The Self-Organization of Stable Perceptual Maps in a Realistic Musical Environment *

Marc Leman, IPEM-Univ. Ghent Blandijnberg 2, B-9000 Ghent, Belgium, Marc.Leman@rug.ac.be
Francesco Carreras, CNUCE/CNR, Via S. Maria 36, I-56126 Pisa, Italy, F.Carreras@cnuce.cnr.it

Abstract

Abstract: A *schema theory* of music cognition has been developed and a computer model has been designed to test basic assumptions in the domain of tone center perception. Thus far, however, the simulations of perceptual learning have been based on limited data-sets such as a set of chords, a series of cadences, or a short piece of music. This paper discusses the results of using a large and more realistic musical environment: Book 1 of Bach's *Das Wohltemperierte Clavier* in a recording by K. Gilbert (CD ARCHIV 413439-2, 1984).

1 Introduction

In the past, music cognition has often been studied with concepts borrowed from linguistics [Monelle, 1992]. Although this approach has contributed to theories of meaning, it turns out that certain aspects of cognitive musical information processing are more related to auditory imagination and topographical representation rather than symbolic conceptualization [McAdams, 1993, Leman, 1995b]. An in-depth analysis, based on music theory, computer modelling, experimental psychology and brain research is necessary to test this hypothesis on empirical grounds.

In this paper, we restrict ourselves to computer simulation but the approach should be considered as part of a larger (interdisciplinary) project which aims to map out the memory structures that underlay music perception and cognition. As a study in computer modelling, the present paper aims show that memories for supposed "high level" cognitive musical functions (such as tone center perception) may be formed by a relatively simple adaptation to the dynamical and statistical structures that are inherent in the sounds of a musical environment. The approach is based on ecological modelling principles. It entails a training of the computer program under real-world conditions and a testing of the program with different stimuli. One of the aims is to use audio stimuli that are identical to the stimuli used in experimental behavioral and brain studies. The idea is to compare the results of the simulations with data from experimental psychology and brain research.

The structure of the paper is as follows: in the first part, we give an overview of the computer model for tone center perception. Then we go deeper into the analysis of perceptual maps of tone center perception. Finally, the output of the model is compared with data of an experimental research. The paper closes with a discussion of the relevance of this work to our understanding of music perception (from both an information processing as neurophysiological point of view).

*Paper presented at JIM96: Journées d'Informatique Musicale 16-18 mai 1996, Caen

2 Modelling Perceptual Learning of Music

Studies which aim to simulate perceptual learning of music have centered on tone center perception [Leman, 1995b] and timbre classification [Toiviainen, 1996, Cosi et al., 1994]. The results of these studies show that an artificial memory, which is exposed to musically relevant patterns, can build up a structure which holds invariant information from the patterns. The structure or schema [Arbib, 1995] is typically considered to be build up by long-term data-driven learning (based on self-organizing properties of the artificial memory) and it may act as an active pool of attractors during short-term (schema-driven) perception [Leman, 1995a]. Up to now, however, most of the simulations of perceptual learning have been based on rather limited data-sets: a set of chords, a set of timbres of musical instruments, or a short piece of music. In the present paper, we report results about extending the simulation towards more realistic proportions using real music played by a real musician on a real instrument. Pilot studies [Carreras & Leman, 1996] have shown that cognitive modelling at this level requires super-computer power in order to obtain results within a reasonable amount of time.

3 Basic Assumptions

Our approach is based on three assumptions about the nature of music perception, representations and the modelling attitude:

1. Music perception is to a large extend: (i) context-dependent (the functional perception of the tones/timbres/rhythms depends on previous tones/timbres/rhythms), (ii) non-denotational (but semantically connotational), (iii) rather unconscious than conscious, and (iv) learned and depending on cultural factors.
2. Perceptual representations are conceived of in terms of *auditory images* rather than *symbols*. Auditory images *reflect* properties of the musical environment and are ultimately (causally) based on the physical characteristics inherent in the sound of that environment. Different types and levels of images are considered (Cfr. Chap. 4 [Leman, 1995b]). Symbols, on the other hand, *point* to the objects in the world without reflecting any of the physical properties of the object.
3. The programmer/modeler incorporates the physical/musical environment into the computer and he restricts his role to defining the conditions on which the interactions between the physical/musical environment on the one hand and the model of human information processing on the other hand can take place. This *ecological* approach differs from symbol-based and connectionist approaches in which the entities of the model and relationships between them are pre-defined (hence treated as symbols).

As a consequence, the nature of the musical environment needs to be defined in terms of the physical properties of the sound (not in terms of symbolic representations thereof).

