In E. Hickmann & R. Eichmann (Eds.) (2004). Orient-Archäologie, Vol. 15
(Studien zur Musikarchäologie, Vol. 4) (pp. 11-28). Rahden: Verlag Marie Leidorf

From *T*oddler's *H*eadturn to *T*schaikovsky and *H*aydn

A Review of Studies on Preferences for so-called Natural Musical Intervals

Reinhard Kopiez

ZUSAMMENFASSUNG

In den letzten Jahren richtete sich die Aufmerksamkeit von Musikpsychologen auf die Frage nach kulturunabhängigen Grundlagen des Musikverstehens, und die Suche nach musikalischen Wahrnehmungsuniversalien gewann an Popularität. Als Folge dieser Entwicklung wurde in einer immer noch zunehmenden Zahl von Studien die Frage untersucht, ob die Existenz eines voraussetzungslosen Musikverstehens angenommen werden kann. Im allgemeinen basieren diese Studien auf Experimenten mit Säuglingen im Alter zwischen zwei und zehn Monaten. Die Auswahl dieser Versuchspersonen basiert auf der idealisierten Annahme, daß diese Individuen unbeeinflußt von kulturspezifischen Hörerfahrungen sind. Wegen der fehlenden Möglichkeit, Reaktionen dieser Versuchspersonen durch verbale Äußerungen zu erfassen, werden Verhaltensreaktionen wie z. B. die Anzahl von Kopfdrehungen hin zu einer Klangquelle oder die Saugfrequenz an einem Schnuller als Indikatoren für das subjektive Gefallen verwendet. Ein Ergebnis dieser Studien war, daß sensorische Basisprozesse wie die spektrale Filterfunktion der *Cochlea* oder die kritische Bandbreite tatsächlich eine wichtige Rolle bei der Erklärung von Musikpräferenzen für konsonante gegenüber dissonanten Klängen spielen. Diese Ergebnisse betonen die Bedeutung sogenannter Bottom-up-Prozesse für die menschliche Wahrnehmung. Allerdings bleibt die Gewichtung dieser Basisprozesse im Verhältnis zu sogenannten Top-down-Prozessen, wie sie etwa typisch für die erwartungsgesteuerte Informationsverarbeitung sind, ungeklärt. Für die ästhetische Erfahrung sind Top-down-Prozesse jedenfalls genauso wichtig oder noch wichtiger als unbewußte Bottom-up-Prozesse. In diesem Beitrag wird deshalb argumentiert, daß die häufig vereinfachte Schlußfolgerung von Basisprozessen auf komplexe Wahrnehmungs- und ästhetische Erlebnisprozesse unzulässig ist. Es wird ein Überblick über die verwendeten Methoden, die in Studien zu sogenannten natürlichen Musikintervallen verwendet wurden, gegeben und es werden Vorschläge für eine stärkere Berücksichtigung kulturspezifischer Faktoren bei der Untersuchung der Musikwahrnehmung gemacht.

1 BACKGROUND AND RESEARCH QUESTION

We have recently seen an increasing number of studies in music perception which tackle the question of whether music can be perceived and enjoyed without prerequisites. The guideline for this kind of psycho-musicological research is the assumption that music could be a kind of 'universal language' which is understood – at least in some of its features – independently of the listeners' cultural background. As Kopiez could show in a review study[1], the motif of a *lingua universalis* can be traced back to biblical origin and is not only limited to music as a particular 'language'. The result of this first review study of musical universals was that there is little convincing evidence for the existence of musically relevant perceptual universals. The present study continues this discussion but with a different viewpoint. My approach is mainly guided through two opposing perspectives: after the first introductory section of this study I will discuss results from so-called bottom-up studies on auditory perception and their interpretation as indicators for 'natural musical intervals'. From this perspective, human perception is

---

[1] Kopiez in press.

understood as being data-driven and determined by general psychophysiological processes, such as the mechanics of the basilar membrane, which should be independent from learning processes or listening experience. In the third section I will change perspective to a top-down approach, whereby human perception is understood as being knowledge and expectation-driven. For both approaches methods and findings will be discussed, and in the fourth and final section I will draw conclusions about the relevance of bottom-up processes. Furthermore, the last section places our findings within the larger framework of cultural psychology and anthropology. It will be argued that aesthetic behaviour plays only a minor role in the explanation of cultural differences. Additionally, the idea of perceptual universals within cultural anthropology is obsolete. Instead of searching for facts that emphasize the commonality of different cultures, it will be demonstrated that modern cultural anthropology promotes the understanding of cultural diversity.

2 BOTTOM-UP PROCESSES

In this paragraph I will give an overview of basic psychoacoustic processes which are of central importance for human auditory perception. The theory of critical bands and its implications for the sensation of dissonance will be discussed as well as the results from cross-species comparisons of auditory perception. The main focus in this section is a review of studies on the perception of 'natural musical intervals' using babies as test subjects. This means that we will search for elementary musical events that are assumed to play an important role within a data-driven approach. Assumptions, results and limitations of these studies will be discussed. Finally, it is argued for a more cautious interpretation of findings from baby-studies in the light of a not yet sufficiently evolved explanation of music sensation.

2.1 SENSORY DISSONANCE AND THE ROLE OF CRITICAL BANDWIDTH

The most fundamental concept in the understanding of human auditory perception is the concept of 'critical bands'. This was first proposed by Fletcher[2], who assumed that the part of a noise that can mask a test tone is that part of its spectrum lying near the tone. Two test tones with a frequency difference smaller than the critical bandwidth can be masked by a narrow band noise. If the difference is larger than the critical bandwidth no masking effect occurs. In other words: subjective responses abruptly change at the critical band boundary. A second consequence of the critical band concept is that two frequencies falling within a critical band are more difficult to discriminate compared to two signals with a frequency distance larger than the critical bandwidth. This means that the basilar membrane is organized as a kind of filter bank comprising 24 regions, so-called critical bands. Below 500 Hz the bandwidth of these regions is 100 Hz and above 500 Hz it is about 20 % of the center frequency, which corresponds roughly to the musical interval of a minor third[3]. Zwicker et al.[4] and Greenwood[5] were the first authors to propose an algorithm for the calculation of critical bandwidth, and Zwicker suggested the term 'Bark' as the unit for bandwidth. A third consequence of the basilar membrane's filter function is that two frequencies within the range of the critical bandwidth can be perceived only with difficulty as two distinct pitches. If frequencies differ only slightly, the sound is characterized by slow beats, and with increasing frequency difference towards the bandwidth boundaries, the sensation of beats changes to the sensation of roughness[6].

In the next paragraph, as an illustration of its mechanic, the basilar membrane's behaviour is demonstrated with two signals of differing frequency distance. Modern computer simulations allow us to compute a simplified simulation of the membrane's behaviour according to the input stimulus. Figure 1b demonstrates the final state of the development of a so-called wandering wave on the basilar membrane after the first 20 milliseconds of a 1 and a 1.5 kHz sinusoid. Simulation is based on the simplifying assumption of a linear behaviour of the basilar membrane according to the model proposed by Diependaal[7] and calculated by the software 'Diep1d' by Brass[8].

As can be seen in Figure 1b, the two sinusoids of 1 and 1.5 kHz produce excitations on the basilar membrane with clearly separated maxima. These frequencies correspond with the musical interval of a perfect fifth (702 cents). Figure 1a demonstrates the basilar membrane's excitation for two frequencies of 1 and 1.1 kHz, which corresponds to a musical interval between a minor and major second (165 cents). Only one maximum of excitation can be observed with a nearly fused side maximum on the right side of the main maximum. The

---

[2] Fletcher 1938.

[3] Zwicker/Fastl 1999.

[4] Zwicker et al. 1957.

[5] Greenwood 1961b; Greenwood 1961a.

[6] Hall 1991, 382.

[7] Diependaal et al. 1987.

[8] Brass 1995.

prediction of the critical bandwidth theory would be that this basilar membrane excitation would correspond with the sensation of a very rough sound with hardly separable pitches. The degree of roughness determines the subjectively perceived degree of a harmonic interval's consonance, and in intervals with a frequency difference of less than 200 Hz, consonance rating decreases rapidly[9]. Physiologically, this sensation is based on the basilar membrane's frequency resolution that is limited by the number of hair cells per millimeter. Frequencies of small distance excite haircells in direct proximity. Due to the fundamental physiological behaviour of the basilar membrane in all human beings, this mechanism should function independently of different listening experiences and cultural background.

