Differences in Spectral and Fundamental Hearing between Absolute and Relative Pitch

Martin Ebeling[a] and Gerald Langner[b] †

[a]*Institute of Music and Musicology, Dortmund University of Technology,*
Emil-Figge-Str. 50, D-44227 Dortmund, Germany martin.ebeling@tu-dortmund.de
corresponding author

[b]*Institute of Biology - Neuroacoustics, Darmstadt University of Technology,*
Schnittspahnstr. 3. D-64287 Darmstadt, Germany
† Prof. Dr. Langner died December 3, 2016.

Keywords: absolute pitch; fundamental hearing; spectral hearing; periodicity detection; inner oscillations
Doi: https://doi.org/10.61712/HNPK7403

Abstract

Consider a stimulus consisting of three sine tones with the same frequency spacing. For some listeners such a stimulus elicits a residual tone. Obviously, their hearing system interprets the three sine tones as three adjacent partials of a complex tone with a fundamental frequency equal to the frequency spacing. On the other hand, some subjects easily hear out some or even all single sine tones.
One can therefore distinguish between two hearing modes: the case of the residue pitch is referred to as fundamental hearing, whereas the ability to resolve single partials is called spectral hearing or overtone hearing.
Subjects may differ considerably in their capability for fundamental or spectral hearing. Depending on factors as the frequency region of the stimuli or the frequency of the residue pitch (missing fundamental) or depending on psychological factors as attention or individual hearing dispositions, subjects may even change between both modes of hearing.
The present study demonstrates that possessors of absolute pitch show significant differences in spectral versus fundamental hearing compared to possessors of relative pitch. In some frequency regions, possessors and non-possessors of absolute pitch prefer either fundamental hearing or spectral hearing quite individually and often may change between both modes of hearing. However, statistically, possessors of absolute pitch show a significantly higher preference for perceiving a residue tone than possessors of relative pitch do.

1. Introduction

1.1. Absolute pitch

Absolute pitch (AP) is defined as the ability to recognize and name musical tones by ear without any external or internal reference tone. Most commonly, possessors of AP are also able to sing the correct pitch of a tone given only its note name without using a reference tone. In contrast, trained possessors of relative pitch (RP) are able to determine intervals, but they cannot name notes without a reference pitch from a pitch fork or a musical instrument.

Due to the phenomenon of octave identification, (almost) all tonal systems are cyclic: the tonal structure and the note names are periodically repeated in distances of an octave. Thus, pitch perception refers to two dimensions of a tone. One dimension is cyclic as described by the note names. This dimension was termed *Tonigkeit* by v. Hornbostel (1926), and later *tone chroma* by Bachem (1950). The other dimension was called *Tonhelligkeit* by v. Hornbostel, and *tone height* by Bachem. It refers to the fact that we can also determine the height of a tone and thus its octave distribution. As a result, we distinguish between tones of different octaves but with the same note name and say that they belong to the same pitch class. Note naming refers to the chroma dimension. It is quite remarkable that possessors of AP are certain and quite fast in naming tones. However, like possessors of RP, they need a moment of reflection before they decide on the octave distribution of the tones. AP refers exclusively to the cyclic chroma dimension of pitch perception.

The root of AP is still quite unclear. Consequently, all theories of AP are yet speculative; they can be divided into four groups, which are *heritage theories*, *unlearning theories*, *learning theories*, and *imprinting theories* (Ward & Burns, 1982).

Innateness is implied by *heritage theories* of AP. Possessors of AP are four-times more likely to have AP possessors among the other family members than non-possessors (Baharloo, 1998). Among the probands with AP of the present study, three sisters reported that their mother, a professional pianist and piano teacher and their grandfather also had AP. In contrast, other AP possessors stated that none of their relatives was known to have AP.

So-called *unlearning theories* claim that the RP is a result of a learning process: as soon as the ability to compare pitches is developed, pitch relations are assumed to become so important that relative pitch judgements dominate and the innate AP is finally given up (Abraham, 1901).

On the other hand *learning theories* argue that the AP is acquired by learning. To support this idea Brady (1970), a non-possessor of AP, trained his own tone identification ability on the basis of memorized anchor tones up to a level comparable to real AP. Finally, he was convinced that an adult can acquire AP at the cost of exhaustive training but also pointed out that “as with many skills, the younger it is learned, the better. Beyond childhood, any improvement in ear training is apparently hard earned.” (see also Corliss' critical objections, 1973).

This remark is close to *imprinting theories* which refer to the observation that possessors of AP usually had early musical training or they had been brought up in environments where music has had a great importance. Deutsch et al. (2006 / 2009) "have found AP to be most prevalent among those who had begun musical training by ages 4–5, less prevalent between ages 6–8, and very rare after age 9". More general, an imprinting phase far before the tenth year of age is assumed in which exposure to pitch based information boosts the genesis of AP.

Music is not the only source of pitch based information. For example, tonal languages such as the Mandarin Chinese involve pitch to convey the meaning of words which otherwise would be ambiguous. The pitch contour of the vowels have a semantic function in tonal languages. Deutsch et al. (2006) found a prevalence of AP in students whose first language was a tonal language. She concluded that the age of onset of musical training meeting the imprinting phase is one decisive factor to acquire absolute pitch. But in case of tonal languages the importance of pitch for the meaning of words may also provoke the acquisition of AP in infants during speech-related critical periods. Having attained AP for speech recognition it is then also available for tone identification. Typically, less than 0.1% of the population in western countries are estimated to possess AP. Profita (1988) reports the observation of music teachers who have taught children for many years, that the incidence of AP is about 1:1500 or less in the general population. However, Deutsch found, that "60 percent of Beijing students who had begun studying music between the ages of four and five years old passed a test for absolute pitch, whereas only 14 percent of the American students did" (Monroe 2004).

The four kinds of theories to explain the genesis of AP point to important aspects of AP but are contradictory in part, such that the auditory and cognitive processes of AP are still unclear.

1.2. Spectral versus fundamental hearing

Tones produced by instruments or the singing voice are commonly perceived as entities characterized by different sensational moments (Stumpf, 1883) to which different tonal attributes such as pitch, timbre, loudness and duration are ascribed. In contrast to the perceived entity, an instrumental tone contains a series of partial tones with frequencies that are integer multiples of a fundamental frequency. The fundamental frequency corresponds to the perceived pitch. Changing the loudnesses of the partials perceptionally elicits a change of timbre but generally does not change the percept of the entity of the tone and its pitch. The perception of a complex tone is an example of primitive grouping (Bregmann, 1990): a plurality of tones is grouped to a perceptional whole or an auditory entity with a few sensational moments. Primitive grouping results from neuro-physiological processes. The perception of the pitch of the common fundamental is due to periodicity coding and periodicity analysis in the auditory system (Langner and Schreiner, 1988; Langner, 2015). Note, that the waveform of a sum of harmonically related sine oscillations is strictly periodic. Its period does not depend on the strength of the single partials or on phase-shifts.

The term 'residue tone' is assigned to the phenomenon that the pitch remains unchanged if the fundamental tone of a series of harmonic partials is missing (Schouten et al. 1962). As long as omitting the fundamental tone (and perhaps also other low partials as in the present experiment) does not affect the period of the waveform, only the timbre and the loudness are changed. If the auditory system exploits the period of the signal waveform to determinate the pitch, the new waveform without the fundamental must have the same pitch as the original signal, namely the pitch of the now missing fundamental. Three (or sometimes even two) adjacent partials may suffice to evoke a residue pitch because they sum to a waveform with a period equal to the period of the fundamental. The test items of the present experiment are composed from stimuli of this kind.