4 Basic Architecture

The framework (called Modelling Environment) thus includes the description of a physical (sound) environment, a information system which simulates perceptual learning, and a musical environment.

4.1 The Physical Environment

Due to standards in audio recording and general knowledge of acoustics we can be very short about the physical environment in which our simulation of perceptual learning takes place.

The acoustical representation is taken from a CD recording, from which we take a digitized waveform using a sampling rate of 20000 sa/s and a resolution of 16 bits.

4.2 The Perceptual Learning System

The computer model of perceptual learning contains two parts: (1) a perception module consisting of an acoustical front end based on an auditory model [VanImmerseel & Martens, 1992] and a pitch extraction part based on a periodicity analysis of neural discharge patterns in the auditory channels of that model [Leman, 1995b], (2) a cognitive model which accounts for *long-term data-driven* perceptual learning and *short-term schema-driven* recognition. Perceptual learning yields a stable schema of tone center perception. Recognition yields an analysis of a piece of music in terms of the stable points that were obtained by self-organization.

- The perception module.

  The acoustical front end of the model has been developed by L. Van Immerseel and J.-P. Martens [VanImmerseel & Martens, 1992]. A pitch extraction part has been added based on a periodicity analysis of neural discharge patterns in the auditory channels of that model [Leman, 1995b]. The acoustical front end implements outer and inner ear filtering and cochlear hydro-dynamical filtering by means of a bank of 20 asymmetric overlapping bandpass filters. The outputs of these filters are transformed into neural firing patterns using half-wave rectification and dynamic range compressing, short-term adaptation and synchrony reduction. The neural firing patterns represent the probability of firing during a defined interval of 0.4 ms. Pitch-shift is not taken into account. For each neural firing pattern (or channel), a periodicity analysis is done by means of a short-term-autocorrelation analysis of 30 ms frames at a rate of 10 ms. The *completion image* (which is similar to the virtual pitch pattern) is obtained by summing the short-term-autocorrelation results of a frame over the 20 channels. The leaky-integrated result of completion images over time is called the *tone context image*. This auditory image gives an account of the temporal dependencies among musical patterns over a time period of about 3 seconds.

- The cognition module.

  The cognitive model accounts for *data-driven* perceptual learning and it is based on the Self-Organizing Map (SOM)-architecture [Kohonen, 1984]. This learning by self-organization is realized at long-term and it yields a schema of tone center perception in terms of a neural network. SOM maps the 56-dimensional *tone context images* onto a two-dimensional grid of artificial neurons but it keeps the topological order of the implied relations that are imbedded in the auditory images. The set of auditory images is thereby mapped onto a limited set of neurons by means of a process of generalization (at the cost of specificity). The mapping realizes a clustering of all images. Multiple auditory images are necessarily compressed onto single representational units with the result that specificity is lost in favour of generality. However, the ability to generalize is a central property of perceptual learning and it is exactly this effect which we want to explore in our studies of tone center perception. In other words the simulation builds up a set of clusters, each representing a tone center, whose reciprocal positions on the map reflect the topological relations of the chords.

4.3 The Musical Environment

The musical environment is here conceived of in terms of the set of cultural constraints imposed upon the acoustical environment. For this study we choose Book I of J.S. Bach's *Das Wohltemperierte Clavier* played by Kenneth Gilbert in a recording which is available on CD ARCHIV 413439-2. The piece contains 24 preludes and fugues in all tonalities. The total duration is 1 hour 50 minutes and 22 seconds. The instrument used in this recording is a harpsichord built by Jan Couchet in Antwerpen, dated 1671. It is tuned in the old pitch of LA=392 Hz.

5 Implementation

The SOM-algorithm can be parallellized in a straightforward way. To speed up computation in 1987, the program was implemented on a Transputer system, first on a four transputer (T800) parallel system using 3L-C and master-slave architecture, later using EXPRESS [Leman & Renterghem, 1989]. For larger simulations, the model has then been ported with minor modifications onto a nCUBE2 using 8 processors (in 1994). It turned out, however, that the esteemed computing time required on the nCUBE system is beyond any reasonable limit for fluent work in cognitive science. A complete revision of the model was necessary in order to optimize the time consuming parts of the code, to reduce the impact of interprocessor communication and to limit disk I/O activity by using large block operations.

In 1995, the model has been implemented on a IBM SP2 which is available at CNUCE CNR Pisa. The system has eight processors, each with 512 Mbytes of memory. The ratio of peak performance between the IBM SP2 processor (266 Mflops peak), based on the POWER2 microprocessor technology, and the nCUBE2 processor is about 113.

