Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

Neuroscientific Measure of Consonance

Adrian Foltyn[*1]

[*]*Department of Composition, Conducting and Theory of Music, F. Chopin University of Music, Poland*
[1]mail@adrianfoltyn.pl

ABSTRACT

The article contains a proposition of new simplified model of neural discrimination of sensory consonance / dissonance at higher stages of auditory pathway. The model regards primarily complex harmonic sounds and is based on periodicity / pitch and its representation in neural discharges. The hypothesis relies on a process involving measuring concentration of neural excitation in *inferior colliculus* in time windows equal to period of sum of the incoming signals. The measure can accommodate pitch deviations via a further mechanism based on harmonic entropy and can be applied to any interval, including microtones and octave enhancements. For simple ratios an algebraic calculation method is available, accounting for several interval relations abstract mathematical consonance measures tended to struggle with. To examine plausibility of the model, a psychoacoustic experiment was carried out, using paired comparison of intervals. One of the resulting dimensions can be clearly identified as consonance – dissonance axis. The proposed modelled consonance values together with 4 other well-known models have been related to experimental results. Logarithmic transformation of the postulated consonance measure displays the highest correlation with the consonance dimension obtained in the experiment out of all examined models ($R^2 \approx 0.8$). Higher degree of correlation versus roughness-based models suggests plausibility of certain pitch-related mechanism underlying basic consonance perception.

I. BACKGROUND

Consonance and dissonance (C/D) sensations have been subject to attempts of modelling since the very Pythagorean times. The notion of musical tension followed by relaxation had profound implications for most prominent composers and entire musical epochs (cf. Voigt, 1985). Along with the advancing knowledge of human psychoacoustics and neurophysiology, sensory dissonance has been separated from its purely musical and culture – dependent sense (Terhardt, 1974).

Sensory dissonance is said to consist of several sensations connected with spectrum and amplitude of the signal (Terhardt, 1974), namely roughness, sharpness, loudness and toneness (opposite of noisiness). Until present, significant part of theoretical models of sensory C/D has relied on roughness, a sensation arising from beating of adjacent partials (e.g. Plomp & Levelt, 1965, Kameoka & Kuriyagawa, 1969, Terhardt, 1974, Hutchinson & Knopoff, 1978). Highest dissonance sensation is observed at approximately one fourth of the critical band (roughly equivalent to width of the auditory filter connected with specific frequency).

However, sensory C/D models based on roughness have provided largely insufficient explanation for psychoacoustic C/D data (Parncutt, 2006). Furthermore, roughness-based theories fail to account for:

- Dissonance of pure tones more than one critical band apart
- Dissonance sensation of distant complex tones (e.g. 3 or more octaves apart), in which case high-order weak partials of the lower complex tone hardly can contribute to any noticeable roughness
- Dissonance of successive tones
- Results of psychoacoustic research where respondents grouped roughness with perceived density, i.e. number of tones, and fusion, as opposed to consonance / dissonance being grouped with euphoniousness and beautifulness (Van de Geer et al., 1962),
- Results of psychoacoustic research of patients with brain lesions, whose consonance perception is impaired in line with pitch perception impairment, contrary to roughness perception remaining nearly unaffected (Tramo et al., 2001)

Conversely, another class of models describe C/D as a pitch-related sensation (e.g. Licklider, 1951, Patterson, 1987, Langner et al, 1988, Meddis & Hewitt, 1991). Based on results of research on non-human mammals, human PET / FMRI examinations and computational models, temporal coding has been confirmed as a viable explanation of numerous psychoacoustic phenomena.

II. AIMS

In line with the latter group of models, the article is grounded in the view that for typical harmonic sounds there exists a central (neural) mechanism which produces a C/D sensation based on periodicity (pitch) analysis in auditory midbrain.

Since none of the sensations constituting sensory C/D in Terhardt's sense stems directly from key perceptive musical element – pitch, I postulate existence of 3 basic levels of consonance / dissonance:

- First-order sensory C/D produced by pitches in non-outermost registers of musical instruments and human voice (approx 35 – 2500 Hz, i.e. F2 – E7)
- Second-order sensory C/D produced by non-pitch sound characteristics (roughness, loudness, sharpness, non-noisiness)
- Musical C/D, as a product of culture, incl. functional and contrapuntal application of it in music (cf. Tenney, 1988).

Consequently, the aim of the article is to hypothesize a simplified mechanism of analyzing neural spike train data which could occur at higher stages of auditory pathway and lead to discerning consonance / dissonance, quantify its output, and examine its plausibility by comparison with psychoacoustic data against several other C/D models.