Note, that three adjacent partials with appropriate phases and amplitudes are equivalent to a sinusoidally amplitude modulated sine (SAM), as its spectrum consists of three components: the carrier frequency f_c and the two side-bands $f_c - f_0$ and $f_c + f_0$ (Hartmann 1998). The residue pitch of the amplitude modulated tone corresponds to the frequency difference f_0 of the partials, which is equal to the frequency of the amplitude modulation.

Despite the unified tone percept, it was always known that a musical tone contains partials that can be singled out. For example, without any scientific acoustic theory the German composer and music theorist Michael Praetorius (1571-1621) mentioned the influence of single partials onto the timbre of musical instruments (Muzzulini, 2006). Since Helmholtz's (1862) examinations, the acoustic principles of complex tones had become clear and provoked the psycho-acoustic question why and under what conditions single partials can be perceived.

In reference to observations of Mersenne (1588-1648), Stumpf (1890) found that the seventh and even ninth partials can be heard out with ease under certain conditions. He pointed out that the attention of the listener is a highly important psychological factor in analysing single overtones and that some listener may switch voluntarily from hearing the fundamental pitch to analysing the overtones.

Modern hearing theory starts from the concept of critical bands (Fletcher 1940). The critical bands are described as auditory filters (Patterson 1976 / 1995) with a frequency dependent width. Critical bands are scaled by the bark scale (Zwicker 1961). Harmonics can be heard out if, and only if, neighbouring harmonics fall into different critical bands (Plomp, 1964; Plomp and Mimpen, 1968) and are thus resolved in different auditory channels (Cheveigné, 2000).

Moore et al. (1984) found:

- A single partial may stand out from the unified tone sensation if its loudness is incremented.
- Intensity difference limes for the lower partials are small but increase abruptly after the fifths or seventh partial.
- The highest partial of a complex tone is quite discernible.
- The pitch of the complex tone is first of all determined by the most resolvable partials which are normally the first to fifth partial. However, note that these partials are also best coded in the temporal domain.

Inharmonicity, which is the mistuning of single partials leads to *quasi*-periodic waveforms and enhances the audibility of the single inharmonic partial. It may also induce pitch shifts of the fundamental (Moore et al. 1985). The difference limen for the detection of mistuning is smaller for lower partials than for higher ones (Moore et al., 1984).

A partial that is alternately turned on and off, stands out as a single pure tone even up to the 20th partial. This kind of enhancement must be due to another effect than resolvability as in the case of higher partials several overtones fall into one auditory filter, masking each other. Obviously, loudness increment unmasks a single partial (Hartmann, 2006).

As already mention, in principle, there are two hearing modes to perceive harmonically related sine tones. Either the residue pitch corresponding to the fundamental frequency is perceived or the individual sine tones are heard. Focussing onto the residue pitch or fundamental is called fundamental hearing (*fm-hearing* or *fm*). For the hearing of the single partials the term spectral hearing (*sp-hearing* or *sp*) has been established, although *overtone hearing* would actually be more appropriate. Reducing the number of lowest partials of the harmonic tone reveals individual preferences for on or the other of both hearing modes: Some subjects hear the fundamental pitch, others discern the individual partials and some perceive both or switch between both hearing modes. Inter-individual differences in the decision for one of the hearing modes can be demonstrated.

In the present experiment, we apply the same kind of stimulus as Schneider et al. (2005) did to classify subjects either to be fm listeners or sp listeners. In their investigation, MRI (magnetic resonance imaging) studies revealed a pronounced leftward asymmetry of gray matter volume in fm listeners and a rightward asymmetry for sp listeners. Additionally, magnetoencephalography studies demonstrated a functional asymmetry of the auditory evoked P50m source activity of lateral Heschl's gyrus in response to harmonic complex tones. P50m magnitude was relatively larger in the left hemisphere for fm listeners, but for sp listeners, it was larger in the right hemisphere, irrespective of musical aptitude.

2. Hearing Experiment

2.1. Test items

Each test item consists of two complex tones T_1 and T_2 presented consecutively. They are constructed in such a way that fm-hearing results in a different pitch percept than sp-hearing. The task is to decide whether the melodic line from T_1 to T_2 ascends or descends.

Each complex tone is composed of three adjacent harmonic partials from an overtone series over the missing fundamental frequencies $f_{0,1}$ of tone T_1 and $f_{0,2}$ of tone T_2. In other words, each tone T_1 and T_2 consists of three sinusoidal tones equally spaced in frequency by $f_{0,1}$ and $f_{0,2}$, respectively. Moreover, the highest partial of T_1 which has the frequency $n \cdot f_{0,1}$ (n a positive integer between 5 and 12) and the highest partial of T_2 which has the frequency $(n+1) \cdot f_{0,2}$, are equal. It follows for the fundamental frequencies:

$$n \cdot f_{0,1} = (n+1) \cdot f_{0,2} \Rightarrow f_{0,1} > f_{0,2} \quad (1)$$

For the lower two partials, a simple calculation shows that:

$$\begin{aligned} (n-1) \cdot f_{0,1} &< n \cdot f_{0,2} \\ (n-2) \cdot f_{0,1} &< (n-1) \cdot f_{0,2} \quad (2) \end{aligned}$$

Proceeding from T_1 to T_2 and taking the corresponding partials as voices, the following tonal movements result (see Fig. 1):

- the residue pitch descends, according to (1) ,
- the common highest partial remains the same;
- both lower partials ascend, according to (2).

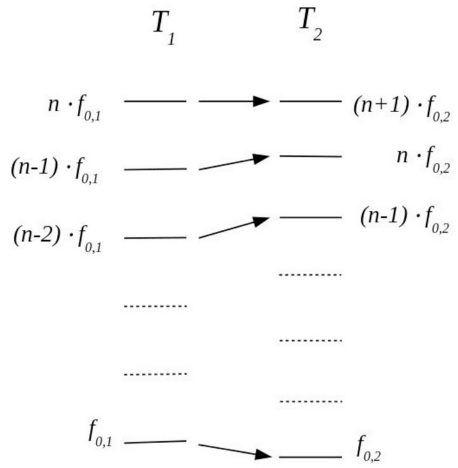

Figure 1. Schematic drawing of the two-tone stimulus. The harmonic numbers of the highest partials with the common frequency are n and n+1. Both lower partials rise in frequency whereas the fundamental frequencies descend. Here, the progression is from n=6 to (n+1) = 7 and the unchanged frequency of the highest partial is: $6 \cdot f_{0,1} = 7 \cdot f_{0,2}$.

Due to the phenomenon of residue pitch, in the case of fm-hearing, two successive low pitches are heard corresponding to the missing fundamental frequencies $f_{0,1}$ and $f_{0,2}$ of the complex tones T_1 and T_2, respectively. In the case of sp-hearing the perception is dominated by the partials so that a high pitch percept (or if completely resolved three high partials) are heard. As sp-hearing focuses on the partials, an ascending tonal movement is perceived in this case, whereas fm-hearing results in a descending pitch percept.

