity of the brain or possibly due to changes of the cellular structure of the ear as a result of hormones. Similarly it has been noted that after the age of 50 or so, music may be heard up to two semitones sharper that what it ought to be (Ward, 1999), this may be due in part to age related changes on the basilar membrane; metabolic or more centrally mediated changes such as epileptic drugs (eg carbamazepine) may also affect accuracy of AP.
3. Why is absolute pitch of interest?

Absolute pitch is of interest for many reasons. To a musician it is a handy skill to have, making it easier to “pick up” the music from any given point when playing in a group or accompanying and making atonal music easier to “pitch” and to play.

From a neurologist’s or psychologist’s point of view, AP is of interest because it may yield some insights into the organization of the brain (Zatorre 1998) as it involves several separate neural processes, namely, pitch perception, classification and naming, long term memory and retrieval (Levitin 1999). These multiple neurological processes make the study and evaluation of AP extremely complex and complicated due to a multitude of variables.

3.1 Neurological pathways

There has long been some curiosity about the possibility of anatomical differences in the brains of persons with Absolute Pitch. Schlaug et al (1995) have documented this in a widely quoted study. The study examined brain structures using Magnetic Resonance Imagery and noted that musicians with Absolute Pitch showed stronger leftward planum temporale (an area between the temporal and parietal lobes, on the surface of the superior temporal gyrus posterior to Heschl’s gyrus containing the auditory association cortex) as well as a larger left planum temporale. In other words, not only was the left side of the auditory association cortex in musicians with perfect pitch larger than the musicians without perfect pitch but the left planum temporale was noted to be larger than the right planum temporale in musicians with AP. This area on the left side is known as Wernicke’s area and is used for interpretation of
speech and language in right-handed persons. One explanation is that whilst there may be brain circuitry that is unique to either music or speech, there may be some aspects of speech or music that involves circuitry common to both. (Deutsch et al, 2004). It may be possible that in persons with AP, this proposed common speech/music circuitry is extremely well developed and coupled with strong verbal associations to pitch class, results in an anatomically enlarged left planum temporale.

A methodological problem of measuring Planum Temporale volumes in MRI is the difficulty in standardizing the boundaries of the Planum Temporale for MRI study. The mesial border is believed to be the point where the Planum Temporale meets the insula and the anterior border marked by the sulcus behind Heschls gyrus or the extension of the first gyrus to the lateral edge (Westbury et al 1999, Zatorre et al 1998.). However, the posterior border of the PT is less easily determined with varying area estimates between research groups (Westbury et al, 1999). A probability map using a voxel painting program (whereupon pixels/voxels can be coloured) to identify and label the PT in horizontal, saggital and coronal planes, and transformed to a stereotactic space has been proposed by Westbury et al in an attempt to standardise the measurements of PT between research groups. However, the differences remain, with some groups choosing to measure the volume of the PT and others electing to measure the surface area of the PT (Zatorre et al 1998, Schlaug et al 1995).

One of the limitations of Magnetic Resonance Imaging (MRI) is that it purely assesses anatomical structures and does not give an indication of function. In order to assess function of the brain, Positive Emission Tomography (PET) is used. Most frequently, it is used to study the degree of blood flow in various parts of the brain. To
enhance its usefulness, PET scans are superimposed on MRI scans to correlate brain blood flow with anatomical structures, and this is commonly known as Functional MRI, this is a useful tool in measuring/observing the in vivo brain activity of perceptual tasks. In addition to merely measuring blood flow, PET has the ability to measure/trace the metabolism of neurotransmitters by using different radioligands.

However, there are limitations to fMRI, and caution should be exercised when examining activation patterns seen in a single plane, as the signal may not be observed in other planes e.g. coronal versus sagittal (Zatorre, 1997). In addition, there may be individual variation in the degree of activation in response to a given input. Nevertheless, fMRI remains an extremely useful tool, however, we need clever behavioural experiments used in conjunction with fMRI to determine underlying processes and/or mechanisms.

3.1.1. The planum temporale

The role of the Planum Temporale in the processing of music has been demonstrated in a study involving music students by Ohnishi et al (2001). Ohnishi et al utilised functional MRI to assess cerebral activity in both musicians and non-musicians. During a passive listening task, the musicians showed increased cortical activity in the left secondary auditory areas, the temporal cortex and left dorsolateral prefrontal cortex, whereas the non-musicians demonstrated increased activation in the right secondary auditory areas. In particular, there was a significant difference in the degree of activation of the bilateral planum temporale, the left secondary auditory area and co-activation of the left posterior dorsolateral prefrontal cortex.
This is of interest as the posterior temporal area, including the planum temporale is involved in pitch processing (Binder et al, 1996, Mazziotta et al, 1982) The earlier study by Zatorre et al (1998) suggested that activation of the left posterior dorsolateral prefrontal cortex might represent a marker for absolute pitch processing and behavioural studies have demonstrated more left lateralized representation of musical processing in musicians as compared to non musicians( Bever &Chiarello, 1974, Mazzucchi et al, 1998, Schlaug et al, 1995). In other words, in gross terms, whilst the right cerebral hemisphere may be important for music processing in non musicians, increasing musical sophistication or training may cause a shift of musical processing from the right hemisphere to the left or flexibility in the dominant hemisphere.