III. MAIN CONTRIBUTION

The following simplified model of C/D perception only applies to pure tones and harmonic complex sounds. It is therefore necessary to assume that the proposed measure of C/D refers to sounds containing no outstanding levels of other spectral (roughness, sharpness, noisiness) and amplitude factors contributing to second-order sensory dissonance.

Furthermore, it is evident that all neural groups discharge not solely in synchrony with the incoming signal's frequencies. In this context, spikes in the proposed model represent higher probabilities of discharge (rather than certainty of eliciting actual responses) as opposed to lower probabilities which have been simplified into non-discharges. The assumption here is that at higher stages of the auditory pathway C/D analysis mechanisms are sensitive only to certain threshold level of excitation which is represented by mentioned spike – non-spike model.

A. Coincidence models.

Transformation of spatial information from basilar membrane into temporal code of neural discharges reflects the frequencies present in the incoming signal. This phenomenon, called phase locking, occurs at nearly all stages of the auditory pathway, including auditory nerve fibre, cochlear nucleus, inferior colliculus and auditory cortex (Tramo et al., 2001). However, additional mechanism is necessary to account for actual pitch perception. Several recent models (e.g. Langner et al. 1988, Meddis & Hewitt, 1991, Nelson & Carney, 2004) provide explanation based on coincidence of neural discharges in inferior colliculus (IC).

By virtue of coincidence, neurons in IC respond with spikes to the modulation frequency, representing pitch, and to its integer multiples. According to mentioned models, pitch is then determined through an autocorrelation of inter-spike intervals (Meddis & Hewitt, 1991) and thanks to inhibition from lateral lemniscus (VNLL – Langner, 1988).

The inhibition is effective after approximately 30 ms by suppressing all responses elicited by neurons tuned to integer multiples of the fundamental frequency (equal to higher partials of the signal) and changing initial comb filter into low-pass one. Low pass feature has further consequence in the fact, that on times equal to integer multiples of the fundamental period T ($2T$, $3T$ etc.) neurons tuned to these periods will still be discharging.

In view of my hypothesis, another important characteristic provided by coincidence models is the apparent orthogonal representation of signal's periodicity (pitch) and fine structure (timbre) in inferior colliculus. It is even possible that separate groups of onset neurons exist in nuclear cochleus, which either provide broad (i.e pitch-oriented) or narrow (i.e. harmonics-oriented) integration windows, allowing for respective accuracy level of encoding signal (Bahmer, 2007). This split could provide additional supporting evidence for postulated threefold understanding of C/D (see II).

B. Concentration window.

Two harmonic complex sounds elicit responses of those IC neurons, whose best modulation frequency (BMF) is equal to period of fundamental or residual pitch of incoming signal, and its integer multiples. As a starting point, the hypothesis requires existence of a mechanism determining length of the period of sum of the two signals, which I will call the *concentration window (CW)*.

Length of the CW can be determined directly from first-order inter spike intervals (i.e. distances between neighbouring discharges) in IC. The first instance of a time interval repeating between responses to two signals denotes their sum's period length, i.e. CW (see Figure 1).

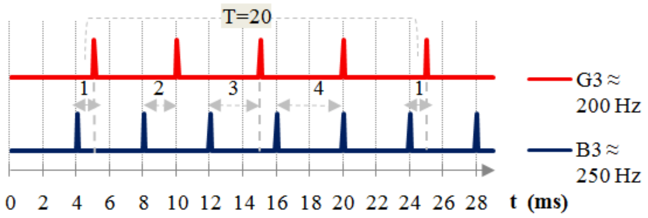

Figure 1. Simplified discharge patterns (fundamentals only) in response to musical sounds close to G3 and B3, forming a just major third. First instance of repeating time interval (e.g. 1 ms) denotes length of concentration window (20 ms).

C. Tolerance factor.

Observation of musical practice leads to the conclusion that C/D judgments are very similar in the case of dyads differing by intonation only (e.g. just vs well-tempered). Since the concentration windows will also fractionally differ between such cases, a certain tolerance factor needs to be assumed to reflect similarity in their perception. The underlying adjustment is most likely performed at higher stages of auditory pathway and reflects the idea of human cognition being able to fit incoming information into known or simple patterns. One plausible concept visualising this mechanism is based on information theory and named harmonic entropy (Erlich, 1997). It denotes the degree of uncertainty a given interval evokes while being analysed by the auditory system.