Analogously, if the test item is reversed, a descending tonal movement is perceived in the case of sp-hearing whereas in the case of fm-hearing the perceived pitch ascends. In short: spectral and fundamental hearing result in the perception of opposite directions of tonal movements.

To mask combination tones, white noise with a level of 45 dB was added during the whole stimulus duration of 3.8 sec. The first tone is introduced 1 sec after noise onset by a linear ramp of 50 ms duration to avoid clicks. The tone is sustained for 500 ms with a level of 75 dB and fades out by a linear ramp of 50 ms duration. After a gap of 600 ms the second tone is presented. It is also introduced and faded out by linear ramps of 50 ms duration and again its sustain time is 500 ms at a level of 75 dB (see Fig. 2). The stimuli were generated with the visual programming language *pure data* (https://puredata.info/).

Figure 2. Time course of one test item. A stimulus consists of two complex tones. Each tone is flanked by 50 ms ramps and sustained for 500 ms. The gap between both tones has a duration of 600 ms . The noise is added to mask combination tones.

2.2 Test design

Throughout the test, the frequency of the highest partial rises from 500 Hz to 10000 Hz in steps of 500 Hz ($n \cdot f_{0,1} = (n+1) \cdot f_{0,2} = k \cdot 500$ Hz, $k = 1, 2, \ldots, 20$). On the whole, we tested twenty different frequencies of the highest partials from 500 Hz up to 10000 Hz. For each of these frequencies eleven pairs (n, n+1) with n=4, 5, … , 14 were chosen to represent the harmonic numbers of the highest partials: (4, 5), (5, 6), … , (14, 15). Thus, we tested 22 residue pitches for each highest partial frequency, two for each of the eleven pairs. The *missing fundamental frequencies* $f_{0,1}$ of the harmonic numbers n range from 500/14 Hz = 35,7 Hz to 10000/4 Hz = 2500 Hz, the *missing fundamental frequencies* $f_{0,2}$ of the harmonic number n+1 range from 500/15 Hz = 33,3 Hz to 10000/5 HZ = 2000 Hz. On the whole, we constructed 20 · 11 = 220 pairs (T_1, T_2) of stimuli. Each pair of tones was presented in the original order $T_1 \rightarrow T_2$ with a descending residue pitch (fw-direction) and in reverse order $T_2 \rightarrow T_1$ with an ascending residue pitch (bw-direction), so that each subject responded to a total of 440 test items. For each participant, the whole test was spread over four separate sessions held on different days and lasting about twenty minutes each, including rests.

We presented the test items in randomized order separating succeeding test items by a break of three seconds. During this short break, subjects had to respond on a questionnaire whether the tonal movement ascended or descended, no alternative responses were admitted (forced choice experiment). None of the subjects complained that the tests were too difficult.

2.3 Participants

Fifty test subjects participated in the tests. Forty-two of them were students with more than ten years of musical education, 8 of them were lecturers with a musical background or instructors of music from music universities or high-schools of music.

Sixteen participants (12 female, 4 male) had absolute pitch (N_{AP}=16), among them two female and one male instructors of music.

Thirty-four participants (14 females, 16 males) had relative pitch, among them two male instructors of music.

2.4 Presentation

For each of the 50 subjects, all tests were presented in four individual sessions via headphone (AKG K 271 MK II Closed-Back Studio Headphones). The tests had been stored on a notebook before and were processed by an external M-audio sound card (*Delta Audiophile 2496)* with a linear frequency response. At the Institute of Electro-acoustics of the Ruhr-University, Bochum, headphone and sound card where calibrated by means of an Articifial Ear Type 4152 and a Real Time Analyser 840 by Norsonic.
The test tone had a frequency of 400 Hz. An amplitude of 1 of the audio editor *audacity* corresponded to 106 dB SPL on the left ear cup and 106.1 dB SPL on the right ear cup. Presented a complex tone, its single partials differed by less than 3 dB between left and right ear cups.

3. Data Analysis

3.1 Preparations for the statistic analysis

The test items are categorized into twenty classes N_k corresponding to the twenty highest partials. Their frequencies are equal to $k \cdot 500$ Hz (k = 1, 2, 3, … , 20). Each class N_k encompasses the eleven test items corresponding to the eleven pairs of harmonic numbers (n, n+1) with n=3, 4, … 14 (see 2.2).
The response to a test item reveals whether the subject focused onto the residue pitches (fm-hearing) or onto the high partials of the stimulus (sp-hearing). We introduced a response index r_n. In case of a response indicating fm-hearing, we rated $r_n = 1$. In the contrary case of sp-hearing, we set $r_n = -1$. All responses r_n were tabled for each subject. For each of the twenty classes and for each participant, the eleven responses r_n are added up to a class response index s_k:

$$s_k = \sum_{n=3}^{14} r_n \quad (3)$$

A negative class response s_k indicates the predominance of spectral hearing in class N_k whereas a positive sum shows that the subject predominantly perceived a fundamental pitch for the items of class N_k. Corresponding to the twenty classes N_k, the whole dataset of individual class responses encompasses 20 sums s_k for each subject. Thus we define an individual response-vector R with the twenty sums of class responses s_k:

$$R = \left(s_k\right)_{k=1,\dots,20} \quad (4)$$

These individual response vectors are tabled and can be graphed (see Fig. 3).
N_{AP}=16 subjects with AP and N_{RP}=34 subjects with RP participated. Each test item was presented in forward and backward direction. Thus, the whole dataset contains $2 \cdot (N_{Ap} + N_{RP}) \cdot 20 = 2000$ individual class responses.
To look for differences between AP and RP, the following classification of individual class responses is applied:

First, we can separately consider the individual class responses for both directions. Then we pool them together which results in twice as many responses. The corresponding data classes are designated as follows:

fw – **forward** direction;
bw – **backward** direction;
bo – data from **both** directions are considered together.
We may also separate f_m-hearing from s_p-hearing or make no distinctions between hearing modes. The corresponding data classes are designated as follows:
fm – **fundamental** hearing
sp – **spectral** hearing
all – **all** data are considered regardless the hearing mode.

By combining both classifications, these distinctions lead to nine classes of data and thus nine statistical tests as listed in Table 1:

Table 1: List of the nine rank sum tests.

	All hearing modes	Fundamental hearing	Spectral hearing
Both directions	test 1: bo_all	test 4: bo_fm	test 7: bo_sp
Forward direction	test 2: fw_all	test 5: fw_fm	test 8: fw_sp
Backward direction	test 3: bw_all	test 6: bw_fm	test 9: bw_sp

Data are classified according to the hearing modes and the directions of the stimuli. In each of the nine classes, a rang sum test has been performed to compare the responses of the AP with those of the RP.
For comparison of AP and RP we must normalize by the numbers N_{AP}=16 of subjects with AP and N_{RP}=34 of subjects with RP respectively. If we jointly consider the individual class responses of both directions of presentation, we have to normalize by $2{\cdot}N_{AP}$=32 for subjects with AP and $2{\cdot}N_{RP}$=68 for subjects with RP. Thus, to consider the average ratings of AP (N_{AP}=16) and RP (N_{RP}=34), the average response-vectors RAP and RRP are defined by:

$$RAP=\frac{1}{16}\sum_{i=1}^{16} R_i \qquad RRP=\frac{1}{34}\sum_{j=1}^{34} R_j \quad (5)$$

where R_i designates the relevant response-vectors of the possessors of absolute pitch AP, and R_j those of the probands with relative pitch RP.
If the responses of both directions are considered together, each proband has responded twice. In this case we have N_{AP}=32 and N_{RP}=68. and define:

$$RAP=\frac{1}{32}\sum_{i=1}^{32} R_i \qquad RRP=\frac{1}{68}\sum_{j=1}^{68} R_j \quad (6)$$

Note, that RAP and RRP are vectors with 20 entries each. The entries are the average ratings of AP and RP respectively for the 20 classes determined by the 20 different highest partials of the test items.