When cognitive theories are tested in a more realistic environment, then the amount of data is huge and the computational efficiency of the model becomes a critical factor in research. Parallelism can be exploited to deal with a realistic input environment [Carreras & Leman, 1996].

6 The Analysis of Self-Organizing Maps

Different techniques for the analysis of the self-organizing maps are available. They include the use of reference patterns, visualization and clustering methods. The general setup and the problem of analysing the obtained perceptual maps is first discussed, then a visualisation method is tested. The latter looks promising in limited data sets but the boundary maps does not seem to emerge straightforwardly in perceptual learning simulations under realistic conditions.

6.1 Using Shepard-Patterns for Testing a 30x30 Map

The SOM maps the 56-dimensional auditory images onto a two-dimensional grid but it keeps the high dimensional topological order of the implied relations that are imbedded in the auditory images. In a number of pilot studies the network was first trained and then tested with Shepard-patterns (processed with the same auditory model). Shepard-patterns are made up of tones which contain spectral components at one octave distance from each other (over the whole frequency range). The spectral energy is shaped according to a bell-like curve which promotes the energy in the region between 500 Hz and 1500 Hz. The test-patterns, which stand for tone center images, are obtained by applying the leaky integration technique to cadences of Shepard-chords (see [Leman, 1995b, Leman, 1994] for details).

In the tests it was shown that the SOM may come up with a structure in which the patterns occurs within a global *circle of fifths*. In the present simulations, we use Preludes of Book I of Bach's *Das Wohltemperierte Clavier* during the training of the map and we use the Shepard-chords for the testing the map. The tests are thus based on patterns which have a spectrum which is quite different from the harpsichord sounds that are used in the training set.

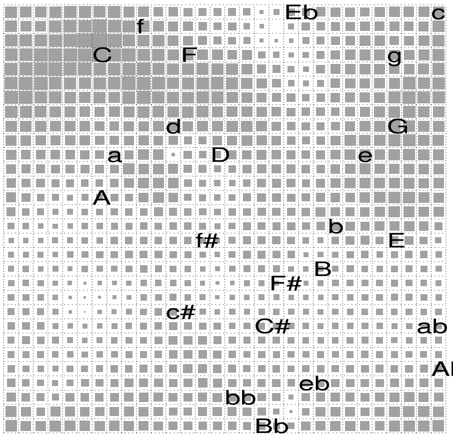


Fig. 1. Topological Map obtained by training using all Preludes of Book I of Bach's *Das Wohltemperierte Clavier*. The map has been trained with harpsichord-patterns from Book I, the test patterns are Shepard-patterns that stand for tone centers (labeled C, C♯, D ..., b♭, b).

Figure 1 shows the way in which these test patterns are mapped onto a trained network. The figure shows a 30 x 30 neurons network on a torus surface (left and right border connect as well as the upper and lower borders). Although the test patterns are quite different from the training data, it is possible to map out a global ordering in terms of circles of fifths. The little boxes stand for neurons, the activation value of the neurons is represented by grey boxes: large boxes indicate high activation, while small boxes indicate low activation. The labels on the map show the places where the corresponding Shepard-patterns generate the highest response. As this map shows the response to C by means of the grey boxes, the label C is on the neuron with the highest grey value. The topological order must be deduced also on the basis of response-regions (grey values) evoked by the test patterns. In general, this structure confirms the earlier studies that used limited data-sets, nl. that the mapping of Shepard-patterns (standing for tone centers) obeys a global topological structure in terms of the circle of fifths.

The simulation, using a realistic musical environment as input, confirms the assumptions of perceptual learning of the schema theory. The self-organization principles, simulated by means of artificial neural network, can be regarded as a model of the self-organization capabilities of the human brain.

However, although the test with reference patterns gives us an idea of the global ordering of the network, there are some further problems:

- First of all, the approach does not give information about the statistical properties of the obtained clusters. It is known, nevertheless, that the SOM may reflect the probability distributions of the inherent classes.
- Secondly, although the reference patterns are mapped onto a particular location of the network, they may not guarantee that these locations are most characteristic for the generalizations that emerge from the environment.
- The Shepard-patterns are somehow artificial and their timbre is highly different from the sounds of the CD-record. It may be that the network develops less specific characteristic features with respect to tone center.
- A 30x30 network may be rather small. Due to the amount of data it is realistic to assume that better results can be obtained with a larger network, e.g. a 100x100 network.