Harmonic entropy (HE) concept employs so called Farey series (an ordered sequence of ratios consisting of natural numbers not higher than n). One of key features of the series is that the simpler the ratio, the further apart it is from their neighbouring elements. Thus, the series can be used to model the probability of perceiving a given interval (ratio i) as the mistuning of a j-th series member (Sethares, 1999). In accordance with definition of entropy (Shannon, 1948), HE is expressed as the sum of products of all such probabilities for ratio i and their logarithms:

$$HE(i) = -\sum_j p_j(i) * \log(p_j(i)) \qquad (1)$$

where $p_j(i) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x \in r_j} e^{\frac{-(x-i)^2}{2\sigma^2}} dx$, i.e. the probability has normal distribution with assumed standard deviation σ equal to about 0.7% (Sethares, 1999).

The graph of HE (see Figure 3) displays clear valleys which occur very close to cent values representing basic just intervals. Hence, each of these intervals has a certain perceptive 'gravity field' which draws perception of a mistuned interval to its tuned version if occurring within that field. Since concentration window is determined on first-order inter-spike intervals, gradual shift from one just interval to a neighbouring one (e.g. from i_1 to i_2) will be reflected in gradual decay of the CW for ratio i_1 and accompanying appearance of CW of i_2 (Figure 2).

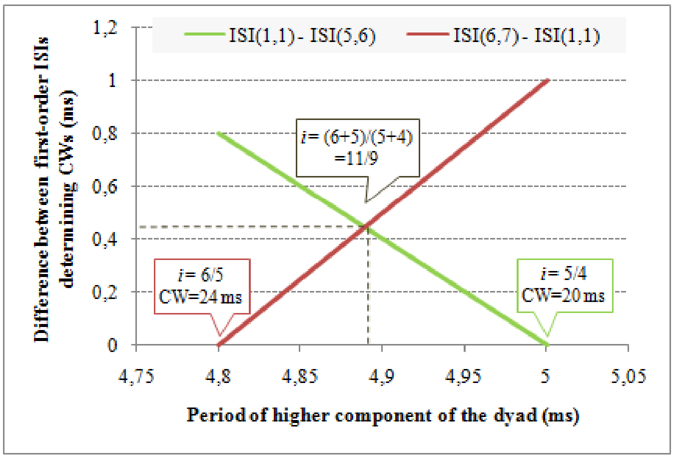

Figure 2. Time differences between first-order ISIs determining concentration windows for ratios between 6/5 and 5/4. Period of the lower component of the dyad is 4 ms. Periods up to 4,89 ms (4 ms * 11 / 9, which is the mediant between 6/5 and 5/4) may be judged more as a mistuning of ratio 6/5, periods above this threshold may be perceived as closer to ratio 5/4.

It is therefore reasonable to assume that the shift of perception from i_1 towards i_2 occurs approximately at mediant between both ratios. Thus, for each possible ratio within one gravity field, we obtain the tolerance factor *TF*, applied further as a correction to consonance measure:

$$TF(i) = \frac{HE(i)}{HE(i^*)} \quad , \qquad (2)$$

where i^* is the central ratio of the 'gravity field'.

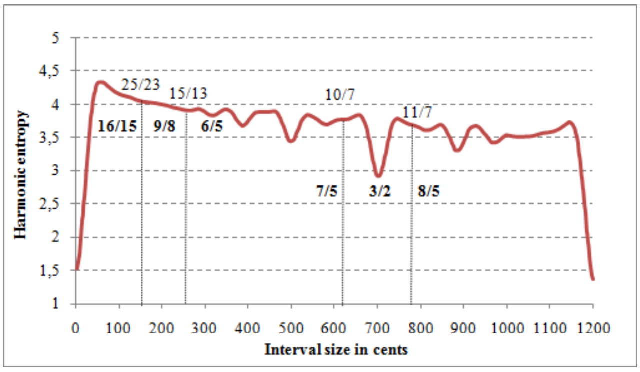

Figure 3. Harmonic entropy computed for intervals within one octave (n = 80, Erlich, 1997). The perceptive 'gravity field' of small-ratio intervals (i.e. distance between neighbouring mediants, shown as non-bolded fractions) is stronger than at higher ratios, as shown in the case of ratios 9/8 (major second) and 3/2 (perfect fifth).

One can also hypothesise that the resolution of HE 'gravity fields' may be subject to learning at higher stages of auditory pathway, which would explain why accomplished contemporary music performers and composers are able to easily discern separate pitches for quarter-tones or other microtones.