Other studies looking at cerebral blood flow (hence, brain activation) have also supported the differences in hemispheric brain activation of when processing/listening to music amongst musicians and non-musicians. Non-musicians were noted to use primarily the Right Hemisphere and musicians showed a preference for the Left Hemisphere when listening to music (Evers et al, 1999). This study utilised transcranial Doppler ultrasound to measure cerebral blood flow instead of PET studies.

Likewise, the degree of activation of the left planum temporale was negatively correlated with the age at which the musicians began musical training (Ohnishi et al 2001). The younger the age of commencement of musical training, the larger the activation of the left planum temporale. This suggests that the degree of functional re-organisation, possibly due to re-routing or even perhaps greater than expected growth of pathways of the brain connections may depend on the age at which musical
training began. There did not seem to be any correlation between duration of training and cerebral activation, suggesting that this possible re-organisation of the left planum temporale or reinforcement of pathways involving the left planum temporale may be related to early commencement of training rather than long-term training.

The findings and conclusions of Ohnishi et al’s study align or conform with a study by Sergeant and Roche (1973) of over 1000 professional musicians, which suggested that the possession of Absolute Pitch related to the age at which musical training began. Ninety five percent of musicians who had begun training before the age of four years possessed perfect pitch, in contrast to only 5% of those who had commenced training between 12 and 14 years of age, implying a sensitive/critical period for development of AP in early childhood.

The physical/ anatomical size of the Planum Temporale appears to be positively correlated with the degree of accuracy of AP (Zatorre et al 1998).

It remains unclear whether musical abilities and structural differences seen in musician’s brains are due solely to learning and training, possibly during the critical periods of brain development and maturation or whether they reflect innate/congenital abilities and capacities that may be expressed by early musical training and exposure (Schlaug et al 2001). The answer is not clear as there is only a correlation (not causation) between early musical training and the acquisition of AP. The consistent anatomical finding in persons with AP is a leftward asymmetry of the Planum Temporale and possibly a larger left Planum Temporale (Schlaug et al,1995).
The question raised by the study of Schlaug et al in 1995 was were these musicians with Absolute Pitch born with larger left planum temporale or did it develop with training?

Keenan et al (2001) attempted to answer this in a study of the size and leftward asymmetry of the Planum Temporale of 27 AP musicians, 27 non musicians and 22 non AP musicians. All the musicians, (bar one in the AP group), had begun musical training at or before the age of 7 and hence, incorporated the possible critical period for acquisition of AP.

In support of the earlier study by Schlaug et al (1995), marked leftward asymmetry was noted in the AP musician group, but there was no significant leftward asymmetry of PT in the non AP musician and non musician groups. Significantly, the marked Leftward asymmetry of the AP musician group was due to the significantly smaller Right PT in this group. The authors postulate that this degree of asymmetry may be due to pruning of the Right PT leading to a functional dominance of the Left PT over the Right, which may be a pre requisite for the manifestation or acquiring of AP.

Furthermore they make the bold statement that whilst there may not be a genetic correlate for AP, there may be a genetic correlate for Leftward PT asymmetry and this coupled with early musical training may determine manifestation of AP. It would be interesting to note if this asymmetry is a neonatal phenomenon, or an acquired change independent of early musical training and perhaps due to some other variable. Cerebral plasticity and repeated musical practice at a critical age may be responsible for the noted anatomical changes, just as violinists have greater cortical representation
of the fingers of the left hand (Elbert et al, 1995) or blind persons with an early age of onset of blindness appearing to have superior pitch discrimination than sighted persons, the younger the onset of blindness, the better the performance. This is accordance with cerebral plasticity being optimal during the early years (Gougoux et al, 2004).
4. Genetics of absolute pitch

Scientists from the University of California, San Francisco have been studying the incidence of Absolute Pitch (AP) amongst musicians and non musicians. Their equipment consists of a laptop computer that fires off 30 tones successively at a subject. The subject has to name the notes almost instantaneously, any more than 2 incorrect answers fails the test (a very strict definition compared to other studies which allow for up to a 30% error rate). Their estimated incidence of AP in the general population varies from 1 in 2000 (in contrast to 1 in 10000 Profita & Bidder, 1988). To date, 600 musicians from the San Francisco Conservatory of Music, the Curtis Institute and the Peabody Conservatory of Music have been included in the study. Within the subgroup of musicians, the incidence of AP was 15%.