A short glance at two standard handbooks of music psychology and music education show how influential this theory has been for the development of sensualistic aesthetics. For example, in his chapter on 'Music and the auditory system', Weinberger reviews several neurophysiological studies and concludes that the total number of auditory nerve fibers that show beating patterns have a high correlation with the perceived dissonance of musical intervals and that "at least one aspect of consonance/dissonance is determined at the most peripheral level of the central auditory system"[10]. As a second example I would like to quote from the chapter on 'Cognitive constraints on music listening' by Thompson and Schellenberg[11]. The authors argue that "sensitivity to sensory consonance and dissonance is thought to be independent of knowledge and enculturation"[12].

The idea of psychophysiological aesthetics can be traced back to the writings of von Helmholtz and his theory of beats between adjacent partials of fundamental tones which cause the sensation of dissonance. An application of this theory to compositional analysis can be found in the study by Plomp and Levelt[13], who explain the distribution of consonant and dissonant musical intervals in the music of J. S. Bach and A. Dvorák by the laws of critical bandwidth, in particular consideration of roughness of sounds. If the authors are right in their unprecedented assumption of a determination of composition through psychophysics, the question as to why we can observe such a great increase in dissonance in music as a general tendency over the centuries from Bach to Mahler and Strauss remains unanswered. If aesthetics is ruled by psychophysics, the development of composition should stop when a balance between consonance and dissonance is reached. At this point in history, music should then remain in a balanced state. Musical reality is different, and as we can already see in von Helmholtz' theory of consonance, it is the result of (a) the degree of roughness produced by fast beats of partials and (b) the relationship between pitches. To date, this two-component theory of sensory consonance is widely accepted in psychoacoustics[14]. However, it is interesting to note that von Helmholtz was far-sighted enough to assume that the relationship between pitches is culturally dependent and thus varies over time and region.

Strong arguments for a psychophysiological theory of consonance/dissonance come from neurobiological research. For example, Tramo et al. [15] analysed the neurophysiology of consonance perception in Western music. The authors assume that auditory information can only be processed in the time domain and that harmonic intervals are characterized by regularities in the temporal fine structure of neural coding. Indeed, neural responses of the auditory nerve fibers reflect the fine structure of the signal in the time domain. Dissonant harmonic intervals such as minor seconds or tritones show irregular interspike intervals throughout the auditory nerve. Consonant intervals (fourths and fifths) show highly structured temporal information: "distinctive acoustic features of consonant and dissonant intervals are translated into distinctive patterns of neural activity"[16]. However, as we will see at the end of this first section, Tramo et al. also emphasize that bottom-up processes are no more than an initial and insufficient attempt at explaining the sensation of pleasant sounds.

2.1.1 Results from Cross-species Comparisons: What Does Human Perception Have in Common with Cats, Bats and Monkeys?

The described characteristics of auditory processing are not unique to humans. As the comparison of human auditory processing with that of other highly developed vertebrates shows, human auditory processing has far more features in common with other species than expected. For example, neurophysiological studies of auditory perception in cats revealed that these animals are able to reconstruct missing fundamentals (the so-called virtual pitch) from incomplete spectra[17]. Even the extremely small auditory cortex in bats can reconstruct missing fundamentals in the ultrasonic

---

[9] Plomp/Levelt 1965.

[10] Weinberger 1999, 67–68

[11] Thompson/Schellenberg 2002, 463–464.

[12] Thompson/Schellenberg 2002, 463.

[13] Plomp/Levelt 1965.

[14] Terhardt 1998, 403.

[15] Tramo et al. 2001.

[16] Tramo et al. 2001.

[17] Heffner/Whitfield 1976; Whitfield 1980.

range[18]. In bats the classification of complex tones and echoes seems to be ruled by the same template for spectral analysis as in cats and humans[19]. As Ehret and Merzenich[20] could show in a neurological study on the reaction of single neurons of the inferior *colliculus* of the cat, the auditory system of this mammal is also based on the perceptual organization into critical bands. However, the number of critical bands in the cat is limited to five (the number of critical bands in humans is 24). As Pickles could show in a comparison between physiological and behavioural data, "the bandwidth of the cat's psychophysical filter is substantially greater than the bandwidth of single fibers of the auditory nerve, and that psychophysical tuning curves are an artificially narrow measure of frequency selectivity"[21]. Cats seem to have a larger mean critical bandwidth (compared to humans) of 410 Hz at 1 kHz and of 690 Hz at 2 kHz. These findings can be taken as evidence of the fact that incomplete spectral information of a complex tone "is sufficient to account for fundamental pitch sensations in the mammalian auditory system"[22].

Does this mean that animals are also capable of discriminating between musically meaningful stimuli and can show preferences for sensory consonance over dissonance? Two studies provide insights: the first by Borchgrevink[23] investigated whether rats show a preference for consonant over dissonant musical chords. Rats were trained to push a switch and received food. Switches were coupled with (a) consonant, (b) dissonant and (c) no sound (control condition). Frequency of pedal pushes was measured as the dependent variable. Observation of behaviour over a period of three weeks revealed a preference for consonant over dissonant chords, and for both over the no noise condition. Unfortunately, no information is given concerning the acoustical structure of the chords and additionally, animals showed great interindividual differences in reaction. Thus, results are difficult to interpret. The second study by Fishman et al.[24] used neuro-physiological methods to test von Helmholtz' roughness theory of sensory dissonance. Authors applied depth electrodes to the auditory cortex (field A 1) of macaque monkeys. Animals listened to harmonically complex consonant (fourth, fifth, octave) and dissonant (major/minor second, tritone, minor seventh) intervals. Neural activity in Heschl's gyrus clearly showed that dissonant stimuli evoke oscillatory activity phase-locked to the difference frequencies, whereas consonant intervals evoked little or no oscillatory activity. Similar results were found by use of the same method applied to human subjects. Results are interpreted as a neurophysiological confirmation of von Helmholtz' roughness theory of dissonance.

2.1.2 Critique of the Concept of Sensory Dissonance

Despite the seemingly compelling arguments for the important role of roughness as the central sensory component in the discrimination between consonant and dissonant sounds, we should bear in mind that there is serious critique against an exclusive use of this concept. For example, as Terhardt argues, the absence of roughness only slightly correlates to the perception of musical intervals, as there are several musical events which cause the impression of dissonance without the presence of roughness (e. g. in successive intervals or in single sounds with inharmonic spectra)[25]. Moreover, the concept of 'psychoacoustic consonance' is not a sufficient basis for the 'valence' impression of intervals. Alternatively, he proposes his theory of 'virtual pitch' as a basic concept of consonance perception[26].

2.2 RESULTS FROM STUDIES ON PREDISPOSITIONS FOR 'NATURAL MUSICAL INTERVALS' (SO-CALLED BABY-STUDIES)

Can the results from animal experiments be transferred to humans? Of course, for ethical reasons, invasive neurophysiological experiments such as the measurement of auditory single neuron activity are not carried out. Thus, in the last decade, methods of behavioural measurement as an indicator of underlying neural processes have been developed in psychoacoustic research. Very young infants seem to be the ideal subjects for these kinds of studies in an investigation of the influence of culture and nurture on human auditory perception. It is assumed that cultural influences are only of low impact in these very young children and that their perception is unshaped by auditory experiences. This paradigm can be described as the so-called *tabula rasa* paradigm, and a review of these studies will follow. It will be demonstrated that this assumption is not a correct description of human predispositions.

An early study on the responses of 2- to 8-month-olds by Standley and Madsen[27] measured the responses to music, as well as to the mother's

---

[18] Preisler/Schmidt 1995.
[19] Krumbholz/Schmidt 1999; Krumbholz/Schmidt 2001.
[20] Ehret/Merzenich 1988.
[21] Pickles 1979.
[22] Preisler/Schmidt 1995, 47.
[23] Borchgrevink 1975.
[24] Fishman et al. 2001.
[25] Terhardt 1974.
[26] See also Terhardt et al. 1982.
[27] Standley/Madsen 1990.

and other female voices by a switch attached to the infant's foot. A cassette player was activated by this switch and the playing time for each of the three stimuli was measured. As a result, no significant overall differences were found between the three stimuli. However, an interaction between age and preference revealed a preference for the mother's voice in younger babies and for the mother's and other female voices in older babies. In an extensive review of the literature, the authors show that evidence of the usage of consonant versus dissonant stimuli in preference experiments cannot be found prior to the 1990s.

Following the above-mentioned predecessory study, the study by Lynch et al.[28] marks the beginning of a new research paradigm. To the best of my knowledge, it was the first instance in which 6-month-olds were used in an experimental setting examining the discrimination abilities for modified native and non-native scales. The authors investigated the influence of acculturation on auditory perception in Western listeners. Short melodies comprising seven notes were constructed using the scale degrees of major, minor, and the Javanese pelog scale. Two versions of these melodies were presented: (a) the original version and (b) a version in which the fifth note was mistuned by raising its frequency by 50 cents. The number of correct head turns in mistuned melodies was counted as the dependent variable. As a result, infants' performance did not differ as a function of the musical scale, but was better than chance. In contrast, the adults' performance was better on the native than non-native scales. Results are interpreted in terms of an "equipotentiality for the perception of scales from a variety of cultures"[29].