6.2 Using Boundary Maps

In this section, we report our results on using boundary maps as an alternative for the analysis of the topological structure of a map. Interestingly, this method does not rely on test patterns but on a direct analysis of the structure of the map. Kraaijveld et al. [Kraaijveld et al., 1992] have proposed a non-linear projection method to visualize high-dimensional data as a two-dimensional image. The first step is based on the SOM-algorithm, then the interpoint distances in the feature space between the units in the network are graphically displayed to show the underlying structure of the data. We have adopted this method to map out the underlying structure of SOM.

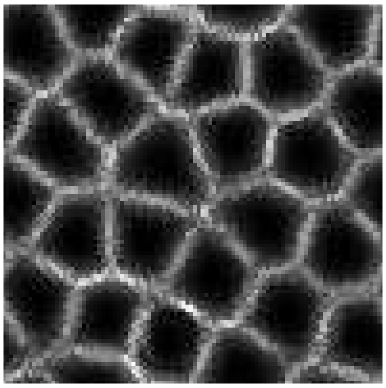


Fig. 2. Boundary map obtained by training a 100x100 SOM with 24 tone center images (using a Shepard-spectrum). The boundary map shows the interpoint distances of the neurons in the map. The white contours show the boundaries of the clusters which, in this case, stand for the 24 images.

A simulation of perceptual learning with the above mentioned Shepard-images indeed shows that the method is capable of visualizing some interesting structures of the map. Figure 2 shows the boundary map after training with 24 Shepard-patterns. The network contains 100x100 neurons on a torus structure. During the training the patterns were presented several times in random order. Obviously, there are quite many neurons compared to the small number of inputs. The generalizing properties of the network are indeed not exploited here because the simulation rather pertains to the ordering of patterns which are clearly visible in terms of a web or maze akin to honeycombs. The highest responses to the Shepard-patterns are typically located in the middle of the spheres that form the honeycomb (not shown in the figure). The labels appear along two parallel (one for the major and one for the minor) spirals in the torus surface, with minor and major tone centers of the same name very close, and tone centers harmonically distant positioned far away.

6.3 The Search for Boundary Maps

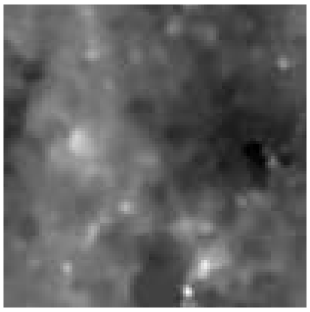
Maps that are trained with the Preludes of Book I instead of the Shepard-patterns do not show a very clear boundary map – as far as our simulations show. In the search for ordered boundary maps, several parameters of the model have been changed. The following section gives a rather general overview of the parameters that have been taken into consideration. A more detailed report is in preparation.

Recall that in the SOM-algorithm [Kohonen, 1995], the following parameters are important:

- neighborhood radius: this is the radius which defines the set of neurons that is adapted during the learning phase. It is recommended that this radius is large in the beginning and gradually decreases when learning proceeds.
- learning rate: this is the rate which defines the speed of adaptation. It is recommended that the speed of adaptation is fast in the beginning, but then slows down as learning proceeds.
- similarity measure: many SOM-architectures rely on the calculation of the Euclidian distance between the input patterns and the neurons in the network. However, it may

happen that other measures are more appropriate, depending on the type of data that are used in the training.

- starting situation of the map: the initial weights of a network are normalized. However, it is possible to bias the weights by starting from a predefined structure.



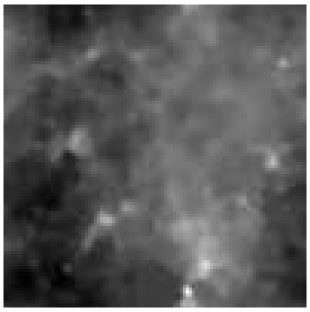


Fig. 3. The response of the network to the Shepard-patterns which stand for the tone center of C-major (figure left) and C♯-major (figure right) after training with the Preludes of Book I of Back's *Das Wohltemperierte Clavier*. The figures show that the network can differentiate between both patterns. The network contains 100x100 neurons.

The simulations are based on a SOM architecture of 100 by 100 neurons. In most of the simulations, the neighborhood radius goes from 55 to 1 according to an exponential distribution of images for each radius value. One third of the total amount of processed images is processed with radius one. The total amount of all images is the set of images taken from the auditory model, multiplied by a number of epochs. The learning rate was made variable from .9 to 0.02 along a number of images given as a parameter. Different choices for the number of epochs and the number of images with variable learning were tested. Given the large amount of data it is reasonable to assume some redundancy in the auditory images. Therefore, some simulations have used only one epoch.

After testing these parameters, the focus was then shifted to test other similarity measures. The fifth order Minkowski metric was first adopted and secondly a metric that took into consideration the differentials of pairs of corresponding components of two images.