Once again, early training seemed to be a critical factor with almost all subjects with AP having commenced their training before age 12. The scientists felt that genetics may have a part to play, with 50% of subjects with AP having a close relative with AP and almost none of the subjects without AP being able to boast of a relative with AP. Blood is also being collected from the subjects with AP in the hope of finding a distinctive pattern of DNA sequence that may be linked to AP. It would then be possible to tract its distribution amongst families.

On the basis of the preliminary evidence from these ongoing studies, Baharloo et al (2000), Baharloo et al (1998) and Gergesen (1998) have postulated that AP has both a genetic and non genetic component. Both are necessary, neither being able to stand-
alone. They speculate that the presumed genetic endowment would require a nurturing or appropriately aurally stimulating environment in order to express itself.
5. Cognitive representation of absolute pitch

Tone/Pitch class AP involves 2 broad sub-skills, long term pitch memory and some form of labelling (eg. C#, D flat, A natural). AP musicians are not necessarily better at discriminating small changes in frequency than non AP musicians (Levitin 1999), for example AP possessors are not able to consistently discriminate between pitches of less than a semitone apart.

5.1 Pitch Perception

One of the theories of AP (Zatorre et al 1998) is that the memory for pitches over a long period of time is due to verbal coding, implying that AP possessors do not actually remember the sound or pitch of a tone, rather they hear the tone, then recall the pitch name. In pitch reproduction tasks, the pitch name is remembered and the subject produces a tone (using auditory feedback to reinforce the association) corresponding to the pitch.

Levitin (1999) suggests that Absolute Pitch Possessors may utilise Categorical Perception with a range of frequencies fitting into a particular category/pitch class. In order to further understand Categorical Perception in the context of AP, Levitin (1996, cited in Levitin,1999) presented AP subjects with focal tones and also modified variations on those focal tones in 20 cent increments. If subjects chose the focal tones as good examples of the category, then this was considered evidence for a “narrow anchor” view of AP if the other choices apart from the focal tones were considered to be “equally as good” then it was felt that this was a “wide anchor”. The
results seemed to support the notion of a “narrow category” with peaks of “correctness” at the focal tones and troughs at the boundaries of the categories.

The definition of AP would suggest an accurate production of a designated pitch class or correct categorisation and labelling of a presented pitch. This seems to be of interest as AP appears to violate the 7+/− 2 short term memory capacity rule which states that subjects are able to categorise stimuli into 5 or 9 categories consistently (Miller, 1956, cited by Levitin 1999). AP possessors who are able to classify over 60 stimuli (5 octaves of 12 tones) seem to break this rule. However, as many AP possessors make octave errors, one may narrow this down to approximately 12 chunks (Levitin, 1999), perhaps a smaller violation!

Whilst it is widely believed in the general community that a pitch belongs to one category or another, there are variations to “intonation”, as players of mean tempered instruments would attest. This is one of the tuning challenges faced by the traditional Piano Trio (the violin and cello capable of “mean tempered” intonation and the modern piano being fixed at equal temperament), mercifully the String Quartet is spared this challenge. The String Quartet faces other challenges relating to how “sharp” or “flat” the thirds and sixths should be tuned in order to emphasize the key (tonality) of the piece and its “colour”, an issue often referred to as “exquisite intonation”.

In keeping with the above, Petran (1939, cited in Takeuchi and Hulse, 1993), claims that AP possessors will accept a range of frequencies in the “internal standard or template” as being of a particular pitch (class). Levitin (1999) postulates that there is a
“wide notion” of pitch, that people will accept a range of frequencies (possibly to a limit of up to 1 semitone) as being of a particular pitch. Amongst musicians, there is a rather fluid definition of pitch/intonation that seems to correlate with the overall tonality of a piece (relative pitch task), however, the pitch category has a relatively narrow anchor. There seems to be a midpoint of the category possibly corresponding to the fundamental frequency, with small variations (smaller than a semi tone) on either side of the midpoint. A variation of a semitone is too large and does not constitute “fine discrimination”. To a musician a semitone is a large interval in a musical context and semitone errors could radically alter the overall tonality of a piece, causing dissonance.

5.2 Pitch memory

Long term pitch memory is the other broad sub skill of AP. Some researchers have attempted to identify “innate or latent AP” in non musicians by eliminating the “pitch labelling” component and relying on presumed long term pitch memory and recognition.

Ross et al (2003) studied 27 experienced musicians (6 with AP and 21 without) as well as 1 person with minimal musical experience who nevertheless claimed that he had AP despite being unable to name notes. Subjects were played a pure sine tone followed by a period of silence that varied between 2, 8, and 16 seconds, after which they were required to reproduce the original tone by using a digital sine generator. No reference tone was provided. The second part of the study involved filling the interstimulus time with distracting tones. There was no difference in the accuracy of the AP and non AP groups in the first part of the study, whereas only the AP group
were able to reproduce the target tone after tonal interference. The non musician who claimed to have AP performed better than the non AP musicians in the tonal interference task. The authors then concluded that the non musician had AP and that this study disproved the hypothesis that early musical training was required for the development of AP.