Continuing these studies, Schellenberg and Trehub[30] investigated whether infants are capable of detecting subtle changes in a sequence of simultaneous intervals. The stimulus consisted of a repeated harmonic interval (e. g. a perfect fifth) followed by the same but slightly compressed interval (diminished by 25 cents). Stimuli were presented via loudspeaker and the number of head turns to the sound source at the instances of interval changes was counted as 'hits'. As a result, infants succeeded in detecting changes to the perfect fifth and fourth interval but not to the tritone. These findings are interpreted as a predisposed preference for simple frequency ratios over more complex ones. The authors assume a biological basis for this predisposition based on the theory of sensory consonance. A further review of literature on auditory predispositions for 'natural preferences' is given by Trehub[31] and Trehub et al.[32].

In a follow-up study by Zentner and Kagan[33], the reductionistic approach by Schellenberg and Trehub was modified in favour of a stimulus with higher validity. Authors presented short melodies in two versions: (a) in parallel thirds and (b) in parallel seconds (see Fig. 2). 4-month-olds listened to the melodies and fixation times of a sound source (loudspeaker) and motor activity was measured accordingly. In summary, infants looked longer at the loudspeaker and showed less motor activity when hearing the consonant versions of the melody. Results are interpreted by the authors as "an innate bias favoring consonance over dissonance"[34].

The most recent study on so-called natural musical intervals was published by Trainor et al.[35]. The authors used the conditioning paradigm as used in previous studies, but used 2- and 4-month-olds as subjects. Results are similar to previous studies, however show a bias: firstly, the drop-out rate of babies who could not be considered due to unreliable behaviour comprised about one third of the subjects, which is interpreted by the authors as a kind of 'denoising of data'; secondly, a sequence effect could be observed, namely that mean looking times were much higher when dissonant trials followed consonant trials compared to consonant trials following dissonant trials. If sensory processes only are responsible for aesthetic appreciation, the sequence of stimuli should have no influence. This is, seemingly, not the case.

2.2.1 Discussion of Studies on 'Natural Musical Intervals'

How can results from studies on 'natural' preferences for consonance be interpreted and what are the critical methodological points in these studies? The following paragraphs point out three noteworthy aspects: namely the question of perception vs. preference, the paradigm of conditioning vs. habituation, and the question of validity.

Perception vs. Preference

As could be shown in the study by Lynch et al.[36], results are only interpreted in terms of infants' ability to discriminate rather than in terms of preferences. However, in the study by Schellenberg and Trehub[37], results from a discrimination task

---

[28] Lynch et al. 1990.

[29] Lynch et al. 1990, 275.

[30] Schellenberg/Trehub 1996.

[31] Trehub 2000.

[32] Trehub et al. 1997.

[33] Zentner/Kagan 1998.

[34] Zentner/Kagan 1998, 491.

[35] Trainor et al. 2002.

[36] Lynch et al. 1990.

[37] Schellenberg/Trehub 1996.

using the same head turn paradigm are interpreted in terms of 'preference' and not as an indicator of discriminability. Secondly, the finding of a non-significant difference in the perception of mistuned, non-Western pelog scales in the study by Lynch et al. and of a significant lower detection of compressed dissonant intervals in the study by Schellenberg and Trehub remains unresolved. The contradition tackles the question of so-called open earedness[38], and equipotentiality (an openness to any auditory experience) vs. predisposition for consonance.

Conditioning vs. Habituation

The above-mentioned authors used a conditioning paradigm. This means that infants were first trained to discriminate between two version of a sound or melody by use of reinforcement through a light-animated toy. However, this paradigm elicits additional 'head-turns' from the subject's spontaneous behaviour, rendering invalid results. If the assumption of a natural preference for consonance is correct, this result should also be found by use of a habituation paradigm. This means that fixation time and number of head-turns towards the sound source as indicators of appreciation should also appear in the case of highly repeated stimuli. In current infant-directed research, the use of a habituation paradigm is a serious and well-established methodological alternative, and results should be independent from the paradigm used. Additionally, physiological indicators for appreciation, such as facial muscle activity, should be considered.

Validity

Until now, there has been only one study that uses entire compositions (instead of isolated intervals) in original and dissonant versions as a valid stimulus. Trainor and Heinmiller[39] produced consonant and dissonant versions of a Mozart Minuet. However, the use of only one composition does not allow for conclusions to be drawn about the effect of dissonance within music of different styles and different degrees of harmonic complexity.

In summary, this short review of research on the psychophysiological and neurobiological correlates of auditory processing convincingly shows common features in elementary auditory pre-processing across species. However, the role of cognitive processes that influence the evaluation of basic signal analysis remains an important and not yet answered question. For example, results from current neurobiological research in the foundations of auditory perception emphasize the importance of cognitive processes. As Tramo et al. argue, the degree of roughness of a harmonic interval is not a sufficient explanation for the consonance rating. The authors' view is based on experimental results from a young stroke patient with bilateral auditory cortex lisions, who rated major triads as mistuned despite the absence of roughness in sound. This leads us to a two-component theory of consonance which is based on elementary neurophysiological mechanisms of auditory processing on the level of auditory nerve activity, as well as on higher level processes (so-called top-down processes). In the authors' own words the current state of research into the complex processes of auditory perception can be summarized as follows: "in view of the likelihood that cognitive representations of pitch hierarchies influence harmony perception in the vertical dimension, we urge that the terms **sensory consonance** and **sensory dissonance** be reconsidered"[40].

3TOP-DOWN PROCESSES

The next paragraphs will discuss the question of how important the basic sensory processes described in the previous first section are for the evaluation and appreciation of sound events, if we consider that bottom-up processes are modulated through higher cognitive processes. For example, processes such as evaluation and attention form a group of perceptual variables, which can be described as 'top-down'. This means that perception is predominantly guided by higher cortical processes and is not data-driven, as assumed in the bottom-up approach. Thus, the following paragraphs will explore the assumption that perception can only be adequately understood if bottom-up as well as top-down factors are considered.

After a short exemplification of the importance of top-down processes using the major-minor chord problem, this section will continue with a discussion of recent neurophysiological findings related to the existence of so-called evaluative networks in the brain. The following comparison of visual and auditory perception and a simulation of the environmental picture produced by the eyes' optical system will clearly demonstrate the importance of higher information processing. This section ends with the discussion of an experiment on the role of attention for the perception of sensory dissonance and an explorative psychoacoustic analysis of the first bars of Wagner's *Tristan Prelude*.

---

[38] For the term 'open earedness' see Hargreaves 1982; LeBlanc 1991; LeBlanc et al. 1996.

[39] Trainor/Heinmiller 1998.

[40] Tramo et al. 2001, 113.

3.1 HELMHOLTZ AND THE PROBLEM OF THE DISSONANT MINOR CHORD

When Hermann von Helmholtz published his famous book *On the sensations of tones* in 1863, he could not have foreseen that his conclusions about the acoustical properties of chords would cause the most disagreement among music theorists. Helmholtz compared the combination tones caused by major and minor chords in root and inverted position[41]. These 'extra' notes are produced by the non-linear transformation characteristic of the middle ear. If we only consider his comparison of chords in root position, we can see that two difference tones (one and two octaves below the fundamental) are clearly perceptible in major chords. However, in minor chords, these first order difference tones (e. g. in C minor: A*b* – E*b* – C from lowest to highest note) do not belong to the harmony. Thus, Helmholtz concluded that these tones "produce a sensible increase of roughness in comparison with the effect of major chords, for all cases where just intonation is employed"[42]. Alexander Ellis, the translator and editor of the book, added in a footnote that this effect also occurs on tempered instruments, despite the omnipresent beats in equal temperament. With regard to the dualism of major and minor, the finding of lower sensory consonance in minor chords was not acceptable for music theorists such as Hugo Riemann. To summarize, from a modern perspective we might conclude that both authors were right: Helmholtz' calculations of roughness are correct and an important finding from a bottom-up perspective. Riemann's claim of a balance of modes which is independent from sensory features is valid from a top-down perspective, where the evaluation of a chord is determined by higher cognitive processes such as its harmonic function or its appearance within a particular context.

3.2 RESULTS FROM NEUROPHYSIOLOGICAL RESEARCH: THE ROLE OF EVALUATIVE NETWORKS

Support for the assumption of an important role of top-down processes is also given by current neurophysiological research. This research produces convincing arguments for at least three networks that modulate perception: (a) a network for the violation of syntax rules, (b) a network for changes in timbre and (c) a network for the evaluation of sounds.