3.2. Results of rank sum tests

To examine for differences between AP and RP in the nine classes of the datasets of table 1, the corresponding vectors *RAP* and *RRP* are compared by a *paired Wilcoxon rank sum test.* Table 2 shows the results of the tests. Consult the figures indicated in the last column.

Table 2: Results of the rang sum tests in the nine classes.

test	class	Rp	Rn	R [%]	p	%	significance	Fig
1	bo_all	170	40	5.00%	0.014	1.36%	s	4
2	fw_all	140	70		0.202	20.24%		5
3	bw_all	189	21	1.00%	<0.001	0.09%	hs	5
4	bo_fm	191	19	0.10%	<0.001	0.06%	hs	
5	fw_fm	184	26	1.00%	<0.002	0.20%	hs	6
6	bw_fm	191	19	0.10%	<0.001	0.06%	hs	6
7	bo_sp	144	66		0.154	15.36%		
8	fw_sp	102.5	107.5		0.940	94.05%		6
9	bw_sp	175	35	1.00%	0.007	0.73%	hs	6

Next to the positive and negative ranks Rp and Rn, the probabilities (probability p and % of probability) and the significance levels (s: significant, hs: highly significant) are listed. We only use the significance level as a descriptive variable and not for decision making. Therefore, we report both significant and highly significant results.Compare the figures indicated in the last column.

In case of fm-hearing (rising partials and falling fundamental), test 5 fw_fm and test 6 bw_fm demonstrate a highly significant difference between AP and RP for the fw-direction ($p = 0.002$, Fig. 6a) as well as for the bw-direction ($p = 0.0006$, Fig. 6b). As a consequence, test 4 for both directions tested together also shows highly significant results in case of fm-hearing ($p = 0.0006$) and in the end causes the significant overall difference between AP and RP as demonstrate by test 1 bo_all (Fig. 4).

In case of sp-hearing (falling partials and rising fundamental), the fw-direction shows no significant difference between AP and RP (test 8: $p = 0.94$, Fig. 6a). In contrast, the bw-direction demonstrates an even highly significant difference between AP and RP (test 9: $p = 0.007$, Fig. 6b). The directional tests reveal that this is due to the directional sensitiveness of RP (see below and Table 3). The average response vector R_{fw} for items presented in forward direction (fw, the residue pitch descends; see 2.2) can be compared to the average response vector R_{bw} to items presented in backward direction (bw, the residue pitch rises). The corresponding response vectors for AP and RP can also be graphed (see Fig. 4 and 5).

Again we applied a *paired Wilcox rang sum test* on to the average response vectors for AP against those for RP to test for differences in directional sensitivity.

The six directional tests are

d-test 1/2: fw_all against bw_all for AP and RP respectively;
d-test 3/4: fw_fm against bw_fm for AP and RP respectively;
d-test 5/6: fw_sp against bw_fm for AP and RP respectively.

Table 3: List of directional tests.

d-test	Test items	Rp	Rn	R	p	%	significance
1	AP_fw_bw_all	131	59		0.15	15.3	
2	RP_fw_bw_all	183	27	0.50%	0.004	0.38	hs
3	AP_fw_bw_fm	126	84		0.44	44.33	
4	RP_fw_bw_fm	161	49	5.00%	0.04	3.83	s
5	AP_fw_bw_sp	132.5	77.5		0.31	31.31	
6	RP_fw_bw_sp	191.5	18.5	0.10%	0.001	0.1323	hs

Data were tested for directional sensitivity by six directional tests. In contrast to AP hearers, listeners with RP show (highly) significant sensitivity to the direction of stimulus presentation. Next to the ranks Rp and Rn the levels of significances are listet. Compare Fig 7.

For RP, differences between the ratings of presentations in fw-direction and bw-direction are highly significant in case of sp-hearing (d-test 6: p = 0.0013), significant in case of fm-hearing (d-test 4: p =0 .04) and highly significant if the hearing modes are not differenciated (test 2: p = 0.0038).
In contrast, the responses of AP show no directional differences between both directions of presentation.

4. Results

4.1 Fundamental versus spectral hearing in AP and RP

To a first approximation our observations confirm the results reported by Schneider et al. (2005) that subjects can be separated into fm-hearers and sp-hearers. In an extensive test series Schneider made use of the same stimulus type (see 2.1) to demonstrate that both contrasting modes of perception (fm- and sp-hearing) correspond to different lateralisations in auditory cortex (*gyrus of Heschel*). We adopted the stimulus type which Schneider applied but extended the frequency range of the highest partial from 4430 Hz to 10000 Hz (see 2.2). Moreover, we focussed on subjects with absolute pitch versus those with relative pitch.

4.2 Observations from individual responses

At first sight the results do not reveal differences between AP and RP in relation to fm- and sp-hearing. Both modes of hearing can be found in the classes of AP as well as RP, as we demonstrate with four examples from the forward direction presentations. Among the possessors of AP, subject AP3 is spectrally orientated over the whole frequency range while AP14 turned out to be consistently fundamentally orientated (Fig. 3a). The same extremes can be found among the possessors of relative pitch: subject RP14 is spectrally orientated over the whole frequency range while RP7 is consistently fundamentally orientated (see Fig. 3a). On the other side, in both classes some subjects alternate their pitch preferences from fundamental hearing in the low frequency range to spectral hearing in the higher frequency range, as subjects AP11 and RP12 do. In contrast subject RP21 changes from spectral hearing in the low frequency range to fundamental hearing in the higher frequency range. Subject AP16 is spectrally orientated in the lower frequency range of the highest partial but ambiguous in the higher frequency range (see Fig. 3b).

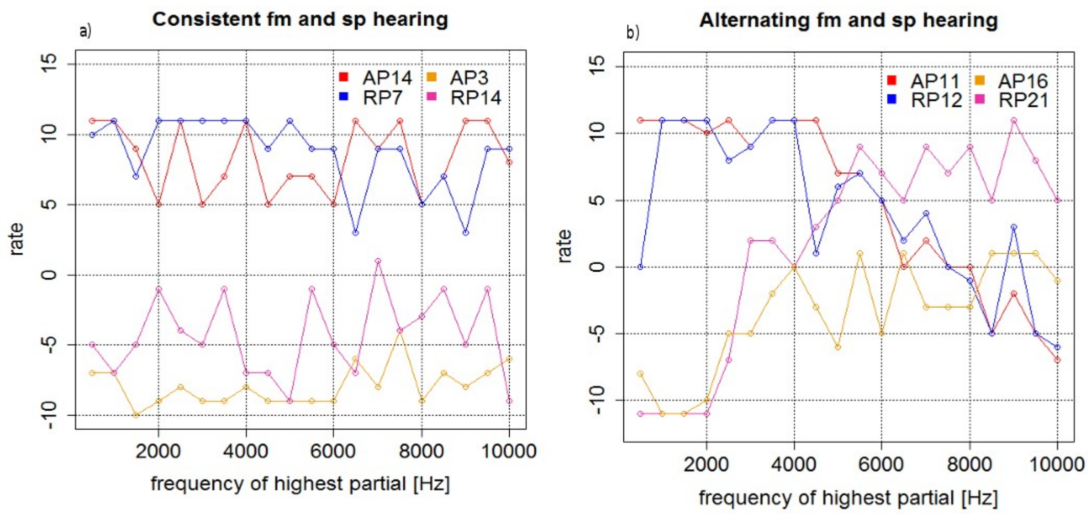

Figure 3. Individual preferences of fm-hearing and sp-hearing. Some subjects are consistently orientated either to fundamental or spectral hearing (left panel). Others change the hearing mode from fundamental hearing to spectral hearing, or still others change from spectral to fundamental hearing. There are subjects of all types among the AP-group as well as the RP-group (right panel).