The weakness of Ross et al’s study is that it had the possibility of reliance of short term memory and possibly echoic memory of several seconds, rather than the long term memory of pitch and associated names/labels. On this basis, it would be difficult to draw firm conclusions about AP, nor can definitive conclusions be drawn on the basis of a single case study.

Other researchers (Levitin, 1996, Schellenberg & Trehub, 2003, Nakata et al, 2004) have tried to design studies that move away from rarefied pitch recognition tasks that require prior musical training. These experiments have tested recognition of the original key of popular TV themes when played either in original form or transposed up or down one or two semitones.

A study by Schellenberg and Trehub (2003) looked at pitch memory in adults over extended periods. In order to do so, they looked at 48 college students (of varying degrees of musical training, none of which reported having AP) with familiarity of 6 TV programs from which excerpts were selected. They were played 5s excerpts of the TV themes at either their original pitch level or shifted upward or downward by 1 or 2 semitones. Performance of identification of the original pitch level exceeded chance levels, with better performance for the larger transpositions. The second part of the
Schellenberg and Trehub’s study was repeated in Japan by Nakata et al this year. They compared a group of 56 Japanese children with 50 Canadian children who were 5 to 6 years of age and 41 Japanese university students with 48 Canadian university students. The Japanese children’s performance was better than chance levels but still poorer than that of Japanese adults. Canadian children performed at chance levels, and whilst both adult groups performed at better than chance levels, Japanese adults were significantly more accurate than Canadian adults. This study also found that good pitch memory for familiar tunes is widespread, and that it was independent of formal musical training.

The authors attributed the cross cultural differences to possible early educational practices in Japan, for example, piano accompaniment for children’s singing or other musical education practices like the Yamaha system and the “fixed do”. However, it
is noted that cross cultural differences were apparent in early childhood, before formal music lessons had begun.

This is in concordance with a statement by Miyazaki (2004) and Vitouch (2003) who note a higher prevalence of AP possessors amongst professional musicians in Japan (at least 50%) than in Europe or North America (1% to 20%). This higher prevalence has been attributed to musical training involving the "fixed do" in early childhood with most music students in Miyazaki’s study (2004) having begun musical training between the age of 3 to 6 years possibly during a critical period. At the same time, the proportion of AP possessors drops as the age of commencement of musical training increases, supporting the “early learning” view of AP. Levitin & Zatorre (2003) suggest that it may not be early musical training alone that leads to AP acquisition, but reinforcement of mapping of musical sounds and their labels, a “fixed do” system being a prime example of such reinforcement.

Not withstanding this, a survey of 1067 music students (Gergersen et al 2000), attempted to establish the nature of musical education before the age of 7, the presence of AP and also the ethnic background. It was noted that for all training environments the probability of AP in Asians exceeded non-Asians, this bias was even greater when taking to account very early musical training and especially training that involved “fixed do” methods. Further statistical analysis predicted that even in the absence of early musical training or exposure before the age of 7, the probability of Asian music student having AP is 0.2; amongst non-Asians, this probability is 0.03. Based on this analysis, Gergersen et al (2000) suggest that there may be some part to be played by genetic factors or assuming that a significant
proportion of the Asian students spoke Mandarin, Vietnamese and Cantonese, tone languages.

A study of American (115) and Chinese (88) Conservatory students (who all spoke the tone language, Mandarin) reported significant differences in the prevalence of AP in students from 2 major music conservatories in the US and China (Deutsch et al 2004). The methodology included on-site direct testing without self-selection from the cohort. Incidentally, once again the prevalence of AP was inversely proportional to the age of onset of musical training.

There seems to be strong evidence from the studies reviewed that the prevalence of AP is inversely related to the age of onset of musical training. Whilst the argument for a critical or sensitive period for the acquisition of AP is probably in the vicinity of 5 or 6 years, critical periods are not “brick walls” and would be expected to be distributed in a “bell shaped” curve (Levitin & Zatorre 2003). Early musical training is directly proportional to the acquisition of AP, but this critical period probably varies with the individual; the bell shaped distribution would account for the acquisition of AP in people who began musical training later than the age of 5 or 6.