D. Concentration of neural excitation.

Models assuming usage of autocorrelation and all-order ISIs (i.e. time distances between all discharges in response to all incoming signals - see examples in III.A) by auditory brain both for determining dyad / chord period and strength of C/D sensation are conceivable, but also certainly requiring substantial working memory capacity. This would correspond well with the type of complex analysis needed for perception of spectral elements of the signal, constituting second-order sensory C/D. Instead, for purposes of determining first-order C/D a simpler process is hypothesized with the following assumptions:

- The response patterns are simplified into spike trains of neurons set to period lengths equal to integer multiples of respective fundamentals (see III.A);
- Discharge patterns are stored in working memory which enables auditory brain to both assess multiple simultaneous signals (usually shifted in phase versus each other) and multiple successive signals.
- Concentration window is established as shown in III.B.
- For dyads, concentration of neural excitation at time point lying one CW length from start of signal is measured for both series of firings. Number of group firings at these 'concentration spots' (CS) is compared to total number of group firings elicited by both signals until CS. High concentration of neural excitation results in consonance perception whereas low concentration (high dispersion) causes dissonant sensation (see example in Figure 4).

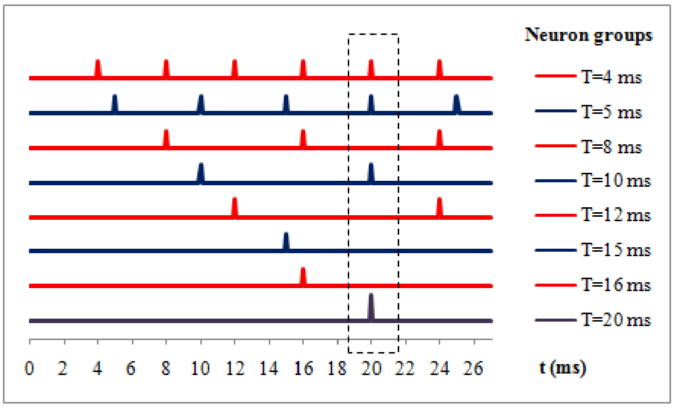

Figure 4. Modelled simplified discharge pattern in response to complex sounds close to G3 and B3, i.e. major third. CW is equal to 20 ms, and there are 5 groups of neurons discharging at CS lying 20 ms from start of both signals (both signals elicit response from neurons tuned to period of 20 ms, so they are double-counted). Since the number of responses until CS is 18, the consonance value is equal to 5/18 (0.278).

- For intervals higher than octave not all firing points will bear the same information content needed to determine CW length, as between any two discharges responding to the lower frequency of the dyad there is more than one firing point in response to higher frequency. This variation is reflected by:
  - Applying weight of *1–1/E(i)* (*E(i)* is the integer part of the ratio *i*) to part of responses to higher frequency equal to *1/E(i)* for all ratios with denominator higher than 1,
  - Applying weight of *1-1/(E(i-1)/2)* to part of responses to higher frequency equal 1/[*E(i-1)/2*] for all ratios containing denominator equal to 1,

- Applying weight of 1 to all other responses.

The same weights are applied to firings at respective CSs, but only in case the denominator of ratio *i* is an even number (because of possible occurrence of weighted firings at CSs). Odd denominators do not need this adjustment as the CS remains unaffected.

- For triads and chords consisting of more than 3 components an analogous process might take place. E.g. in triad case, each pitch evokes a firing pattern with two concentration sub-windows related to each of the other pitches. For each of the pitches, the concentration sub-window which is longer in time, denotes the length of overall concentration window. Consonance degree is then expressed as sum of discharges at all concentration spots in relation to all discharges in the overall CW.

E. Quantifying C/D.

The modelled consonance measure for dyads can be analytically presented in form of:

$$C(i) = \frac{CSF_A + CSF_B * w}{CWF_A + CWF_B * w} * TF \quad , \qquad (2)$$

where CSF_A, CSF_B are firings in response to signal A and B respectively at CS (B is the signal of higher frequency), CWF_A, CWF_B are total number of firings in response to signal A and B within the CW, *w* is the weight computed as shown III.D (for intervals within an octave it's equal to 1), and *TF* is the tolerance factor, as shown in III.C.

Theory of numbers allows for optimising calculations for integer and relatively simple proportions, e.g. just ($TF = 1$) or equally tempered ($TF \approx 1$) intervals. Using the Lejeune – Dirichlet formula (Sierpinski, 1948), for a simple ratio $i = p/q$ consonance degree can be derived as:

$$C\left(\frac{p}{q}\right) = 1 - \frac{T(p-1) + T(q-1)}{T(p) + T(q)} \quad , \qquad (3)$$

where *T(x)* is the sum of number of divisors of natural numbers not greater than *x*. First ten elements of *T(x)* series are presented in Table 1.