In a sequence of studies using modern methods of brain imaging, Koelsch[43] revealed the existence of brain networks that are specialized in the detection of syntax violation in language and music. For example, when major triads with a slightly detuned major third were presented in a sequence of tuned triads, the electrical brain response (measured by the so-called mismatch negativity) in a non-attentive listening condition showed a distinct reaction in professional musicians, but not in non-musicians. This means that identical sensory input can elicit different brain reactions in different subjects on the pre-attentive level, depending on the listeners' musical expertise. In other words: perception can be modulated on the sensory level by top-down processes such as listening experience and expectation.

A second study by Koelsch used unexpected (transposed) chords in a piano sonata to elicit event-related brain potentials[44]. Clear EEG mismatch potentials show that networks specialized in the auto-detection of musical syntax violation are always activated even in inexpressive musical stimuli. Mismatch negativity is also elicited in the case of unexpected harmonic changes such as the Neapolitan chord at the end of a cadence. This is an interesting stimulus because this chord is characterized by the same degree of sensory dissonance as any other triad of the cadence. However, the elicited reaction of the syntax network indicates that nearly identical sensory features are modulated by harmonic expectation and knowledge. This behaviour of the brain, namely the highly automized search for rules and rule violations, seems to be a general charateristic of the human brain. It is less domain-specific than previously assumed, and violations of musical syntax elicit activity in the same areas known for syntax violation in language (e. g. the Broca area)[45]. Studies also show evidence of the existence of an additional network specialized in the detection of timbre differences: when subjects heard the last chord of a sequence with a different timbre mismatch, negativity was also elicited in non-musicians[46].

The third proposed network in music listening is the so-called evaluative network. It is assumed that this network is specialized in decisions made about 'pleasantness' vs. 'unpleasantness' of a sound. In a recent study by Koelsch et al.[47], the authors found that contrasting consonant and dissonant versions of a composition elicited activation in the parietal operculum. This area seems to

---

[41] Helmholtz 1954, 215.

[42] Helmholtz 1954, 216.

[43] Koelsch et al. 1999.

[44] Koelsch/Mulder 2002.

[45] Koelsch et al. 2001.

[46] Koelsch 2002.

[47] Koelsch et al. in preparation.

be relevant for the evaluation of a sound's 'valence' dimension (pleasantness).

3.3 THE CONSTRUCTION OF REALITY IN VISUAL PERCEPTION

It is difficult to demonstrate the importance of higher cognitive processes in auditory perception. Of course, a simulation of the filter function of the basilar membrane would be possible with modern digital sound processing. However, I prefer to show the validity of the post-processing of sensory information with an example using visual perception.

As Itti[48] could show in a computer simulation of human visual perception, what we perceive as a clear image of a surrounding does not correspond with the optical features of the human eye. As Figure 3a shows, the eye is not a camera, even one of poor quality: the only region in focus is centered around the spot of sharpest vision (*fovea centralis*) in the middle of the picture. The black spot is caused by the exit of the optic nerve from the retina where no optic receptors exist. Despite these obvious severe lacks in projective qualities of the eye, our subjectively perceived impression of the same scene corresponds to Figure 3b, which is produced by a camera. We should note that this impression of a picture in focus is based on cortical compensation mechanisms, namely the brain's successful processing of incomplete sensory information. In other words: visual and auditory perception is not a copy of the real world but the result of reconstruction.

3.4 THE ROLE OF ATTENTION IN AUDITORY PERCEPTION: ATTENTIVE AND NON-ATTENTIVE LISTENING

As could be demonstrated in the previous paragraphs, sensory processes in auditory perception are modulated by higher cognitive factors, and attention plays an important role. If perception were only data-driven (bottom-up), the sensation of pleasantness would be independent from the degree of attention allocated to the auditory event. However, two important experiments investigating the role of attention for the evaluation of auditory perception have been carried out recently: a study about real-time detection of 'wrong' notes by Janata et al.[49] and a study about non-attentive listening by Wolpert[50]. With the above-mentioned studies by Koelsch in mind, we can emphasize that auditory perception is directed by three different degrees of attention: pre-attentive, attentive, and non-attentive listening. The next paragraph will summarize some important findings about the relationship between particular degrees of attention and the perception of auditory features.

In a most recent study, Janata et al.[51] used an 'attentive listening online probetone paradigm' to investigate, whether 'wrong' notes (meaning notes that are non-diatonic and do not belong to the main scale degrees of a key) can be detected more easily than diatonic ones. An ongoing melody was used and indications had to be indicated in real time. Depending on the tonal distance between test tones and tonal key, it was predicted that test tones would blend into some keys and 'pop out' in others. Subjects were instructed to use a footpedal to indicate notes which seemed to be 'wrong' or 'out-of-place'. As expected, indication of 'tonal fit' increased with tonal distance from current key. Results were interpreted as an effect of long-term memory for tonal hierarchy, learned through lifelong exposition to Western tonal music. This familiarity with a musical system establishes the basis for reliable judgements of tonal fit – even in non-expert listeners. Based on results from neuroimaging methods, the authors assume tonality-specific networks located in the rostromedial prefrontal cortex.

It remains contentious as to whether results could change with different degrees of attention. To study the question of 'what **do** people hear' (rather than 'what **can** people hear'), Wolpert used a non-directive listening paradigm. A jazz composition was prepared in three versions: (a) melody and accompaniment in the same key, (b) accompaniment transposed down a major second, (c) accompaniment transposed up a major second. Subjects listened to all three versions and were simply asked to note any differences. The results were as follows: 100 percent of the musicians perceived the clash of keys in the transposed versions, however, only 40 percent of the non-musicians recognized the difference between melody and accompaniment. Furthermore, 50 percent of the non-musicians indicated non-existent differences (e. g. in instrumentation or tempo).

To summarize, we might conclude that sensory information is a valid but insufficient explanation for the sensation of valence. As could be demonstrated by use of experimental methods, bottom-up processes are strongly influenced by higher cognitive processes such as attention or expectation. This finding is relevant for everyday listening of music: in an exploratory study by Sloboda et

---

[48] Itti 1998.

[49] Janata et al. 2003.

[50] Wolpert 2000.

[51] Janata et al. 2002.

al.[52], it could be revealed that listening to music as a main activity accounted for only 2 percent of all episodes measured. The majority of episodes was characterized by a non-attentive music listening attitude (listening while driving, at work etc.).

BOTTOM-UP VS. TOP-DOWN PROCESSES: WAGNER'S *TRISTAN PRELUDE* AS A TEST STIMULUS FOR THE PREDICTIONS OF SENSORY DISSONANCE THEORY

In the last paragraph of the third section I would like to test the predictions made by sensory aesthetics through the psychoacoustic analysis of Wagner's *Tristan Prelude* and the calculation of a 'sensory pleasantness curve'. The theoretical background is given by Zwicker and Fastl's[53] model of 'sensory pleasantness'. I intend to demonstrate a test of the predictions made by the sensory pleasantness theory which is based on the equation by Zwicker and Fastl (see Fig. 4). Strictly speaking, this equation must be applied cautiously to music, as the parameters are standardized to time-invariant and quasi-stationary sounds (such as noise). However, I believe it is an adequate approach to compare the predicted pleasantness curve (a) with the subjective experience of pleasantness (e. g. the experience of musical tension) and (b) with the musical structure of the score. This explorative method serves as an informal test of how far the assumptions of sensory aesthetics can be applied to musical stimuli. The basic parameters according to this equation are as follows[54]:

Sensory pleasantness: predicts whether a sound will be accepted as pleasant or unpleasant (measured as a relative value 'P/P_0'). For scaling reasons (original P values only cover an extremely narrow range), the original pleasantness curve has been modified by applying an amplification function (calculation of fourth root of original P value) resulting in the curve 'sensory pleasantness modified'. However, original and modified P values have been included in Figure 4.

Loudness: psychoacoustic loudness that considers frequency-dependent loudness sensitivity of the human ear and masking effects (measured in 'sone'). Doubling of sone value corresponds to subjectively perceived doubling of loudness.

Roughness: amplitude and frequency modulations with modulation frequencies > 15 Hz (measured in 'centi Asper')

Sharpness: proportionality between low- and high-frequency components (measured in 'acum')

Tonality: proportion of harmonic components to noise in a signal. Calculation is currently unclear and estimation is recommended by Zwicker and Fastl within the original equation. Due to this lack of precision in the calculation, the parameter tonality is not included in Figure 4.