4.3 Remarkable introspections of some subjects

During the test, subject AP17 reported that she sometimes heard both, a high pitch or sometimes even more than one high pitch and simultaneously a fundamental pitch. But she would focus onto the fundamental tone as she had learned that this pitch is the important tone in music. Another possessor of AP (AP1) was able to switch voluntarily from fundamental to spectral hearing and vice versa. During a short brake, she remarked: “At the end I had the impression that I could also hear it the other way round”. So, we repeat the last part of the test and indeed, she switched the hearing mode perfectly. Again she responded quickly and decidedly and asserted that she had only one distinct perception which was not at all ambiguous. A pianist with AP (AP6) was able to name the tones of the stimuli and she could replay them, all three spectral tones and the fundamental tone, on the piano. All three participants were instructed to follow the dominating perception in each case, which they did without hesitation. After the test, they reported that they had no doubt about their decisions.

The other participants affirmed that they always heard only one clear pitch. Sometimes, but quite seldom, other participants reported to perceive weak additional tones in the extreme frequency ranges.

4.4 Results from test statistics

4.4.1 Class bo_all

In contrast to the first diffuse impression a *paired Wilcoxon rank sum test* reveals a **significant dominance of fundamental hearing in possessors of absolute pitch compared to listeners with relative pitch** ($p = 0.014$, test 1). This difference between AP and RP is mainly due to the frequency region below 5000 Hz as can be seen from Figure 4. Accordingly, a rank sum test regarding only stimuli with highest partials below 5000 Hz not surprisingly proves an even highly significant dominance of fundamental hearing in AP compared to RP ($p = 0.002$). Note, that the periodicity analysis of the auditory system is restricted to the frequency region below 5000 Hz (Zwicker & Fastl 1990, Langner 2015).

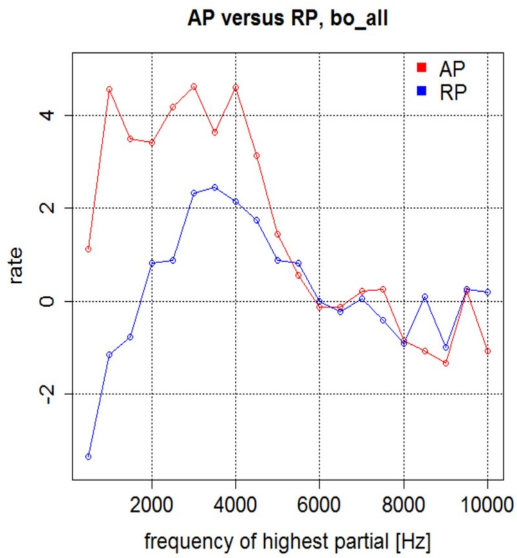

Figure 4. Dominance of fundamental hearing in possessors of AP compared to RP. A rank sum test applied to the pooled data demonstrates a significant difference between AP and RP: in the frequency range below 5000 Hz listeners with AP prefer fm-hearing much more extensively than listeners with RP.

4.4.2 Classes fw_all and bw_all

Separating the data obtained in the fw-direction – the missing fundamental descends – from the data obtained in the bw-direction – the missing fundamental rises – and testing all data by a rank sum test also shows the same preference for fm-hearing in AP for both directions (see Fig. 5). For frequencies above 2500 Hz the difference between AP and RP becomes smaller and smaller for the fw-direction, so that the rank sum test only confirms a tendency (test 2: $p = 0.2$, see Fig. 5 a). In contrast, for the bw-direction the difference is highly significant (test 3; $p = 0.0009$, see Fig. 5b).

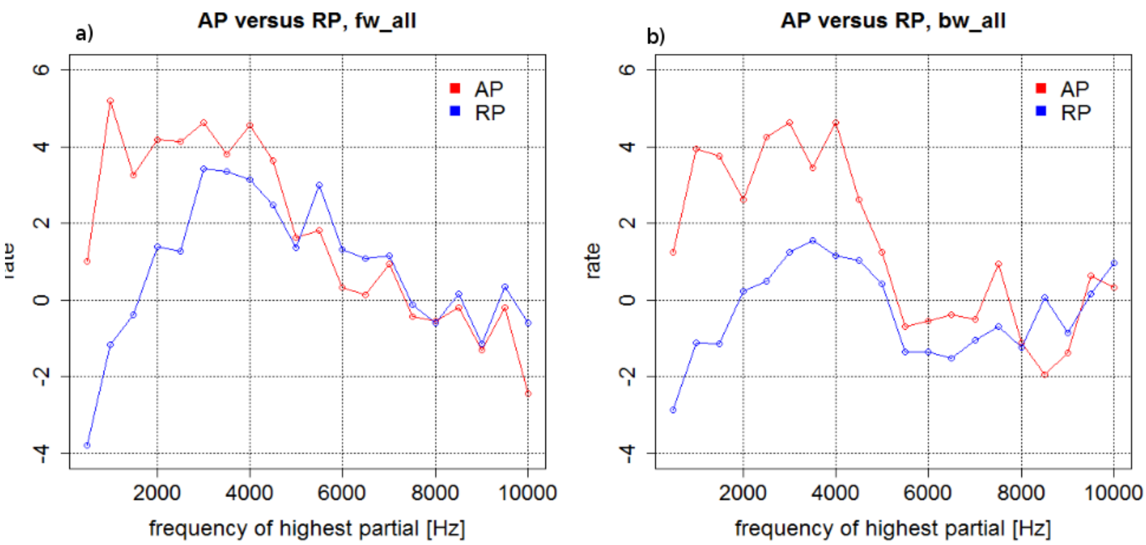

Figure 5. More extensive fm-hearing in AP than in RP. The higher degree for the preference of fm-hearing in AP is obvious in the frequency range below 4000 Hz. The difference between AP und RP is highly significant in the bw-direction (Fig. 5b) but not significant in the fw-direction (Fig. 5a).

4.4.3 Classes fw_fm and bw_fm

As mentioned above, listeners either prefer fundamental listening (fm-hearing) or spectral listening (sp-hearing). Thus, we further differentiated the data and separated fm-responses from sp-responses and compared AP to RP in both groups. The top lines of Fig. 6 graph the response rates of absolute pitch possessors with preferred fundamental listening (AP_fm) versus listeners with relative pitch and preferred fundamental listening (RP_fm). The bottom lines of Fig. 6 show the response rates of spectral listeners with absolute pitch (AP_sp) versus the response rates of spectral listeners with relative pitch (RP_sp). In the group fm, the ratings for fundamental hearing are much higher for AP compared to RP (highly significant) as test 5 demonstrates for the fw-direction ($p = 0.002$, Fig. 6a) and test 6 for the bw-direction ($p = 0.0006$, Fig. 6b). If the common highest partial is below 5000 Hz the statistical tests always prove a highly significant dominance of fundamental hearing in AP compared to RP. Above 5000 Hz the differences between AP and RP vanish for all both groups.