Table 1. First ten elements of the T(x) series.

x	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
T(x)	1	3	5	8	10	14	16	20	23	27

Analytically, T(x) is expressed as:

,
$$T(x) = 2\sum_{n=1}^{\sqrt{x}} E\left(\frac{x}{n}\right) - [E(\sqrt{x})]^2 \qquad (4)$$

Consonance values computed upon these assumptions are displayed in Figure 5. The values are by default normalised between 0 and 1 (maximum assigned to unison), and clearly represent a scale which is subject to logarithmic perception, similarly to pitch or acoustic pressure. Hence in comparisons it is more convenient to use the transformation to dissonance: $D = -\ln(C)$.

For triads the formula (2) can be transformed into:

$$C(i_3) = \frac{\sum_{j=1}^{3}\sum_{k \neq j}^{3} CSF_{jk}}{\sum_{j=1}^{3} CWF_j} * TF \qquad (5)$$

where:

- i_3 is the triad consisting of relations: a_1/b_1, b_2/c_1, c_2/a_2 ;
- CSF_{jk} is the number of firings in response to *j*-th pitch at the CS with *k*-th pitch,
- CWF_j is the total number of firings in response to *j-th* pitch until end of its overall CW,
- *TF* is the tolerance factor, computed as the product of tolerance factors stemming from each of the 3 dyads the triad consists of.

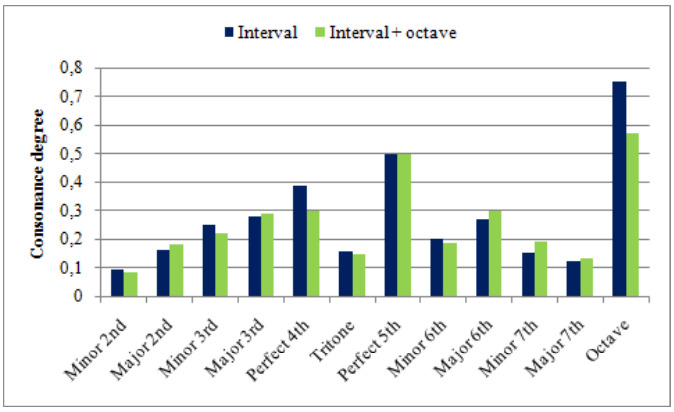

Figure 5. Consonance raw values for intervals up to two octaves, computed according to the postulated model. In most cases the values for a given interval and its octave enhancement differ from each other. The differences tend to be more meaningful as number of octaves increases (not shown).

For maintaining simplicity, weights have been omitted in the formula, however they obviously apply in the same way as for dyads, i.e. in the case of any component interval of the triad exceeding octave.

The consonance values of triads are directly incomparable to dyads or any other chords (cf. Figure 6). Their eventual perception-driven logarithmic transformation needs to take into account doubled number of concentration spots per pitch, e.g. by using a factor 3/2 as the logarithm's power.

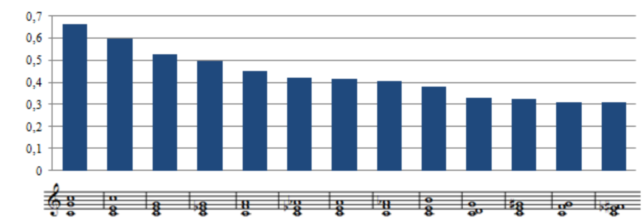

Figure 6. Consonance degree of computed for selected triads. Raw consonance values are not comparable between chords of different number of components. Similarly to dyads, the scale can be subject to logarithmic transformation to reflect non-linear perception.

IV. METHOD

To verify if the model yields plausible results, an experiment was carried out on a sample of 50 music-trained respondents (students and graduates of the music university). Respondents wearing standardised headphones were confronted with a computer interface which produced sound stimuli and recorded participants' answers.

119 intervals (dyads) had been selected as the experiment stimuli. Interval were then split into 3 groups (A, B, C) based on a pilot study, the groups representing most probable high, medium and low consonance values. The middle group (B), consisting of 50 intervals, shared 5 intervals with each of the outer groups having 40 intervals each.