The first measures of a recording of Wagner's *Tristan Prelude* served as acoustical stimulus[55]. Psychoacoustic analysis was conducted using the software dBSONIC[56] (SR = 44.1 kHz, 16 bit). The equation of the calculation of sensory pleasantness is displayed above Figure 4, whereby 'R' stands for roughness, 'S' for sharpness, 'T' for tonality (meaning the pitch to noise ratio of a signal; not included in this figure) and 'N' indicates loudness. If we trace the curve 'sensory pleasantness modified', it can clearly be seen that this curve mirrors the roughness curve. For our musical stimulus roughness seems to be the best predictor of pleasantness in the beginning of the *Tristan Prelude*. Other parameters, such as loudness, play no significant role. The dominance of roughness can be explained on the background of instruments determining the orchestral sound of the first 17 seconds: the timbre is governed by string instruments, and as Hall[57] explains, sound production on string instruments (the sliding of the bow across a string) is ruled by the 'stick-slip mechanism'. This means that in contrast to wind instruments, the timbre of string instruments is not characterized by formants but by permanent spectral fluctuations which are considered in the parameter 'roughness'. The pleasantness curve is barely influenced by the overall loudness development.

What else can we learn from this psychoacoustic analysis? If we compare the psychoacoustic analysis to a listener's sensation of pleasantness (e. g. in terms of musical tension), we might conclude that the maximum tension is reached when loudness reaches its maximum at roughly t = 7–8 s, and the first 17 seconds of this recording are characterized by a clear global maximum. Additionally, tension increases continuously over the first 8 seconds. The global climax of tension corresponds to the first chord in measure 2. If we compare this sensation of pleasantness to the features of the score (see arrows pointing to the corresponding events), we will find a clear correspondance between the climax of tension and the structural feature in form of the famous altered seventh chord – the so-called Tristan chord[58] – in measure 2.

---

[52] Sloboda et al. 2001.

[53] Zwicker/Fastl 1999, 245.

[54] For the definition of parameters see Zwicker/Fastl 1999.

[55] Wagner 1994.

[56] dBSONIC 2002.

[57] Hall 1991, 202.

[58] Cook 1987, 208.

To summarize, we can conclude that the pleasantness curve does not represent the sensation of pleasantness at all. The most obvious difference between subjective experience and displayed psychoacoustic curve is the 'near-sighted' plesantness curve, which is shaped by a sequence of local maxima and minima of roughness, but not by a global development. However, experiencing music is a different process: the current listening position is influenced by the music's past as well as by its' future (e. g. in form of expectation). This means that the listening experience of a human listener can only adequately be understood if we consider past, present and future in listening as well. These are simply top-down processes which consider, for example, musical expertise, familiarity with the particular style and musical memory. Utilizing this simple explorative approach to the *Tristan* opening we can conclude that sensualistic explanations of pleasantness are oversimplified and psychoacoustic parameters can only deliver an insufficient explanation of musical sensations.

4 DISCUSSION

The previous sections gave us insight into the role of data-driven (bottom-up) and context-driven (top-down) processes in music perception. However, an important aspect remains unconsidered: the role of cultural experiences and their relationship to sensory processes. This aspect will be discussed in this last section and we will also attempt to find a balanced view on the extreme positions of cultural relativism versus universalism. To illustrate the importance of this 'clash of positions' let us remember that the assumption of music as a universal language is a widespread belief. Surprisingly, the proponents of this *lingua universalis* position seem to be uninfluenced by the enormous diversity of human cultures. However, the current state of research in music psychology and the state of knowledge in related disciplines such as anthropology and culture psychology finds very little support for this 'naive' view. Moreover, as Kopiez[59] could rather reveal, this idea of music as a universal language can be explained as utopian belief in Western thinking.

As we can also conclude from the first three sections, the current state of psychological research enables only limited insight into the 'nature' of music sensation. This is due to several factors: firstly, research on sensory dissonance is limited to the parameter of pitch. Although the theory of harmonic intervals has a long tradition in Western music theory, this does not mean that the frequency domain is the only relevant perceptual domain determining pleasantness. This category only became a focus of psychological research due to an elaborated theory of auditory perception. At the current state of research we have to point out that there is a serious lack of theory in the time domain. In other words, there is no framework within which phenomena such as 'rhythmic dissonance' could be explained. There is evidence in non-Western musical cultures of the important role of 'rhythmic dissonance' which could, for example, be illustrated in terms of low vs. high rhythmic complexity or density (which is a key concept in African percussion music and labeled as 'échauffemement'[60]). Secondly, there is no systematic research into the way in which perceptual processes are modulated through cultural conventions. For example, let us consider the process of reading: results from dyslexia research show that the rate of reading disorders (dyslexia) differs between languages[61]: 4.3–6.4 % of Germans[62] and 5.3–11 % of Americans[63] suffer from this reading disorder with a mean international rate of 4–5 %. A possible explanation for these differences between languages could be that languages which have a high congruency between the written and spoken language, such as Italian, show a lower rate of dyslexia compared to languages which are characterized by significant deviations between spoken and written language.

At this point we should examine other related disciplines such as cross-cultural psychology or cultural anthropololgy and see what we can further learn about the importance of cultural conventions and the state of theoretical discussion concerning the relationship between 'nature' and 'nurture'.

4.1 CROSS-CULTURAL PSYCHOLOGY

Let us begin our short excursion with an overview of cultural anthropology, represented, for example, by the textbook *Humanity* by Peoples and Bailey. In this book, 'culture' is given a radical relativistic definition: "the culture of a group consists of shared, socially learned knowledge and patterns of behavior"[64]. Within this context, "shared"

---

[59] Kopiez in press.

[60] It is important to note that the term 'échauffemement' is not primarily related to the description of structural musical features but to the description of the increased movement intensity of the dancers within a performance. During the échauffemement the drumming ensemble reacts to the solo dancer by increasing tempo and loudness of drumming. For details see Beer 1991.

[61] For a review see Schulte-Körne/Remschmidt 2003.

[62] Haffner et al. 1998.

[63] Katusic et al. 2001.

[64] Peoples/Bailey 2000, 17.

means that culture is collective and shared by a group of people (e. g. Western culture); "socially learned" means that biological differences do not explain cultural differences; "knowledge" means the sum of a people's belief, rules etc. that lead to meaningful behaviour, and "patterns of behavior" means the observable action. From this definition we might conclude that within anthropology, the idea of culturally independent universals currently plays no role in the understanding of cultural behaviour. This finding should prevent us from rashly overemphasized interpretations of experimental results from psychological research. In People's book, not a single investigation into sensory processes is deemed worthy of discussion and music is merely described in the holistic framework of ritual practise such as dance, song and religion.

The second influential book, *Cross-cultural psychology* by Berry et al.[65] emphasizes that cross-cultural experiments can help us discover principles of perception and provide a better understanding of the importance of cultural conventions. Although most research reviewed is related to visual perception, the basic insight, namely that sensation and perception are always modulated by cultural experiences, remains fundamental. Phenomena considered include processes such as depth perception and visual illusions. The emphasis on cultural conventions is not trivial because they can determine artistic styles for centuries. However, cross-cultural experiments in music perception are rare and results are inconsistent. For example, in a cross-cultural study Balkwill and Thompson[66] investigated whether the mood intended in Hindustani ragas can be recognized by Western music listeners who are not familiar with this musical language. Authors claim that "listeners are sensitive to musically expressed emotion in an unfamiliar tonal system". However, statistical results are not that clear, and despite the restrictive experimental setting (subjects could only chose between four target emotions), a great deal of confusion could be observed: the intended emotion 'joy' was often rated as 'anger', and 'anger' received frequent ratings for 'peace'.

To summarize the cross-cultural approach, I would like to point out three theoretical positions in an attempt to explain cultural similarities and dissimilarities[67]: namely absolutism, universalism and relativism. The absolutist position views psychological behaviour as being independent from the cultural background and sees no problems with ethnocentrism in research methods. Differences are explained quantitatively (e. g. 'less intelligent'). The universalist position assumes that basic psychological processes are similar everywhere, but that their manifestations can be influenced by culture. Experimental methodology should be adopted to local cultural knowledge. Findings are interpreted in terms of 'weak , strong or strict universals'. Relativists believe that context-free measurement of psychological concepts is impossible and should not even be attempted, and that an effort should be made to understand people 'in their own terms' (this means 'in their own categories' and 'with their own values'). Thus, comparative approaches are avoided. This position is characteristic for modern anthropology.

4.2 Anecdotal Reports

Finally, I would like to close this last section with some anecdotal reports about attempts to understand foreign cultures. These reports illustrate the pitfalls of investigating non-Western cultures using Western methods and the sources of misunderstandings. One of the very first ethnological experiments in music was made by Erich M. von Hornbostel. In 1906, when a group of Hopi native Americans were invited to Berlin, Hornbostel took the chance and prepared some psychological investigations, including tests in rhythm, harmony and melody to test the 'musicality' of the subject. The written report revealed non-existing major-minor discrimination and inconsistent appreciation of Western musical concepts. A second report was given by the ethnomusicologist Rüdiger Schumacher as a reminiscence to his field research in Indonesia. When he played a Beethoven Symphony to the hotel staff to observe spontaneous reactions, listeners responded: "some nice melodies – but why all concurrently?"[68]. From this report we can learn that familiarity with the Western concept of polyphonic composition is a prerequisite for an adequate understanding of a Beethoven Symphony.