Baharloo et al (1998 cited by Deutsch et al, 2004) notes that the prevalence of AP in persons who began musical training before the age of 4 is 40%, compared to 27% for those who commenced between the age of 6 and 9, 4% who began between 9 and 12, and just 2.7% after the age of 12. This seems to indicate a correspondence between the time frames for the acquisition of AP and speech and language (Vitouch, 2003, Henthorn & Deutsch, 2004).
Deutsch et al (2004) suggest that speakers of tone languages (Mandarin, Cantonese and Vietnamese) use absolute pitch, as they claim that words in tone languages take on different meanings depending on the pitch that the words are enunciated. They argue that pitch height and pitch contour define lexical tones in these languages and also hypothesise that different individuals of the same sex who speak the same dialect in these populations should match up in absolute pitch terms when words are enunciated.

Deutsch et al (1999) conducted studies of Vietnamese and Mandarin speakers who read out a list of 10 words twice on 2 separate days. For each individual, the average pitch of each word was analysed by computer, and comparisons were made between the average pitches produced for the same word on different days. The pitch difference averaged across words, was within a semitone. A similar experiment was repeated by Deutsch et al (1999), this time using a list of 12 words with 20 seconds between the 2 readings and repeated on different days. Again, the results showed that the pitch differences, averaged across words within and across sessions, was within 1 semitone. The same experiment was repeated using native English speakers. This time, they found that pitch consistencies within a session were significantly greater than that derived between sessions. On this basis, Deutsch et al (1999, 2004) claim that tone language speakers use absolute pitch or an internal pitch template to enunciate words.

The experiments of Deutsch et al (2004) rely on production of speech within individuals over a number of sessions. It could be argued that everyone is fairly consistent in their production of speech possibly partly because of vocal chord muscle
memory (see next section) and feedback. Whilst Deutsch et al argue that different individuals of the same sex who speak the same dialect in these populations should match up in absolute pitch terms when words are enunciated, the experiments described have not demonstrated this, with analysis being directed at each individual’s production of speech. Their studies so far do not give a solid argument as to why speakers of tone languages should require AP, as speakers of all languages need to perceive speech with a broad range of frequencies and speech pitch, hence relational pitch would be more important in the perception of speech (McMullen & Saffran, 2004, Saffran, 2003)

5.3 Good pitch memory is widespread

Levitin (1994), studied 46 university students who were asked to reproduce by singing the start (typically a 4 bar phrase) of some popular tunes (eg Hotel California by the Eagles), their reproductions were recorded on Digital Audio tape (to avoid potential pitch and speed fluctuations of analogue tape) and compared to the artists’ original tones on CDs. Octave errors were not penalised. 24% of subjects reproduced the original pitch and 51% were within 1 semitone of the original.

Levitin states that this provides evidence for the Two Stage theory of Absolute Pitch, that of long term pitch memory and pitch labelling. It is believed that there is a very high incidence of good pitch memory within the general community regardless of musical training but only a few have acquired pitch labelling, a process for which musical training is a pre-requisite.
One of the criticisms levelled at this study was the reproduction of melodies vocally, possibly involving some vocal chord “muscle memory”, presumably from multiple repetitions of “singing along” to the song on the radio or record. Whilst there is always some degree of muscle memory in a vocal reproduction task (Ward & Burns, cited in Levitin, 1994) and the initial pitch of a vocal tone is dependant on muscle memory, studies have shown that muscle memory for pitch is not very accurate. Ward & Burns (1978, cited in Levitin, 1994) removed auditory feedback to professional singers and found that the singers made pitch production errors of up to a minor third. Hence, overall, muscle memory would not be expected to have played a vital role in the reproduction of familiar melodies Levitin’s study (1993).

5.4 Pitch associations

In keeping with the pitch labelling component of the Two component theory of Absolute Pitch, a study by Zatorre et al (1998) compared music processing by 20 AP and Non AP musicians using functional MRI.

Three PET scans were performed on each subject, the first being a baseline (control) scan using noise bursts which were acoustically matched to tones that were to be heard in the other 2 scans and the subjects were instructed to press a key after each pair. The subjects were then presented with pairs of tones and instructed to press a key after each pair. Finally, the same pairs of tones were presented again (this time in a different order), subjects were asked to make a judgement on whether the intervals were major or minor, (a task that requires relative pitch), and to press the appropriate key. Subjects were also scanned with an MRI and the size of the planum temporale determined.
The strongest difference between AP and RP (relative pitch) groups was the strong activation of the left posterior dorsofrontolateral (DFL) cortex in only the AP group when listening to tones. However, it was active in both groups in the major/minor judgement activity. The DLF has been implicated in conditional associative learning, including visual conditional associative tasks (Petrides, 1993, cited by Zatorre et al 1998). The postulated explanation for this is that the DLF activation may be due to verbal tonal associations, the AP group making that association between the pitch class and labelling and both the RP and AP groups making that association between hearing and naming an interval.

The size of the left planum temporale corresponded directly with the degree of accuracy of AP in the screening tests.