4.4.4 Classes fw_sp and bw_sp

In the group bw_sp, the ratings for spectral hearing is much higher (higher absolute values of the ratings) for RP compared to AP (Fig. 6b) as test 9 shows with high significance ($p = 0.007$), whereas test 8 shows no differences between AP and RP for group fw_sp (Fig. 6a). Thus, we conclude that possessors of AP much more willingly follow the rising fundamental than hearers with RP. If the common highest partial is below 5000 Hz the statistical tests always prove a highly significant dominance of fundamental hearing in AP compared to RP. Above 5000 Hz the differences between AP and RP vanish for all both groups.

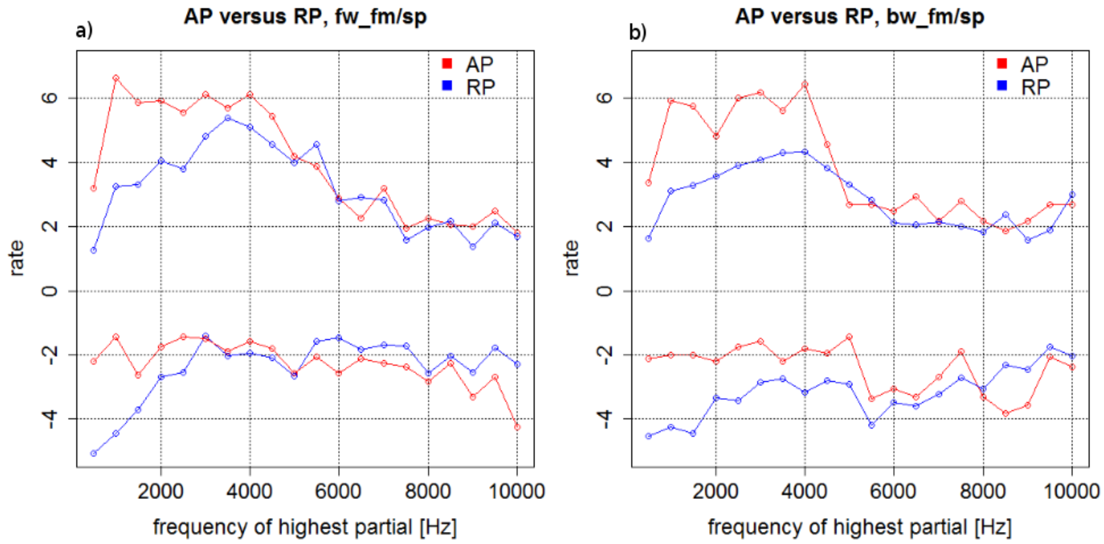

Figure 6: Data splited into fm-hearing and sp-hearing for both directions. In the frequency range from 1000 HZ to 5000 Hz the degree of fm-hearing is much higher for AP than for RP as well in the fw-direction (Fig. 6a) as in the bw-direction (Fig. 6b). In the range below about 3000 Hz sp-hearing dominates for RP in the fw-direction (higher absolute values of rates). In the range below even 7000 Hz sp-hearing dominates for RP in the bw-direction. In contrast, AP gains high rates of fundamental hearing even for frequencies below 1000 Hz (both panels).

4.5 Directional sensitivity

Only in listeners with relative pitch do the response rates to the presentations in forward direction differ significantly (d-test 4: $p = 0.04$) or even highly significantly (d-test 2: $p = 0.004$ and test 6: $p = 0.001$) from the response rates to the presentations in backward direction (see Fig. 7b).

If the highest partials of both complex tones are in a frequency range between about 2000 Hz to 8000 Hz, listeners with RP have a significant higher preference for hearing a missing fundamental in the forward direction (ascending higher partials and a descending fundamental) compared to the backward direction. In contrast, the ratings of possessors of absolute pitch are not influenced by the direction of presentation (d-test 1: $p = 0.15$, d-test 3: $p = 0.44$, d-test 5: $p = 0.31$). Observe the scale of the ordinate of Fig. 7 to realize that the ratings of possessors of AP are overall higher than those of listeners with relative pitch in the frequency range up to 5000 Hz.

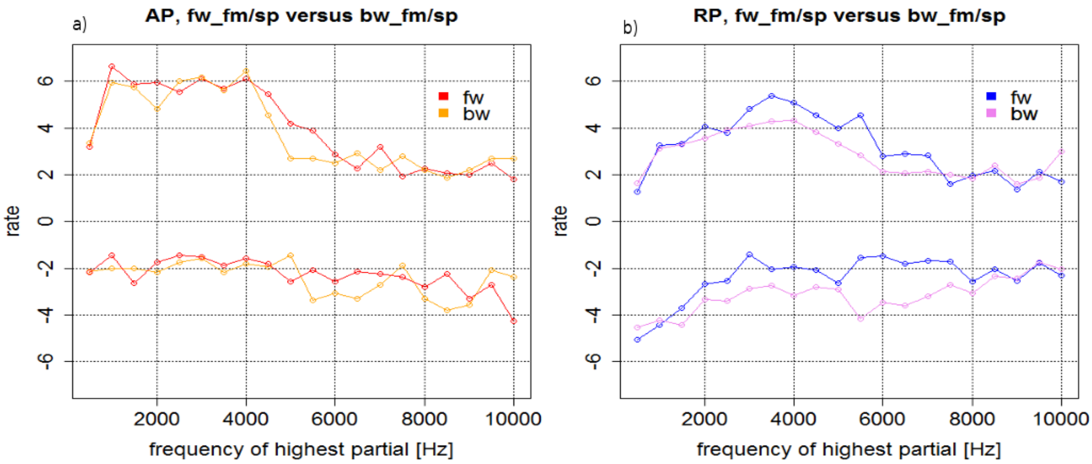

Figure 7: Differences in directional senitivity between AP and RP. In both panels, the upper two lines represent both directions of fm-hearing and the lower lines represent both directions of sp-hearing. The responses of AP (Fig. 7a) show no significant differences between presentations in forward versus backward direction for both hearing modes. In contrast, the responses of RP to presentations in forward direction (Fig. 7b) are significantly higher between 2000 Hz and 7000 Hz compared to presentations in backward direction.

5. Discussion

Except the note naming ability which just defines absolute pitch, no differences in hearing between AP and RP have been found up until now. Therefore, to our knowledge, the present study is the first to reveal an additional difference in hearing between AP and RP: **possessors of absolute pitch preferably extract the fundamental tone from harmonically related stimuli below 5000 Hz.** Note, that periodicity detection is mainly restricted to frequencies below 5000 Hz (Zwicker & Fastl, 1990, Langner 2015). This indicates a connection between

- the strength of periodicity detection,
- fm-hearing, and
- absolute pitch.