In all the previous tests, the weights of the synapses of the neurons were given random values. A simulation was therefore designed to test the impact of biased weights. To this end, a network has been trained that started with the structure of the boundary map (Fig. 2).

Although a detailed analysis of these simulations is currently underway, the results show – with some differences due to the different approaches – a tendency of the network to organize itself into a global order. This order was mapped using the method of reference patterns. The direct method, based on boundary maps, however, did not lead to very clear results. Figure 4 shows the boundary map that is obtained from the map used in Fig. 3a-b.

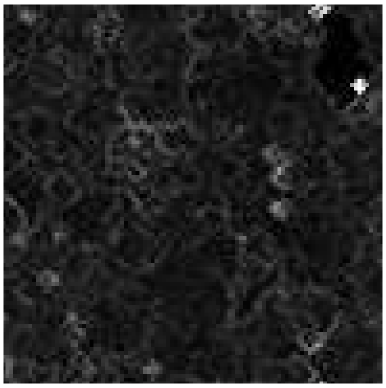


Fig. 4. Boundary map corresponding to the underlying structure of the map used in Fig. 3a-b.

The context images present a large variety of configurations and the classes which are inherent in the images may be too much overlapping to allow a clear separation of the inherent clustering. The context images obtained from the perception module indeed reflect all the complexities of the interactions among the vibrating strings of the harpsichord when more keys are pressed in a very rapid sequence. Further analysis, using clustering analysis may provide more information about the underlaying structure. Moreover it was noticed that the map contains some spots in which the responses to different patterns is indifferent. For example, on Fig. 3, such spots are seen in the lower region and in the higher right quadrant. It is unclear where these spots come from, but they may have something to do with deformations during the preprocessing phase. At the beginning of each piece, the auditory model (including integration process) constructs some patterns which, when normalized, give a rather flat spectrum.

6.4 Using the Model for Psychological Tests

The present section shows that although the direct analysis of a map may not be completely satisfactory, some responses of the map are nevertheless significant within a comparative analysis. Significant correlation with psychological data has been obtained even when the topological structure looks somewhat complex.

6.5 The Inherent Structure of the Map

Figures 3a-b show the response of the network to the Shepard-pattern which corresponds with the tone center of C and C♯. The pictures show the response of the network to the Shepard-patterns that stand for the tone center of C-major and C♯-major. The network responds to these tone centers in a different way. One method of analysis, used above to map out the topological representation of tone centers, is to find the location with the highest response (here represented by the black areas) and label it accordingly (See also Fig. 1).

The evidence for topology, however, may not be very straightforward at first sight. For example, the tone center9 of G is apparently closer to F♯ than it is to C, although F♯ is further away on the circle of fifths. In a similar way, D should be closer to A than to B. Obviously, the organization of the map may be such that the borders between G and F♯ are bigger than between G and C. It is reasonable to interpret some of the emerging lines of the boundary map of Fig. 4 in that way, but further work is needed to obtain a more clear picture.

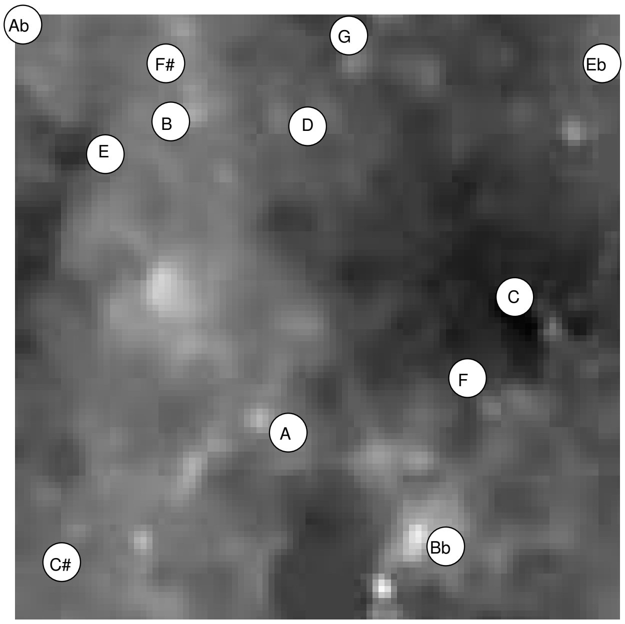


Fig. 4. CN-map. The map shows the response to the Shepard-pattern that stands for the tone center of C-major, but in addition, the labels of all major tone centers are mapped. As shown, C is mapped on the location with the black spot.