At the very end of this essay we might conclude that it is implausible to explain the diversity of musical cultures with simple psychophysiological concepts such as 'sensory dissonance'. There is much more convincing evidence of the dominance of culture over nature and for the modulation of basic sensory processes through higher cognitive processes. Although we cannot exclude that basic perceptual processes are very similar across cultures, we are uncertain whether they can ever be revealed in adequate experimental settings and question the significance of this finding for the understanding and sensation of 'real' music cultures. As long as these questions remain unan-

---

[65] Berry et al. 2002, chapter 8.

[66] Balkwill/Thompson 1999.

[67] Berry et al. 2002, 324.

[68] Schumacher 1999.

swered, elementary explanations such as the concept of ‘natural musical intervals’ are of questionable value, and the more realistic scenario for the experience of music from an unfamiliar culture is illustrated in the cartoon of Figure 5. In other words: without sufficient familiarity with the aesthetic concepts and the cultural background of a particular musical culture one always listens to music with one’s own ears – independent of the presence of ‘natural’ or ‘unnatural’ intervals.

BIBLIOGRAPHY

Balkwill, L.-L./Thompson, W. F. 1999
A Cross-cultural Investigation of the Perception of Emotion in Music: Psychophysical and Cultural Cues. Music Perception 17 fasc. 1, 43–64.

Beer, J. 1991
Comments to CD recordings. In: A. Simon (ed.), Rhythmen der Malinke – Guinea [Audio CD]. Museum Collection Berlin 18. Berlin.

Berry, J./Poortinga, Y. H./Segall, M. H./Dasen, P. R. 2002
Cross-cultural Psychology. Research and Applications^{2}. Cambridge.

Borchgrevink, H. M. 1975
Musikalske akkordpreferanser hos mennesket belyst ved dyreforosok [Musical chord preferences of humans, in the light of perceptual experiments with animals]. Tidsskrift for den Norske Laegforening 95 fasc. 6, 356–358.

Brass, D. 1995
Diep1d.exe [Computersoftware for the simulation of wandering waves on the basilar membrane. Available via http://www.oae-ilo.co.uk]. Hatfield, UK.

Cook, N. 1987
A Guide to Musical Analysis. Oxford.

dBSONIC 2002
Version 3.0 [Computersoftware for psychoacoustic analysis. For description see http://www.cortex-instruments.de]. Regensburg.

Diependaal, R. J./Duifhuis, H./Hoogstraten, H. W./Viergever, M. A. 1987
Numerical Methods for Solving One-dimensional Cochlear Models in the Time Domain. Journal of the Acoustical Society of America 82 fasc. 5, 1655–1666.

Ehret, G./Merzenich, M. M. 1988
Complex Sound Analysis (Frequency Resolution, Filtering and Spectral Integration) by Single Units of the Inferior Colliculus of the Cat. Brain Research Reviews 472 fasc. 2, 139–163.

Fishman, Y. I./Volkov, I. O./Noh, M. D./Garell, P. C./Bakken, H./Arezzo, J. C./ Howard, M. A./Steinschneider, M. 2001
Consonance and Dissonance of Musical Chords: Neural Correlates in Auditory Cortex of Monkeys and Humans. Journal of Neurophysiology 86 fasc. 6, 2761–2788.

Fletcher, H. 1938
The Mechanism of Hearing as Revealed through Experiments on the Masking Effect of Thermal Noise. Proceedings of the National Academy of Sciences 24, 265–274.

Greenwood, D. D. 1961a
Auditory Masking and the Critical Band. Journal of the Acoustical Society of America 33 fasc. 4, 484–502.

Greenwood, D. D. 1961b
Critical Bandwidth and the Frequency Coordinates of the Basilar Membrane. Journal of the Acoustical Society of America 33 fasc. 10, 1344–1356.

Haffner, J./Zerahn-Hartung, C./Pfuller, U./Parzer, P./Strehlow, U./Resch, F. 1998
Auswirkungen und Bedeutung spezifischer Rechtschreibprobleme bei jungen Erwachsenen – empirische Befunde in einer epidemiologischen Stichprobe. Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie 26, 124–135.

Hall, D. E. 1991
Musical Acoustics^{2}. Pacific Grove, CA.

Hargreaves, D. J. 1982
The Development of Aesthetic Reactions to Music. Psychology of Music, Special Issue, 51–54.

Heffner, H./Whitfield, I. C. 1976
Perception of the Missing Fundamental by Cats. Journal of the Acoustical Society of America 59 fasc. 4, 915–919.

Helmholtz, H. v. 1954
On the Sensations of Tone as a Physiological Basis for the Theory of Music. 2nd English ed. by A. J. Ellis. New York.

Itti, L. 1998
Seeing the World through a Retina. Retrieved, from the World Wide Web: http://www.klab.caltech.edu/~itti/retina/index.html.

Janata, P./Birk, J. L./Tillmann, B./Bharucha, J. J. 2003
Online Detection of Tonal Pop-out in Modulating Contexts. Music Perception 20 fasc. 3, 283–305.

Janata, P./Birk, J. L./van Horn, J. D./Bharucha, J. J. 2002
The Cortical Topography of Tonal Structures Underlying Western Music. Sciene 298, 13 December, 2167–2170.

KATUSIC, S. K./COLLIGAN, R. C./BARBARESI, W. J./SCHAID, D. J./JACOBSEN, S. J. 2001
Incidence of Reading Disability in a Population-based Birth Cohort, 1976–1982, Rochester, Minn. Mayo Clinic Proceedings 76 fasc. 11, 1075–1077.

KOELSCH, S. 2002
Bach Speaks: a Cortical 'Language-network' Serves the Processing of Music. Neuroimage 17 fasc. 2, 956–966.

KOELSCH, S./FRITZ, T./GUNTER, T. C./CRAMON, D. Y. V./ZYSSET, S./LOHMANN, G./FRIEDERICI, A. D. in preparation
Investigating Emotion with Music: an fMRI Study.

KOELSCH, S./MAESS, B./GUNTER, T. C./FRIEDERICI, A. D. 2001
Neapolitan Chords Activate the Area of Broca. A Magnetoencephalographic Study. In: R. J. Zatorre/I. Peretz (eds), The Biological Foundations of Music. Annals of the New York Academy of Sciences 930, 420–421. New York.

KOELSCH, S./MULDER, J. 2002
Electric Brain Responses to Inappropriate Harmonies during Listening to Expressive Music. Clinical Neurophysiology 113 fasc. 6, 862–869.

KOELSCH, S./SCHROGER, E./TERVANIEMI, M. 1999
Superior Pre-attentive Auditory Processing in Musicians. Neuroreport 10 fasc. 6, 1309–1313.

KOPIEZ, R. in press
Der Mythos der Musik als universell verständliche Sprache. In: C. Bullerjahn/W. Löffler (eds), Mythos Musik. Hildesheim.

KRUMBHOLZ, K./SCHMIDT, S. 1999
Perception of Complex Tones and its Analogy to Echo Spectral Analysis in the Bat, Megaderma Lyra. Journal of the Acoustical Society of America 105 fasc. 2, 898–911.

KRUMBHOLZ, K./SCHMIDT, S. 2001
Evidence for an Analytic Perception of Multiharmonic Sounds in the Bat, Megaderma Lyra, and its Possible Role for Echo Spectral Analysis. Journal of the Acoustical Society of America 109 fasc. 4, 1705–1716.

LEBLANC, A. 1991
Effect of Maturation/Aging on Music Listening Preference: a Review of the Literature. Ninth National Symposium on Research in Music Behavior, March 7–9. Cannon Beach, Oregon.

LEBLANC, A./SIIVOLA, C./OBERT, M./SIMS, W. L. 1996
Music Style Preferences of Different Age Listeners. Journal of Research in Music Education 44 fasc. 1, 49–59.

LYNCH, M. P./EILERS, R. E./OLLER, D. K./URBANO, R. C. 1990
Innateness, Experience, and Music Perception. Psychological Science 1 fasc. 4, 272–276.

PEOPLES, J./BAILEY, G. 2000
Humanity. An Introduction to Cultural Anthropology^5. Belmont, CA.

PICKLES, J. O. 1979
Psychophysical Frequency Resolution in the Cat as Determined by Simultaneous Masking and its Relation to Auditory-nerve Resolution. Journal of the Acoustical Society of America 66 fasc. 6, 1725–1732.