However, the perception of the fundamental tone is not at all the privilege of AP. Everybody without hearing disorder may perceive the fundamental pitch and in case of a missing fundamental each normal auditory system produces the residue pitch. This is due to the periodicity detection mechanism of the auditory system between the *nucleus cochlearis* of the brain stem and the *inferior colliculus* of the midbrain (Langner, 1981, 1983; Schreiner and Langner, 1988; Langner,1992; Ebeling, 2008. Langner, 2015).

Be it inherited, learned or imprinted, AP must have neuronal correlates and most likely it depends on a certain neuronal mechanism as a pre-condition for AP. Such a neuronal mechanism would also support arguments that AP is not a specific kind of memory. To bring to mind imagined pitch from long term memory or determining tones by comparison to a memorized reference would be a time consuming and perhaps uncertain process with a latency of at least one to two seconds (Külpe, 1912). In contrast, judgements from a “genuine absolute pitch” as Bachem (1954 / 1955) calls the real AP is “fast, decided, and certain”. In informal tests the authors found times of less then 800 ms from tone onset to the end of verbal responses by most possessors of AP. This finding is in line with an investigation by Miyazaki (1990). Thus, it is quite unlikely that long-term memory is decisive for AP (Bachem 1954).

The assumption of a neuronal process in the auditory system on the level of the above mentioned periodicity detection mechanism between brain stem and midbrain as a basis of AP could explain the determined and quick responses of possessors of AP to tone identification tasks.

The assumption of a pre-cortical mechanism for AP is supported by an investigation by Waymann (1990). In response to sinusoidal stimuli he found a faster P3 ERP in possessors of AP compared to musicians and non-musicians with RP. But as these fast P3 responses of AP possessors had remarkably lower amplitudes, he concluded that AP possessors "do not involve the higher brain pathways that evoke the P3 in the same way as they do for persons without AP ability".

The finding of our study that possessors of AP prefer fm-hearing has a correspondence in results of Keenam et al. (2001) and Schneider et al. (2005) who found leftward cortical asymmetries for possessors of AP and fm-listeners, respectively. Keenam had demonstrated an increased leftward asymmetry of the Planum temporale in AP musicians compared to non musicians. In comparison, in musicians without AP no asymmetry of the Planum temporale could be observed even despite early onset of musical training. He also observed that the AP musicians almost always started musical training before the age of seven. These observations support the hypothesis that imprinting in a critical early phase to acquire AP becomes only effective on the basis of a neuronal predisposition. Quite remarkably, Schneider found a pronounced leftward asymmetry of grey matter volume and P50m activity within the pitch-sensitive lateral Heschl's gyrus (which is a part of the Planum temporale) for fm-listeners.

The relation between AP and fm-listening is not straightforward. Some participants with AP showed a preference for spectral hearing (AP3) or were able to perform both. Some test hearers with AP demonstrated that they could voluntarily switch from sp-hearing to fm-hearing to determine pitch (see 4.4). Already Carl Stumpf (1890) knew that directed attention may help to focus either onto the overtones or the fundamental.

We speculate, that the hypothesized mechanism that allows for absolute pitch determination is based on intrinsic oscillations by chopper neurons in the cochlear nucleus (Bahmer and Langner, 2007) which might also lead to the preference for fundamental pitch. The coincidence neurons in the IC are organized in a laminar structure along two approximate orthogonal axis (Schreiner and Langner, G. 1988, 1997). On the one axis the neurons are tonotopically graded according to their best frequency (tonotopy). On the orthogonal axis they are successively organized according to their best period (periodotopy, Schreiner and Langner 1997, Langner 2015). A corresponding orthogonal representation of tonotopy and periodotopy can also be found in the auditory cortex (Langner et al., 1997; 2009). As fm-listening refers to the period of the stimulus, one may assume that in case of AP, the intrinsic oscillations with periods that are multiples of 0.4 ms couple to the periodotopy axis. As a result, fm-listening is enhanced with the consequence of an

increased leftward cortical asymmetry. Listeners with preferred sp-listening and AP may switch from sp-listening to fm-listening to determine the pitch absolutely. But it is also conceivable that the intrinsic oscillations with a period of 0.4 ms can also couple to the tonotopy axis in these relatively rare cases of our test.
The reason for the directional sensitivity of RP in contrast to AP is not clear to us. We speculate that this is an example of how musical context becomes important at an elementary level. It has been established that autistic people perceive auditory stimuli as well as or better than neurotypical listeners. On the other hand, they find it more difficult to recognize the musical context. Wenhart et al. (2019) have shown “both shared and distinct neural features between AP and autistic traits”. One may speculate that the observed lack of directional sensitivity of AP is a result of autistic traits in the possessors of AP.
Further investigations are necessary to examine the directional sensitivity of RP and to test our hypotheses.

6. Conclusion

From previous investigations and the present study, it becomes clear that there are three pairs of correspondences:

1. AP corresponds to a leftward asymmetry of the planum temporale,
2. fm-listening corresponds to a leftward asymmetry of the lateral Heschl's gyrus,
3. AP corresponds to an enhanced fm-listening.

As the detection of the fundamental frequency is performed by a neuronal periodicity analyses, the root of AP could be a certain property of this periodicity detection, which results in an enhanced leftward cortical representation of periodotopy.

7. Acknowledgment

The following lecturers, their institutes and their students supported the tests:
Prof. Matthias Franke †, HfMT Cologne / Germany
Prof. Hervé Laclau, HfMdK Frankfurt a. Main / Germany
Prof. Dr. Rötter, TU Dortmund, Institute of music and musicology, Germany
Prof. Markus Roth, Folkwang UdK Essen / Germany

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

8. References

Abraham, O., 1901. Das absolute Tonbewusstsein. Psychologisch-musikalische Studie. *Sammelbände der Internationalen Musikgesellschaft*, 3. Jahrg., H. 1, 1-86; https://archive.org/details/SammelbndeDerInternationalenMusikgesellschaft021900-01.
Bachem, A., 1954. Time Factors in Relative and Absolute Pitch Determination,

Absolute pitch, *J. Acoust . Soc. Am.* 26, 751-753; DOI:http://dx.doi.org/10.1121/1.1907411.

Bachem, A., 1955. Absolute pitch, *J. Acoust . Soc. Am.* 27, 1180-1185; DOI:http://dx.doi.org/10.1121/1.1908155.

Bahmer, A. and Langner, G. 2007. Simulation of oscillating neurons in the cochlear nucleus: A possible role for neural nets, onset cells, and synaptic delays. In: *Hearing–From Sensory Processing to Perception*. Berlin Heidelberg Springer: 155-164.

Bregmann, A., 1990. *Auditory Scene Analysis.* Cambridge: MIT.

Baharloo, S., Johnston, P.A., Service, S.K., Gitschier, J., and Freimer, N.B., 1998. Absolute Pitch: An Approach for Identification of Genetic and Nongenetic Components. *Am. J. Hum. Genet.* 62:224–231; DOI: http://dx.doi.org/10.1086/301704.

Brady, P.T., 1970. Fixed-Scale Mechanism of Absolute Pitch. *J. Acoust. Soc. Am.* 48, 883-887; Doi: http://dx.doi.org/10.1121/1.1912227

Cheveigné, A. de., 2000. The auditory system as a “seperation machine”. *CNRS-Ircam*, 1 place Igor Stravinsky, 75004, Paris, France; http://vanilla47.com/PDFs/Cognition/General/The%20auditory%20system %20as%20a%20separation%20machine.pdf.