Interestingly, the posterior portions of the superior temporal plane, (including the planum temporale) contains the auditory association cortex that directly projects to the posterior portion of the dorsolateral frontal lobe. This is the area that was strongly activated in the AP group in both test conditions of tone and interval naming. Zatorre then postulated that AP might result from the interaction between the auditory association (including Wernicke’s) area and the dorsolateral frontal cortex (conditioned associative learning) which is involved in retrieval and manipulation of various types of associations to the pitch of a note. Whilst his study makes a persuasive case for AP being a verbal association process, Zatorre also raises the possibility of AP resulting from interactions in the superior temporal area and a network of brain regions including the DLF cortex, involving the retrieval and
manipulation of a variety of associations (not exclusively verbal) to the pitch of a tone.

Reviewing the evidence thus far, one may well subscribe to the Early Learning theory of AP. The early learning theory of AP is based on the assumption that AP must be learnt from exposure to music education (pitch labelling), possibly during a critical or sensitive period at a young age (less than 5 or 6 years of age) as musicians who began training later have a much lower incidence of AP (Keenan et al, 2001, Sergeant & Roche, 1973). Educational methods (eg fixed do or Yamaha method) may increase the possibility of developing AP (Nakata et al, 2004) and a genetic component may play a part in the equation (Baharloo et al, 2000, Schlaug et al, 2001.

It is generally believed that Relative Pitch may be the main mode by which most adults perceive pitch (Takeuchi and Hulse, 1993); this may be an adaptive strategy as detection of pitch relations in various transposed guises is more useful than absolute pitch in most circumstances, eg recognising the same piece of music played in a different key or sentences spoken by various people at differing pitch and timbre.

Some studies have suggested that absolute pitch may be present in infancy (Saffran 2003, Saffran and Griepentrog, 2003, Saffran et al 2000). Infants and adults were presented with tone sequences in patterns representing Absolute Pitch sequences and others with Relative Pitch sequences and their tracking of the tones was measured by timing the attention given to the sequences. Infants showed a preference for absolute pitch patterns whereas adults only detected absolute pitch patterns when given tonal structure, relative pitch patterns were recognised regardless of tonality. Infants showed a preference for tonality but tonality per se did not interfere with their ability
to track absolute pitches. Saffran makes the important point that perception of absolute pitches in the auditory world does not necessarily equate to the ability to label pitches, her working definition for Absolute Pitch in the context of her studies was “the encoding of pitch independent of its relation to sounds.” (page 35, 2003)

Even if infants have a predisposition for Absolute Pitch processing, it appears that there is a diminishing of this ability as the child develops verbal/language and musical skills. Saffran (2001) argues that there is an “unlearning” of AP in most people. Perceptual learning as reflected in language and musical structure seems to emphasise distance between frequencies rather than frequencies themselves per se, ie pitch relations. She argues that without relative pitch, musical structure and possibly phonetic structure would elude most listeners.
6. Intermodal effects

Although there is some consensus regarding the definition of tone AP being the ability to name or reproduce a given pitch class without an external reference, it is possible that the experience of AP differs amongst possessors, to quote the illustrious violin pedagogue Alice Waten “everyone experiences or has a different method for what ends up being Absolute or Perfect Pitch”(2001, personal communication.)

This is likely to be the case, as stated by Zatorre et al (1998) raising the possibility of AP resulting from activation and interactions within a network of brain regions including the DLF cortex and the superior temporal lobe including the planum temporale, involving the retrieval and manipulation of a variety of associations (not exclusively verbal) to the pitch of a tone. He suggests that retention of note names by AP possessors are not limited by verbal encoding but by multiple codes such as auditory, kinaesthetic and visual imagery (1989).

Music can be a multi sensory and highly emotive experience, as Australian composer Peter Scutlhorpe confirms in a radio interview of “how the relationships between the elements of his music evoked for him the deepest feelings (Spurling, 1999). Pitches, and the relationship between them are of particular interest to this artist. ‘C’, he said, ‘is like the pure white light of God; E, is like eternity’. He referred to Mahler’s Song of the Earth and the special significance, the sadness, of the relationship between A and G flat in the structure of the tensions and releases in that work.”(Stevens & McKechnie, under review). Hence it may be possible that individuals associate certain notes or keys with specific emotions, this may act as a cue in recalling pitch names.
It has been claimed that various keys have their own colour and mood, for example “the softness of D flat major, the brilliancy of A major” (Bachem, 1995 cited in Ward, 1999), and AP possessors may utilize this as part of the intermodal process of pitch class naming/labeling.

Zatorre et al (2001) illustrates this in a PET study of musicians, that charted CBF changes in response to music that elicited highly pleasurable responses of “shivers down the spine” or “chills. These chills were also associated with increased heart rate, electromyelogram and respiratory rate implying psychophysiological activity. There was increased cerebral blood flow in the brain areas thought to be involved in reward/motivation such as the ventral striatum, midbrain, amygdala, orbitofrontal cortex and ventral medial prefrontal cortex. These brain areas are associated with other euphoria inducing stimuli such as sex, food and party drugs. Hence, it is possible that psychophysiological feedback may also play a role in the intermodal process of AP.