PLOMP, R./LEVELT, W. J. M. 1965
Tonal Consonance and Critical Bandwidth. Journal of the Acoustical Society of America 38 fasc. 4, 548–560.

PREISLER, A./SCHMIDT, S. 1995
Virtual Pitch Formation in the Ultrasonic Range. Naturwissenschaften 82, 45–47.

SCHELLENBERG, E. G./TREHUB, S. E. 1996
Natural Musical Intervals: Evidence from Infant Listeners. Psychological Science 7 fasc. 5, 272–276.

SCHULTE-KÖRNE, G./REMSCHMIDT, H. 2003
Legasthenie – Symptomatik, Diagnostik, Ursachen, Verlauf und Behandlung. Deutsches Ärzteblatt 100 fasc. 7, 396–406.

SCHUMACHER, R. 1999
Personal communication, 2 August.

SLOBODA, J. A./O'NEILL, S./IVALDI, A. 2001
Functions of Music in Everyday Life: An Exploratory Study Using the Experience Sampling Method. Musicae Scientiae 5 fasc. 1, 9–32.

STANDLEY, J. M./MADSEN, C. K. 1990
Comparison of Infant Preference and Responses to Auditory Stimuli: Music Mother, and other Female Voice. Journal of Music Therapy 27 fasc. 2, 54–97.

TERHARDT, E. 1974
Pitch, Consonance, and Harmony. Journal of the Acoustical Society of America 55 fasc. 5, 1061–1069.

TERHARDT, E. 1998
Akustische Kommunikation. Berlin.

TERHARDT, E./STOLL, G./SEEWANN, M. 1982
Pitch of Complex Signals According to Virtual-pitch Theory: Tests, Examples, and Predictions. Journal of the Acoustical Society of America 71 fasc. 3, 671–678.

THOMPSON, W. F./SCHELLENBERG, E. G. 2002
Cognitive Constraints on Music Listening. In: R. J. Colwell/C. Richardson (eds), The New Handbook of Research on Music Teaching and Learning. New York, 461–486.

TRAINOR, L. J./HEINMILLER, B. M. 1998
Natural Musical Intervals: Evidence from Infant Listeners. Infant Behavior and Development 21 fasc. 1, 77–88.

TRAINOR, L. J./TSANG, C. D./CHEUNG, V. H. W. 2002
Preference for Sensory Consonance in 2- and

4-month-old Infants. Music Perception 20 fasc. 2, 187–194.

TRAMO, M. J./CARIANI, P. A./DELGUTTE, B./BRAIDA, L. D. 2001
Neurobiological Foundations for the Theory of Harmony in Western Tonal Music. In: R. J. Zatorre/I. Peretz (eds), The Biological Foundations of Music. Annals of the New York Academy of Sciences 930. New York, 92–116.

TREHUB, S. E. 2000
Human Processing Predispositions and Musical Universals. In: N. L. Wallin/B. Merker/S. Brown (eds), The Origins of Music. Cambridge, MA, 427–448.

TREHUB, S. E./SCHELLENBERG, E. G./HILL, D. 1997
The Origins of Music Perception and Cognition: A Developmental Perspective. In: I. Deliège/J. Sloboda (eds), Perception and Cognition of Music. Hove, 103–128.

WAGNER, R. 1994
Wagner Orchestral Music. CD BMG Classics 74321-17893-2 Performed by E. Ormandy/Philadelphia Orchestra.

WEINBERGER, N. M. 1999
Music and the Auditory System. In: D. Deutsch (ed.), The Psychology of Music. San Diego, 47–88.

WHITFIELD, I. C. 1980
Auditory Cortex and the Pitch of Complex Tones. Journal of the Acoustical Society of America 67 fasc. 2, 644–647.

WOLPERT, R. 2000
Attention to Key in a Non-directed Music Listening Task. Music Perception 18 fasc. 2, 225–230.

ZENTNER, M. R./KAGAN, J. 1998
Infants' Perception of Consonance and Dissonance in Music. Infant Behavior and Development 21 fasc. 3, 483–492.

ZWICKER, E./FASTL, H. 1999
Psychoacoustics. Facts and Models^2. Berlin.

ZWICKER, E./FLOTTROP, G./STEVENS, S. S. 1957
Critical Band width in Loudness Summation. Journal of the Acoustical Society of America 29 fasc. 5, 548–557.

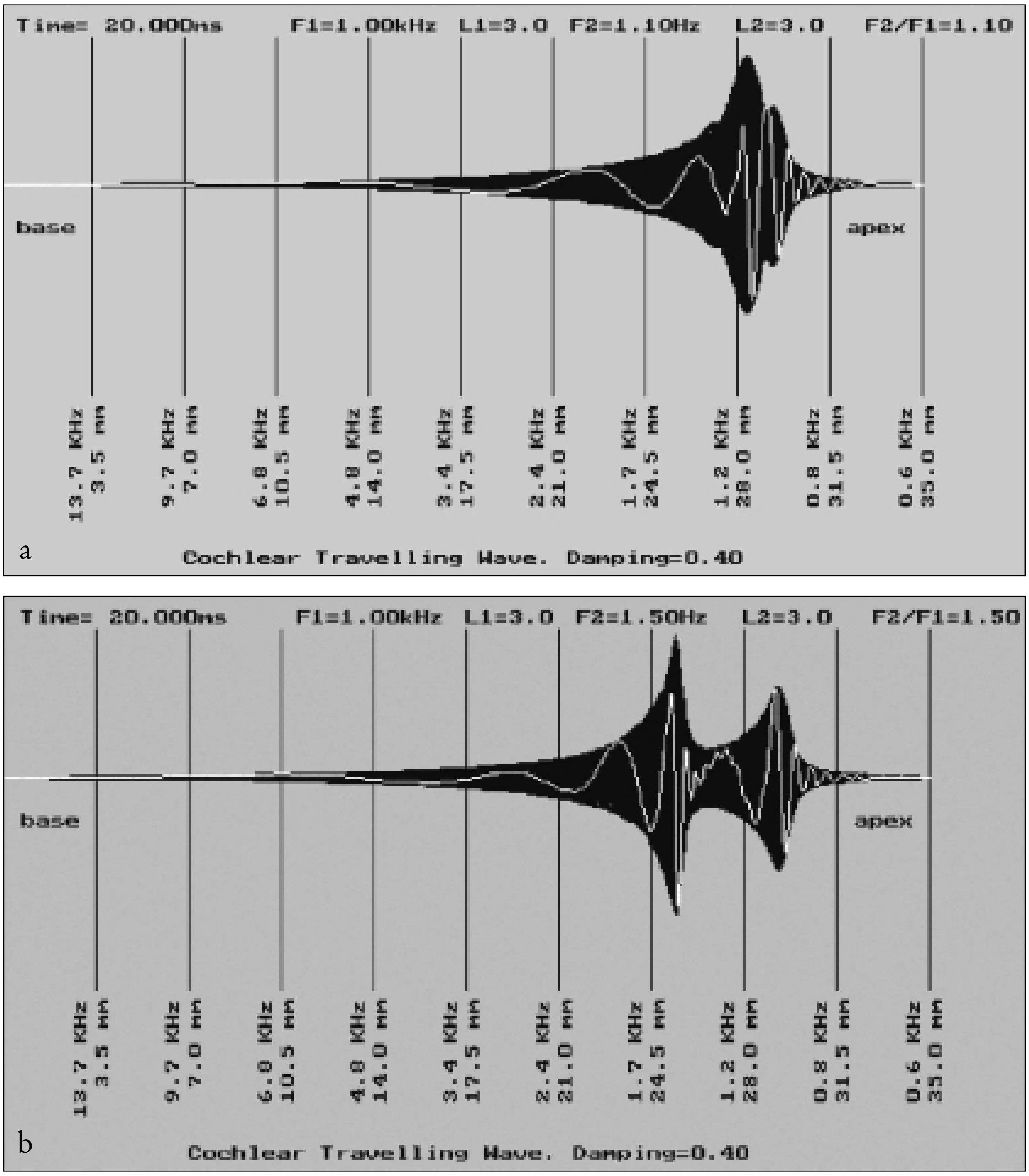


Fig. 1 Simulation of the first 20 ms of a wandering wave. – a. of a simultaneous 1 and 1.1 kHz. – b. of a 1 and 1.5 kHz sinusoid on the basilar membrane.