In fact, it is quite obvious from Fig. 4 that G is more akin to C than, e.g. B♭, although B♭ may be closer in distance. The dark cloud seems to overlap G, whereas B♭ seems to occur in region which is more white and thus less activated. The circle of fifths thus appears in bands over the map. Starting with C, we go to F, B♭, E♭ – which is in the top right quadrant, but which is actually very close because of the torus structure. E♭ is also very close to A♭ and then by C♯ (lower left quadrant) we go to F♯, B, E and then make a big jump to A, D, G and again a rather big jump back to C.

The underlying response structure can be mapped out by a systematic comparison of the responses of the network. Figure 5a shows a graph in which the responses to all tone centers are compared with the response to the tone center of C-major. The correlation coefficient is used to calculate the degree of overlapping of the response clouds. Figure 5a shows a graph in which the responses to all tone centers are compared with the response to the tone center of C-minor. The dotted curve are the computer results, the full curve are the data obtained by Krumhansl [Krumhansl, 1990]. The correlation between both curves is 0.94 for the C-major curve and 0.97 for the C-minor curve. This shows a very significant correlation between the response structure of the computer model and the data obtained by psychological experiments.

7 Using the Model for Psychological Tests

The above analysis aims to clarify the structure of the schema by using reference patterns and direct analysis by means of boundary maps. Given the adopted ecological approach to modelling, in which a computer program is exposed to a sound environment, it is straightforward to consider this program on similar terms as an information system which can undergo a psychological test. This section explores the idea to test the model with musical sequences that are used in psychological experiments with humans and to compare results directly with these data. Although another report is underway to describe the results in detail, we show here some preliminary results. In order to be sure about the input stimuli, we repeated the experiments done by Krumhansl [Krumhansl, 1990] and Bruhn [Bruhn, 1988]. Listeners were asked to rate the similarity between an open cadence (degrees I IV V in C-major) and a chord which follows the cadence. The open cadence was held fixed, while the chords at the end were variated. In this experiment, this variation was limited to all major triads.

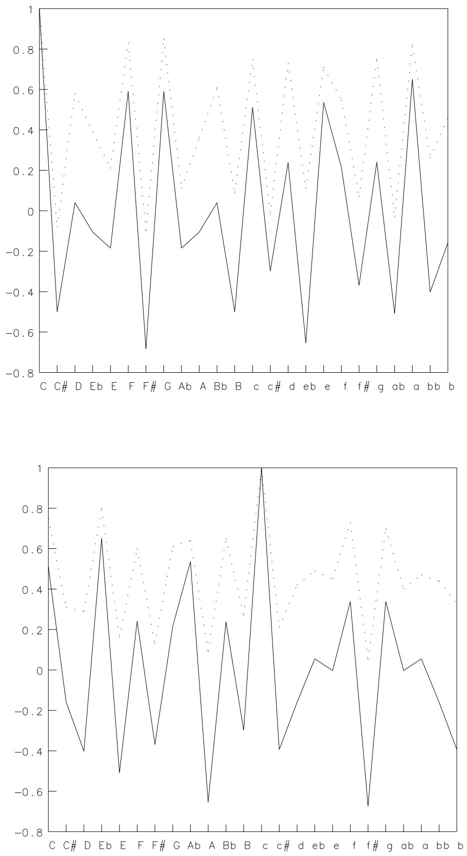


Fig. 5 a-b. The correlations between the responses of the map are shown by the dotted curve. The full curve is based on data provided by [Krumhansl, 1990]. In the top figure, responses to tone centers are compared to the response to C. In the low figure, responses are compared to the response to c.

The resulting 12 sequences were based on Shepard-tones. These sequences are exactly the same stimuli which we used in the computer simulation. Normally, a chord sequence follows a trajectory on the map. Up to the end of the cadence, however, the trajectory is the same for all chord sequences. To simplify matters, we took the response of the map at the end of the open cadence and compared this with the response of the map at the end of the complete sequence. The resulting profile is shown in Fig. 6. Twenty s11ubjects were divided in three groups: experts (with musical training), novices (without any training) and subjects with an initial training. The profile of the computer simulation is also shown. The comparison is shown in Table 1. Our own experiments are here compared with the results on similar stimuli obtained by Krumhansl and Bruhn. The simulation shows the highest correlation with the group of experts and is in good agreement with the results by Bruhn and Krumhansl. The mean (exp.int,nov) is based on taking the mean of the profiles.

8 Discussion

With the publication of the Handbook for Neural Networks and Brain Research [Arbib, 1995] it becomes clear that the new tendencies in cognitive science are directed towards the integration and empirical corroboration of computer models. The idea of a convergence between results in computer modelling, physiological acoustics (psychoacoustics), experimental psychology and brain research is in line with this tendency.