For purposes of the experiment, 88 equally tempered real piano samples (i.e. preserving non-harmonicities of piano strings) were split into 3 groups representing 3 registers:

- Very high, i.e. pitches above 2500 Hz, where most probably phase locking mechanism decreases its accuracy, which may also be due to the smallest possible periodicity preference of neurons in cochlear nucleus equal to 0.4 ms (Langner, 2007). Taking into account, that piano samples carry slightly different sound qualities in highest register due to lack of key dampers, this group was extended to start around 2300 Hz (from D7);
- Typical, i.e. pitches between 40 and 2300 Hz;
- Very low, i.e. samples below 40 Hz (below E1), which again provide difficulties for listeners in determining their pitch; one of the reasons could be the decreased occurrence of long integration periods of phase-coupled neurons above approx. 25 ms (Davis et al., 2010).

The samples were also unified in terms of attack. Sound pressure level was optimised between the registers to facilitate perception of very large intervals. The dyads were selected according to the following setup (Table 2):

Table 2. Intervals used in the experiment.

<table>
<tr><th>Register / Type of occurrence</th><th>Very low (A0- E♭1)</th><th>Typical (E1 - E♭7)</th><th>Very high (D7-C8)</th></tr>
<tr><td>Both pitches within one register</td><td>6 intervals, 1 - 6 semitones</td><td>69 intervals, 1 - 69 semitones</td><td>10 intervals, 1 - 10 semitones</td></tr>
<tr><td rowspan="3">Pitches belonging to different registers</td><td colspan="2">13 intervals, 13-24 semitones and 1 semitone</td><td></td></tr>
<tr><td></td><td colspan="2">18 intervals, 13 - 24, 70 - 72, 75, 78 and 79 semitones</td></tr>
<tr><td>3 intervals, 83 - 85 semitones</td><td></td><td>(same register)</td></tr>
</table>

The experiment was carried out in modified incomplete paired comparison mode, with each respondent assessing two series of 65 direct comparisons of 2 intervals. Each series consisted of 20 comparisons in group A, further 25 in group B and 20 in group C. The order of comparisons and intervals within one comparison as well as pitches within predefined ranges were randomly selected by a Max/MSP patch, which also served as an interface for the respondents.

Each comparison consisted of presentation of two harmonic 3-second-long dyads, separated by 1-second silence. Respondents were asked to judge, which of the two intervals sounds more consonant to them, without using any information they might possess as a result of music theory education. The assessments were marked on a continuous scale, the ends of which represented maximal difference in consonance sensation (in either interval's direction), whereas the middle was assigned to zero difference between the two dyads. Respondents were given unlimited possibility of repeating first, second or both dyads. However, data records show that no pair of intervals was repeated more than 3 times.

Results obtained in the described procedure were standardised for each respondent. Subsequently, a separate transitivity analysis was performed for each respondent's answer values. The analysis assumed that if interval X was judged to be more consonant than interval Y, as well as Y was evaluated as more consonant than Z, then X is also more consonant than Z. The X-Z difference value was computed as the mean distance between start of hypothetical vector XY and ending of vector YZ over all 2-dimensional spaces the two vectors can constitute.

Both initial and derived distance values served as input to create a dissimilarity matrix, filled to 90,4 % with mean judgements of distance between the intervals used in the experiment. The dissimilarity matrix was then subject to multidimensional scaling employing ALSCAL algorithm in interval mode.

V. RESULTS

2-dimensional solution was selected as Young's stress value did not change significantly between 2 and 3 dimensions. Total stress calculated for 2-dimensional solution by means of Kruskal's formula 1 was relatively high (see Table 3), however for a large number of variables it can still be considered acceptable, especially since stress values significantly decrease with falling number of variables (cf. Borg et al., 1997).

Table 3. Kruskal formula 1 stress values for selected various number of variables (i.e. number of intervals subject to multidimensional scaling procedure).

Number of intervals	Stress value
119	0,342
72	0,326
48	0,281

The 2-dimensional solution has been rotated, which helped in identifying the horizontal dimension as a possible representation of consonance sensation (Figure 7). The second (vertical) dimension is not clearly interpretable, however we can observe, that intervals traditionally clearly assigned to either consonance or dissonance group, are found in the upper half of the graph. On the contrary, consonance ratings of intervals in the lower half of the graph have been ambiguous and fluctuating in music history, e.g. fourth, minor third and sixth or major seventh.

Values of the horizontal dimension were subsequently used to evaluate the postulated sensory consonance degree, as well as selected other models' C/D merits. Intervals between 0 and 24 semitones were taken into account in the comparison. In the case of models which do not differentiate between intervals and their octave enhancements, respective C/D values were repeated in both octaves.