Corliss, L.R., 1973. Remark on “Fixed-scale mechanism of absolute pitch” [P. T. *Brady*, J. Acoust. Soc. Am. 48, 883–887 (1970)]. *J. Acoust. Soc. Am.* 53, 1737-1739; Doi: http://dx.doi.org/10.1121/1.1913527.

Deutsch, D., Henthorm, T., Marvin, E., and Xu, H., 2006. Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *J. Acoust . Soc. Am.* 119, 719-722; DOI: http://dx.doi.org/10.1121/1.2151799.

Deutsch, D., Dooley, K., Henthorn, T., and Head, B., 2009. Absolute pitch among students in an American music conservatory: Association with tone language fluency. *J. Acoust. Soc. Am.* 125, 2398-2403; DOI: http://dx.doi.org/10.1121/1.3081389.

Ebeling, M., 2008. Neuronal periodicity detection as a basis for the perception of consonance: A mathematical model of tonal fusion. *J. Acoust . Soc. Am.* 124, 2320-2329. Doi: http://doi.org/10.1121/1.2968688

Fletcher, H., 1940 “Auditory patterns,” *Rev. Mod. Phys.*12, 47–65; DOI:http://dx.doi.org/10.1103/RevModPhys.12.47.

Hartmann, W.M., 1998. *Signal, Sound, and Sensation.* New York: Springer.

Hartmann, W.M., Goupell, M.J., 2006. Enhancing and unmasking the harmonics of a complex tone. *J. Acoust. Soc. Am.* 120, 2142-2157; DOI:http://dx.doi.org/10.1121/1.2228476.

Helmholtz, Hermann v., 1862/1913. *Die Lehre von den Tonempfindungen*. Reprint 1983 Hildesheim: Georg Olms

Hornbostel, E.M. v., 1926, Physiologie der Gehörserscheinungen. In: *Handbuch der normalen und pathologischen Physiologie*; 11. Receptionsorgane I. Berlin: Springer, 701-730.

Keenam, J.P., Thangaraj, V., Halpern, A.R,, and Schlaug, G., 2001. Absolute Pitch and Planum Temporale . *NeuroImage* 14, 1402-1408; Doi: http://dx.doi.org/10.1006/nimg.2001.0925.

Külpe, Oswald 1912. Über die moderne Psychologie des Denkens. *Internationale Monatsschrift für Wissenschaft Kunst und Technik* 6, 1912, 1069-1110.

Langner, G., 1981. Neuronal mechanisms for pitch analysis in the time domain. *Experimental Brain Research* 44: 450-454; Doi:10.1007/BF00238840.

Langner, G., 1983. Evidence for neuronal periodicity detection in the auditory system of the guinea fowl: implications for pitch analysis in the time domain. *Experimental Brain Research* 52: 333-355; Doi:10.1007/BF00238028.

Langner, G. and Schreiner, C.E., 1988. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *Journal of Neurophysiology* 60: 1799-1822; http://jn.physiology.org/content/60/6/1799.

Langner, G., 1992. Periodicity coding in the auditory system. *Hearing Research,* 60, 115-142; Doi: http://dx.doi.org/10.1016/0378-5955(92)90015-F.

Langner, G. et al., 1997. Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography. *Journal of Comparative Physiology* 181: 665–676, DOI: 10.1007/s003590050148

Langner, G., Dinse, H.R., and Godde, B., 2009. A map of periodicity orthogonal to frequency representation in the cat auditory cortex. *Frontiers in integrative neuroscience*, 3. Doi: http://dx.doi.org/10.3389/neuro.07.027.2009

Langner G., 2015. *The neural code of pitch and harmony.* University Press: Cambridge. Doi: https://doi.org/10.1017/CBO9781139050852

Miyazaki, K. 1990. The speed of Musical Pitch Identification by Absolute Pitch Possessors. *Music Perception* 8/2, 177-188; DOI: 10.2307/40285495.

Monroe, D., 2004. *Speaking tonal languages promotes perfect pitch.* http://www.scientificamerican.com/article/speaking-tonal-languages/ (20.09.2016).

Moore,B. C. J., Glasberg, B. R., Shailer, M. J., 1984. Frequency and intensity difference limens for harmonics within complex tones. *J. Acoust . Soc. Am.*, 75, 550-561, DOI:http://dx.doi.org/10.1121/1.390527.

Moore,B. C. J., Peters, R. W., Glasberg, B. R., 1985. Thresholds for the detection of inharmonicity in complex tones. *J. Acoust . Soc. Am.*, 77, 1861-1867; DOI:http://dx.doi.org/10.1121/1.391937.

Muzzulini, D. 2006. Genealogie der Klangfarbe. Bern: Peter Lang. http://www.muzzulini.ch/publications/dm_2004_GdK_2004.pdf.

Patterson, R. D., 1976. Auditory filter shapes derived with noise stimuli. J. Acoust. Soc. Am. **59**, 640
DOI: http://dx.doi.org/10.1121/1.380914

Patterson, R. D., 1995. The auditory filterbank: Fletcher's functional model of hearing. *J. Acoust. Soc. Am.* 97, 3378,
DOI: http://dx.doi.org/10.1121/1.412622
Plomp, R., 1964. The Ear as a Frequency Analyzer. I. *J. Acoust . Soc. Am.* **36**, 1628-1636.
DOI: http://dx.doi.org/10.1121/1.1919256.
Plomp, R. and Mimpen, A,M., 1968. The Ear as a Frequency Analyzer II. *J. Acoust . Soc. Am.* **43**, 764-767;
DOI:http://dx.doi.org/10.1121/1.1910894.
Profita, J., Bidder, T. G., Optiz, J. M. and Reynolds, J. F., 1988, Perfect pitch. *Am. J. Med. Genet.*, 29: 763–771.
Doi: http://10.1002/ajmg.1320290405
Sachs, L. and Hedderich, J., 2006. *Angewandte Statistik. Methodensammlung mit R*. Berlin / Heidelberg: Springer.
Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H.J., Dosch, H.G., Bleeck, S., Stippich, Chr., Rupp, A., 2005. Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference, *Nature Neuroscience* **8**, 1241-1247;
Doi: http://dx.doi.org/10.1038/nn1530.
Schouten, J.F., Ritsma, R.J., Lopes Cardoz, B., 1962. Pitch of the residue. *J. Acoust. Soc. Am.* 34, 1418-1424.
DOI:http://dx.doi.org/10.1121/1.1918360.
Schreiner, C.E. and Langner, G., 1988**.** Coding of temporal patterns in the central auditory nervous system. In: *Auditory Function* (Edelmann GM, Gall WE, Cowan WM eds), pp 337-361. New York: J. Wiley & Sons, Inc.
Schreiner, C.E., and Langner, G., 1997. Laminar fine structure of frequency organization in auditory midbrain. *Nature* 388: 383-386.
Stumpf, C. 1883/1890. Tonpsychologie. Leipzig: Hirtzel.
Ward, W.D., Burns, E. M., 1982. Absolute pitch. In: Deutsch, D. *The psychology of music.* Orlando: Academic Press, 431-451.
Zwicker, E., 1961. Subdivision of the auditory frequency range into critical bands (Frequenzgruppen). *J. Acoust. Soc. Am.* **33**, 248
Doi:10.1121/1.1908630
Zwicker, E., Fastl, H., 1990. *Psychoacoustics – Facts and Models.* Berlin*:* Springer