Recent research in neurophysiology and neurpsychology indeed shows that relevant information about high level musical information processing can be obtained by studies of the brain. Articles such as [Schreiner & Langner, 1988a, Schreiner & Merzenich, 1988] provide neurophysiological evidence that a number of nuclei in the brainstem up to the cortex are organized in a topographic way, reflecting the tonotopic properties of the cochlea.

Research by C.E. Schreiner and G. Langner [Schreiner & Langner, 1988a] [Langner & Schreiner, 1988, Schreiner & Langner, 1988b, Langner, 1983] provides evidence that the inferior colliculus (IC) transforms temporal signal aspects in the residual pitch range into a spatial representation of periodicity-tuned units. The method, based on multiple electrode recordings, cannot be used for research on human perception but the results show, however, that the neural representation of the pitch is consistent with the perception of "the missing fundamental" or "virtual pitch" – a basic concept in the theory of tonal perception.

Table 1. Correlations of mean harmonic hierarchies for major ˝tritones in the expert, intermediate and novice groups, compared with findings by Bruhn, 1988, Krumhansl ˝1990, and a simulation in the 100x100 neural net.

	simulation	Bruhn, 1988	Krumhansl,1990
exp	0,843659*	0,904935*	0,760091*
int	0,714664*	0,524316	0,626279*
nov	0,693066*	0,587222*	0,669436*
mean (exp, int, nov)	0,807235*	0,729814*	0,738648*
simulation	1	0,856515*	0,619846*
Bruhn, 1988	0,856515*	1	0,696953*
Krumhansl, 1990	0,619846*	0,696953*	1

*Significant at $p < .05$.

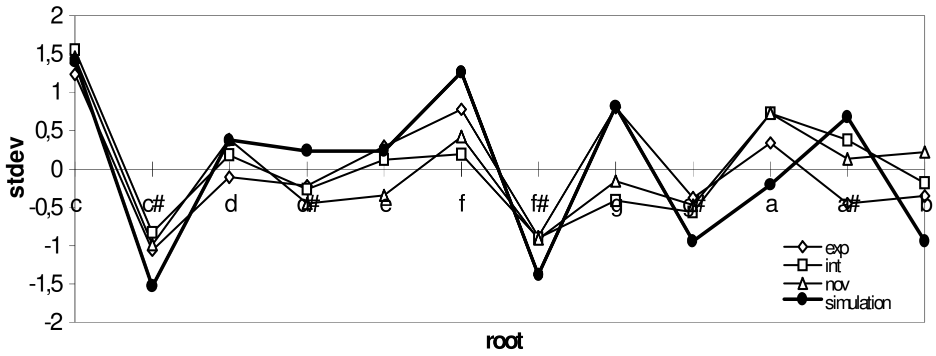


Fig. 6. Profile from experimental tests with musical experts (exp), novices (nov), subjects with initial training in music (int) and the computer (simulation).

A number of studies in cognitive neuropsychology and neurophysiology, moreover, show that non-invasive (and invasive) techniques, based on electroencephalogram (EEG) may provide interesting information about higher-level music processing. The EEG measures the collective electrical behavior of cortical neurons and of the various sensory systems, the components of the auditory evoked potential have been most extensively studied [Martin, 1991]. Musical relevant studies have been based on long latencies potentials that are generated by the neurons in the auditory cortex [Cohen & Erez, 1991, Janata & Petsche, 1993] [Tervaniemi et al., 1993]. Of particular relevance to the present work is P. Janata [Janata, 1995], who studied expectancy violations produced by a chord sequences by means of ERP-measurement. The area and peak amplitude of the component P300 thereby varied as a function of the degree of expectancy violation. The results of this study call for a larger and deeper analysis using source-localization, more subjects (making the distinction between musical trained and non-musical trained) and with more variation in the parameters of expectancy violation.

The above results show that the simulation of music perception, carried out in a realistic environment, show significant correlations with data from psychology. The model, based on physiological acoustics and some basic principles of self-organization, generates some highly interesting hypothesis about the way in which the brain may process tonal relevant information. The hypothesis is that some center in the temporal cortex of Western listeners may show a similar functional organization of neurons. The model provides evidence for the fact that the representation of tone center is at a level which is quite close to sensorial cortical structures. Neurological studies will be necessary to falsify this hypothesis.

9 Conclusion

Metaphors of dynamical processes can be cast in terms of mathematical models whose complex systems dynamics can be studied by means of computer implementations in the domain of music cognition. The claim that the input/output of computer simulations should converge with the input/output data from experimental psychology and brain research studies is central in our approach to modelling. The present work is seen as a step towards the accomplishment of that paradigm. The role of computer modelling is seen as a provider of working hypothesis and ideas for empirical tests. In return, brain research and experimental psychology may provide the ultimate foundation for issues in music perception, music theory, and perhaps also music composition.