It is highly probable that in addition to the final stage of auditory verbal association and pitch labelling, each individual AP possessor has their own intermodal experience that may involve psychophysiological phenomena such as heart rate, respiratory rate, muscle tension, etc, visual imagery, and spatial imagery each weighted according to that individual’s makeup (genetic and/or environmental). Hence, AP may be a general term encompassing a number of responses and processes with the end point being the labelling of pitches without an external reference.
7. Why is AP unresolved?

The main difficulties of evaluating the studies of AP are the differing definitions and small sample sizes used by various research teams.

Possibly because of the stated rarity of AP, (the commonly cited figure by Profita & Bidder, 1988 being 1 in 10000 people) the sample sizes in the AP studies have been very small, ranging from 3 subjects (Balzano, 1984, cited in Takeuchi and Hulse 1993) to 22 subjects (Miyazaki, 1988). This makes it very difficult to draw statistically significant conclusions from the studies with individual differences possibly skewing the results and the high probability of a heterogeneous group instead of a homogenous mix. Hence it would be difficult to study general trends in a small heterogeneous group of AP possessors.

Then there is there problem of definition; Baharloo’s criteria is certainly much more strict than most of the other researchers; that is, subjects are tested with a 30 tones, (studies have varied between testing 13 to 180 tones) a time constraint (“almost instantaneous”), subjects needed to record greater than 90% correct responses, other studies have accepted a huge range of between 27% to 100% with a means from 71% and 99% correct responses.

Twenty seven percent (Zatorre & Beckett, 1989 cited in Takeuchi and Hulse, 1993) of correct responses from an AP possessor seems to be an extremely low percentage and one has to question how this person could qualify as an AP possessor. Even a mean of 75% correct responses (Zatorre & Beckett, 1989 cited in Takeuchi and Hulse 1993) seems to be too low. Whilst 27% to 75% is a higher percentage that that of chance
(8.3%), and may be useful in an experimental setting, in practical terms, 27% to 75% correct in a musical setting would probably be unacceptable.

Furthermore, the definition of “correct” varies, with some groups accepting errors within 1 semitone and others being more strict and accepting pitch class alone, excluding octave errors.

When it comes to the evaluation of pitch memory in the general population (Levitin, 1993, Schellenberg and Trehub, 2003, Trehub, et al, 2004), the definition of recognition becomes even more loose, with a range of 1 semitone to 2 tones being acceptable as “recognition or “good memory”.

If there was a universally accepted standardised measure of AP, meta-analysis could be used to combine various studies to come to some statistically significant statements or conclusion.

In addition, there may be a place for sub classification of AP, eg presence of octave errors, pitch range, pitch timbre, pitch class, one may also consider inclusions of visual imagery, spacial imagery and emotional responses. However, this sub classification may be useful, but cumbersome to implement.
8. Summary and conclusions

Studies in absolute pitch have been confounded by lack of consensus in definition, methodology of AP testing and small sample size.

Notwithstanding this, there remains a fascination with AP by the public and psychologists alike. Some would argue that the inability to name pitches is a problem of verbal association, not unlike colour anomia; most people are able to name colours, yet are unable to name pitch chroma, why might this be the case (Levitin & Zatorre 2003)?

Firstly, AP needs to be defined. Perhaps the most useful definition of AP is that of Levitin (1999), who states that AP is the ability to identify the chroma (pitch class) of any isolated tone and or to reproduce any given tone without an external reference (page 255). This is a process involving pitch perception, classification and naming, long term memory and retrieval.

Whilst “true AP” would conform to the definition above, perhaps it is not an “all or nothing effect,” on the contrary, AP may be the end of a continuum of skills which evolve from total absence to latent AP, to imperfect, to practically perfect AP (Vitouch 2003).

Studies have shown that “good pitch memory is widespread” (Schellenberg & Trehub, 2003), both for recognition of the original pitch of popular tunes (an exercise in tonality) (Schellenber & Trehub, 2003, Nakata et al) and production (Levitin,
Researchers are tempted to call this latent AP or in the case of recognition, “absolute tonality” (Takeuchi & Hulse, 1993). It has been noted that whilst good pitch memory is very common, only a few are able to put labels to these notes, hence, falling short of “practically perfect AP”.

With regards to “imperfect AP”, it has been noted that AP possessors tend to be more accurate in identifying pitches in the middle pitch register, with timbre playing a part (musicians are more accurate at identifying pitches produced by their primary instrument), white notes on the piano being more easily identified than black. In addition, hormonal fluctuations such as the menstrual cycle (Wynn 1970) and metabolic changes due to drugs (eg. Carbamazepine) affect the accuracy of AP (Ward, 1999).