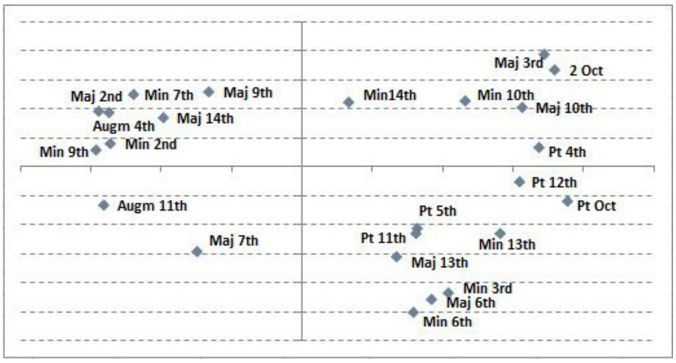

Figure 7. Multidimensional scaling results for selected group of 24 intervals within first 2 octaves. The horizontal dimension clearly corresponds to intuitive consonance rating.

Experimental results (horizontal dimension) were compared to both raw and logarithmically transformed C/D measures in selected models. The outcome reveals that the postulated model bears very strong correlation with experimental ratings (see Table 4). 79% per cent of the experimental consonance dimension's variability in first 2 octaves is explained by the logarithmically transformed postulated C/D degree. In both one and two-octave ranges the proposed C/D model achieves highest conformity with experimental data out of the selected known measures of consonance.

It is also evident that pitch-based consonance models (Hofmann-Engl, 2010, Ebeling, 2007, author's model) provide better agreement with experimental consonance ratings, than in the case of roughness-based models. This may indicate that indeed a pitch-related mechanism underlies perception of first-order sensory consonance, which is also the basic assumption of the model proposed herewith.

Table 4. R square (coefficient of determination) computed for 5 selected C/D models against experimental data from presented multidimensional scaling procedure. Asterisks denote maximum R square value for intervals within 1 octave and within 2 octaves. Calculations of first two models are based on (Mashinter, 2006).

Range	**1 octave**		**2 octaves**	
Measure	***'C/D'***	***ln('C/D')***	***'C/D'***	***ln('C/D')***
Kameoka and Kuriyagawa (1969)	0,34	0,34	0,30	0,31
Hutchinson and Knopoff (1978)/ Parncutt (1989)	0,56	0,54	0,59	0,53
Hofmann-Engl (2010)	0,75	0,74	0,75	0,76
Ebeling (2007)	0,46	0,69	0,43	0,59
Postulated model	0,65	0,81*	0,62	0,79*

VI. CONCLUSIONS

Experimental results have proven that the proposed simplified neural discharge model can be regarded as one of plausible hypotheses explaining C/D sensation of chords built by complex harmonic tones. This obviously does not constitute the model's verification or acceptance, as further research and more profound understanding of processes happening along the auditory pathway are necessary.

Nevertheless, the proposed consonance degree brings several interesting features which might make it useful for musicologists and composers even at such early development stage:

- It enables computations for virtually any interval, both microtonal and large-sized (far beyond octave);
- It can accommodate slight pitch deviations, as it is in the case of different tunings;
- Basing on working memory, it can also easily explain the phenomenon of melodic C/D;
- Set against a reference value (e.g. average consonance degree for all possible equally-tempered chords containing *n* components) it can serve as an analytical tool to follow flow of harmonic tension and formal progress of a musical piece;
- Employed in composition, it can be e.g. a means of more conscious control over the progress of harmony.

The model also raises further questions which need to be addressed by experimental verification, both on psychoacoustic and neurophysiological ground. The issues of interest include:

- Relation of simplified (modelled) and actual train spikes, including the mechanism of discerning key discharges with large probability from noise of all other firings;
- Notion of first-order sensory consonance, its plausibility and eventual underlying mechanism allowing for perception of C/D in chords consisting of harmonic complex tones, especially in view of apparent orthogonality of representations of pitch and timbre;
- Proving or disproving existence and usage of concentration windows;
- C/D perception mechanism for intervals in outermost registers; e.g. for the highest register the working hypothesis is that intervals exceeding 6 octaves converge to a similar C/D value;
- Verification of tolerance mechanism for slight pitch deviations;
- More accurate mathematical description of processes leading to C/D sensations, including use of biocomputational modelling of neural activity;
- Comparing proposed C/D degree with other psychoacoustic data, esp. in the case of chords with more than 2 components.

The above shortlist of next steps leads to the conclusion that the proposed model is still far from being considered as a fully tested and implementable solution. However, results presented in this article suggest we might be on the brink of establishing a sensory consonance model which does find confirmation in psychoacoustic data, and hence could become widely applicable in music-related disciplines.