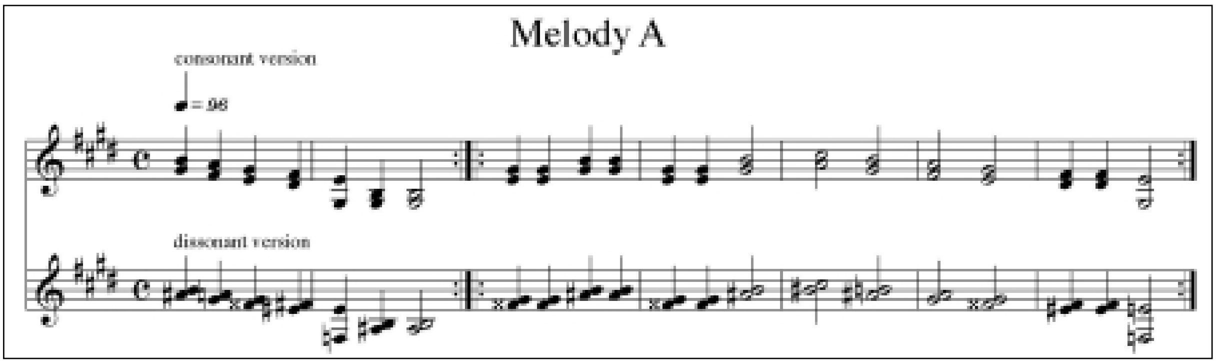
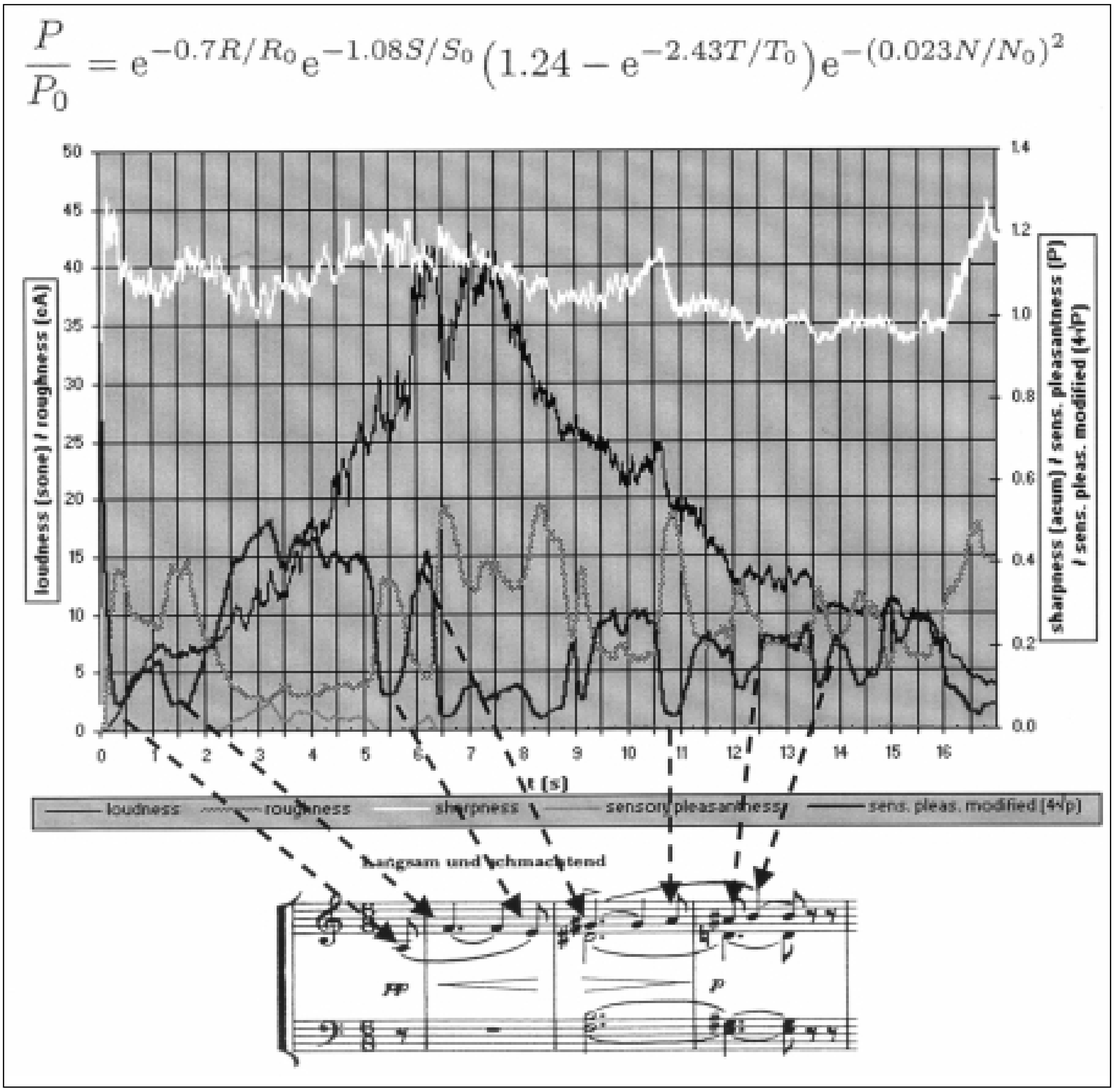


Fig. 2 Consonant and dissonant version of a melody as used in the experiment by Zentner/Kagan 1998.

Fig. 3 Simulation of the reproduction quality of the human eye (upper figure). The spot in focus in the middle represents the area of sharpest vision on the retina, the so-called *fovea centralis*. The black dot corresponds with the so-called blind spot (exit of the optic nerve). The lower figure is a reproduction of the same scene by a camera and corresponds to our perceived impression of a picture in focus (based on Itti 1998).



$$\frac{P}{P_0} = \mathrm{e}^{-0.7R/R_0}\mathrm{e}^{-1.08S/S_0}\left(1.24 - \mathrm{e}^{-2.43T/T_0}\right)\mathrm{e}^{-(0.023N/N_0)^2}$$

Fig. 4 Predicted sensory pleasentness curve (according to Zwicker and Fastl's equation) and related psychoacoustic parameters for the first 17 seconds of Wagner's *Tristan Prelude*.

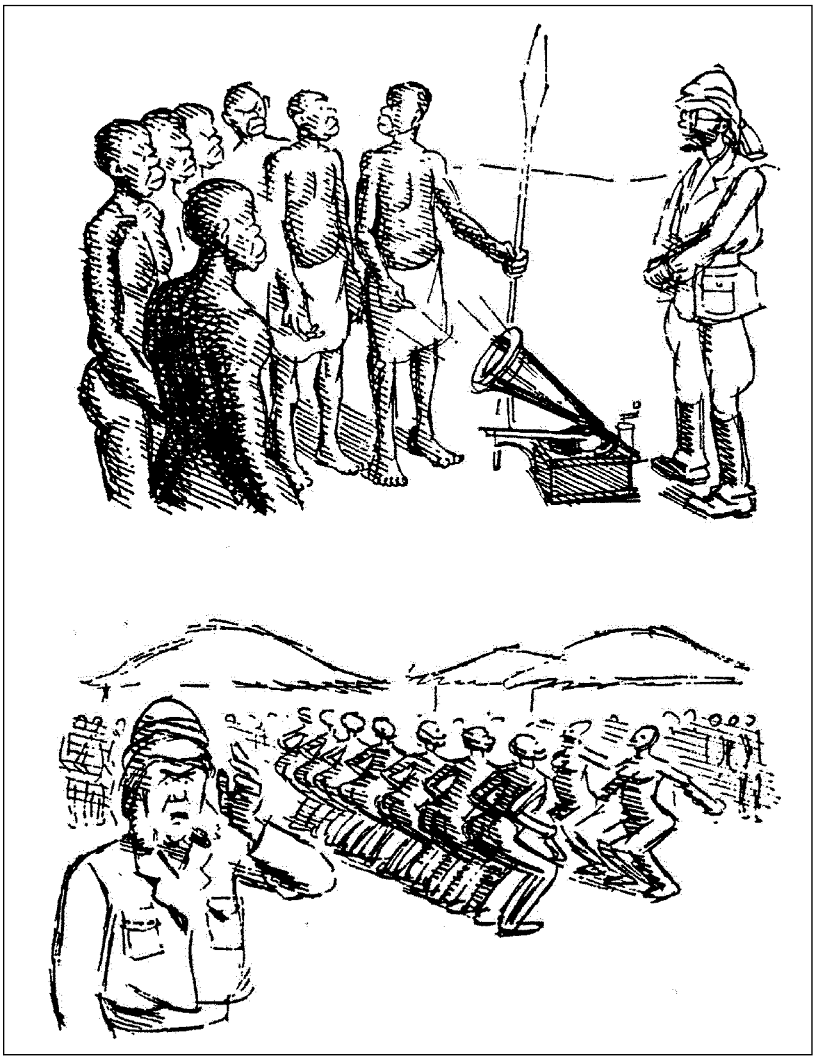


Fig. 5 One always listens to music with one's own ears (based on Williams 1967, 18–19).