Finally, “practically perfect AP” seems to vary according to the research group, with accuracy of identification varying from 27% to 75% correct (Takeuchi & Hulse, 1993), “correct” being up to half a tone wrong! This is in contrast to Baharloo et al’s (1998) group who have a strict criterion of “almost instantaneous” response, 93% correct and without the “semitone” leeway.

Like the acquisition of language, there seems to be a “critical period” for the acquisition of AP, with most possessors having commenced musical training before the age of 5 or 6 (Deutsch et al, 2004, Takeuchi & Hulse, 1993), the incidence of AP varying inversely with the age of commencement of musical training (Brown et al, 2003, Deutsch et al, 2004, Levitin & Zatorre, 2003).
A genetic component has been proposed as 50% of AP possessors have a relative with AP (Baharloo et al, 1998). Miyazaki (2004) and Vitouch (2003) note that the incidence of AP amongst professional musicians in Japan is greater than 50% as compared to North America and Europe where the estimated incidence is between 1% and 20%). This may be attributed to early musical education methods amongst the Japanese musicians that involve the “fixed do”, or a genetic component. Gergersen et al (2000) conducted a survey of 1067 music students and found that the probability of an Asian student having AP is 0.2; amongst non-Asians, 0.03 even in the absence of early musical training, on this basis Gergersen et al (2000) propose that there is a genetic component to the acquisition of AP.

AP possessors have been noted on MRI to have a larger left planum temporale and greater leftward asymmetry than non AP possessors, the cause of which is unknown (Schlaug et al, 1995). Of significance, it has been found on MRI and fMRI that the physical size (Zatorre et al, 1998) of the PT is positively correlated with the degree of accuracy of AP and the degree of activation of the left planum temporale was positively correlated with AP accuracy but negatively correlated with the age of commencement of musical training (Oinishi et al, 2001). In spite of this, there was no correlation of duration of musical training and cerebral activation. This correlates with the notion of a critical period for the acquisition of AP, cerebral plasticity and repeated musical practice at an early age may be responsible for the noted changes, in line with cerebral plasticity being optimal during the early years (Gougoux et al, 2004).

The posterior portions of the superior temporal plane containing the planum temporale contains the auditory association cortex (involved in pitch processing...
Binder et al, 1996, Mazziotta et al, 1982) and projects to the dorsolateral frontal lobe. Zatorre et al (1998) postulated that AP may result from the interaction between the auditory association area (including the left planum temporale/ Wernicke’s area) and the dorsolateral frontal cortex (conditioned associative area) which is involved in the retrieval and manipulation of a variety of associations to the pitch of a note. Zatorre et al (1998) make a persuasive case for AP not just being a verbal association process, but raises the possibility of a variety of associations (not exclusively verbal) to the pitch of a tone. Thus, AP possessors may not be limited by verbal encoding, but by multiple codes such as auditory, kinaesthetic and visual imagery, with the enlarged left planum temporale simply being a “marker” for multiple interactions within a network of brain regions involving the left PT and the DLF cortex.

Whilst there is very suggestive evidence for a critical period for the acquisition of AP, some researchers may argue that AP is present in infants (Saffran, 2003, McMullen & Saffran, 2004). When presented with tone sequences involving AP or RP sequences, infants appeared to show a preference for AP pitch patterns. Saffran makes the point that whilst infants may have a preference for perceiving absolute pitches in their auditory world, it did not equate to the ability to label pitches, hence, one may argue that infants have a form of “latent AP” (Vitouch, 2003). Although there may be an “unlearning” of AP pitch processing in favour of RR pitch processing (Saffran 2003), many people retain very good pitch memory (Levitin, 1993, Saffran 2003, Saffran & Griepentrog, 2001)

This brings us back to Vitouch’s (2003) suggestion that the acquisition of AP may be a continuum from latent, to imperfect to almost perfect AP. This acquisition may
involve a genetic component coupled with a critical period of musical instruction (possibly involving the “fixed do” system) before the age of 6, possibly reflected anatomically in an enlarged left planum temporale, and a leftward asymmetry of the planum temporale. Despite the correlation of the critical period for the acquisition of AP with the acquisition of language, AP may not be a purely verbal- pitch association and may involve other associations with pitch such as emotional, psychobiological changes and visual imagery, with the end point being labelling of pitches without an external reference.
References


Nakata, T, Trehub, S.E., Schellenberg, E.G.,(2004), Cross-cultural Perspectives on Pitch Memory


Schellenberg, E.G., and Trehub, S.E., (2003), Good Pitch Memory is Widespread. *Psychological Science, 14* (3) 262-265


Vitouch, O., (2003), Absolutist models of absolute pitch are absolutely misleading, Music perception 21, (1), 111-117