ACKNOWLEDGMENT

The author's research has been inspired by content input from Prof. Gerald Langner (TU Darmstadt) and Dr. Martin Ebeling (P. Cornelius Conservatory of Music, Mainz). Furthermore, Prof. Andrzej Miśkiewicz (F. Chopin Music University) and Paweł Hendrich (Academy of Music, Wrocław) provided invaluable help in carrying out the experiment.

REFERENCES

Bahmer, A. (2007). *Computer simulation of chopper neurons: intrinsic oscillations and temporal processing in the auditory system.* Doctoral dissertation at TU Darmstadt, unpublished.

Borg, I. & Groenen P. (1997). *Modern Multidimensional Scaling: Theory and Applications.* New York: Springer.

Davis, K. & Hancock, K. & Delgutte, B. (2010). Computational Models of Inferior Colliculus Neurons. In R. Meddis & E. Lopez-Poveda & A. Popper & R. Fay (Eds.), *Computational models of the auditory system* (pp. 129-176). New York: Springer.

Ebeling, M. (2007). *Verschmelzung und neuronale Autokorrelation als Grundlage einer Konsonanztheorie*. Peter Lang: Frankfurt am Main.

Erlich, P. (1997). *On harmonic entropy.* Retrieved May 10, 2011 from http://sonic-arts.org/td/erlich/entropy.htm.

Hofmann-Engl, L. (2010). Consonance/dissonance – a historical perspective. In S. Demorest & S. Morrison & P. Campbell (Eds.) *Proceedings of the 11th International Conference on Music Perception and Cognition (ICMPC 11)* (pp. 852-856). Seattle, Washington, USA.

Hutchinson, W. & Knopoff, L. (1978). The Acoustic Component of Western Consonance. *Interface*. Vol. 7, No. 1, 1–29.

Kameoka, A. and Kuriyagawa, M. (1969). Consonance theory, part II: Consonance of complex tones and its computation method. *J Acoust Soc Am* 45, 6, 1460–1469.

Langner, G. & Schreiner, C. (1988). Periodicity coding in the inferior colliculus of the cat: I neuronal mechanism. *J Neurophysiol* 60, 1799–1822.

Langner, G. (2007). Die zeitliche Verarbeitung periodischer Signale im Hörsystem: Neuronale Repräsentation von Tonhöhe, Klang und Harmonizität. *Zeitschrift Audiologie* 46 (1), 8–21.

Licklider, J.C.R. (1951). A Duplex Theory of Pitch Perception. *Experientia* VII/4, 128-134.

Mashinter, K. (2006). Calculating Sensory Dissonance: Some Discrepancies Arising from the Models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff. *Empirical Musicology Review* 1, 2, s. 65-84.

Meddis, R. & Hewitt, M. (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. *J Acoust Soc Am* 89, 6, 2866–2882.

Nelson, P. & Carney, L. (2004). A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *J Acoust Soc Am* 116, 2173–2186.

Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach*. Berlin: Springer-Verlag.

Parncutt, R. (2006). Commentary on Keith Mashinter's "Calculating Sensory Dissonance: Some Discrepancies Arising from the Models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff". *Empirical Musicology Review* 1, 4, 201-203.

Patterson, R. (1987). A pulse ribbon model of monaural phase perception. *J Acoust Soc Am* 82, 1560-1586.

Plomp, R. & Levelt, W. (1965). Tonal consonance and critical bandwidth. *J Acoust Soc Am* 38, 548-560.

Sethares, W. (1999). *Tuning, Timbre, Spectrum, Scale*. London: Springer.

Shannon, C. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal* 27, s. 379–423, 623-656.

Sierpiński, W. (1950). *Teoria liczb*. Warszawa: PWN.

Tenney, J. (1988). *A History of Consonance and Dissonance*. New York: Excelsior Music Publishing Company.

Terhardt, E. (1974). Pitch, consonance, and harmony. *J Acoust Soc Am* 55, 1061-1069.

Tramo, M. & Cariani, P. & Delgutte, B. & Braida, L. (2001). Neurobiology of harmony perception. In I. Peretz & R. Zatorre (Eds.) *The Cognitive Neuroscience of Music* (pp. 127-151), New York: Oxford University Press.

Van de Geer, J. & Levelt, W. & Plomp R. (1962) The connotation of musical consonance. *Acta Psychologia* 20, 308–319.

Voigt, W. (1985). *Dissonanz und Klangfarbe: Instrumentations-geschichtliche und experimentelle Untersuchungen.* Bonn: Verlag für systematische Musikwissenschaft.