In view of this so-called *Convergence Paradigm*, the framework used for modelling is based on an ecological approach. The approach aims to integrate psychoacoustics and Gestalt theory within a single framework for sensorial, perceptive and cognitive information processing.

Acknowledgement

Part of this work has been supported by the Belgium National Fond for Scientific Research (NFWO) and by the National Research Council of Italy (CNR). The psychological tests have been realized by B. Herreman at IPEM.

References

[Arbib, 1995] Arbib, M. (1995). Schema theory. In M. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA: The MIT Press.

[Bruhn, 1988] Bruhn, H. (1988). *Harmonielehre als Grammatik der Musik*. München: Psychologie Verlags Union.

[Carreras & Leman, 1996] Carreras, F. & Leman, M. (1996). Distributed parallel architectures for the simulation of cognitive models in a realistic environment. In E. DHollander, G. Joubert, F. Peters, & D. Trystram (Eds.), *Parallel Computing: State-of-the Art Perspective – Proceedings of the International Conference PARCO95, Gent, 19-22 September 1995*. Amsterdam: Elsevier.

[Cohen & Erez, 1991] Cohen, D. & Erez, A. (1991). Event-related-potential measurements of cognitive components in response to pitch patterns. *Music Perception*, 8(4), 405–430.

[Cosi et al., 1994] Cosi, P., Poli, G. D., & Lauzzana, G. (1994). Auditory modelling and self-organizing neural networks for timbre classification. *Journal of New Music Research*, 23(1), 71–98.

[Janata, 1995] Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, 7(2), 153–164.

[Janata & Petsche, 1993] Janata, P. & Petsche, H. (1993). Spectral analysis of the EEG as a tool for evaluating expectancy violations of musical contexts. *Music Perception*, 10(3), 281–304.

[Kohonen, 1984] Kohonen, T. (1984). *Self-organization and associative memory*. Berlin: Springer-Verlag.

[Kohonen, 1995] Kohonen, T. (1995). *Self-organizing Maps*. Berlin, Heidelberg: Springer-Verlag.

[Kraaijveld et al., 1992] Kraaijveld, M., Mao, J., & Jain, A. (1992). A non-linear projection method based on kohonen's topology preserving maps. In *Proc. 11 ICPR, Int. Conf. on Pattern Recognition* (pp. 41–45). Los Alamitos, CA: IEEE Comput. Soc. Press.

[Krumhansl, 1990] Krumhansl, C. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

[Langner, 1983] Langner, G. (1983). Neuronal mechanisms for a periodicity analysis in the time domain. In R. Klinke & R. Hartmann (Eds.), *Hearing - Physiological bases and psychophysics*. Berlin: Springer-Verlag.

[Langner & Schreiner, 1988] Langner, G. & Schreiner, C. E. (1988). Periodicity coding in the inferior colliculus of the cat. Part I. Neuronal mechanisms. *Journal of Neurophysiology*, 60(6), 1799–1822.

[Leman, 1994] Leman, M. (1994). Schema-based tone center recognition of musical signals. *Journal of New Music Research*, 23(2), 169–204.

[Leman, 1995a] Leman, M. (1995a). A model of retroactive tone center perception. *Music Perception*, 12(4), 439–471.

[Leman, 1995b] Leman, M. (1995b). *Music and Schema Theory - Cognitive Foundations of Systematic Musicology*. Berlin, Heidelberg: Springer-Verlag.

[Leman & Renterghem, 1989] Leman, M. & Renterghem, P. V. (1989). Transputer implementation of the Kohonen feature map for a music recognition task. In *Proceedings of the Second International Transputer Conference: Transputers for Industrial Applications II* Antwerpen: BIRA, Belgian Institute for Automatic Control.

[Martin, 1991] Martin, J. (1991). The collective electrical behavior of cortical neurons: The electroencephalogram and the mechanisms of epilepsy. In E. Kandel, J. Schwartz, & T. Jessell (Eds.), *Principles of Neural Science (Third Edition)*. Amsterdam: Elsevier.

[McAdams, 1993] McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), *Thinking in Sound–the Cognitive Psychology of Human Audition*. Oxford: Oxford University Press.

[Monelle, 1992] Monelle, R. (1992). *Linguistics in Semiotics and Music*. Harwood Academic Publ.

[Schreiner & Langner, 1988a] Schreiner, C. E. & Langner, G. (1988a). Coding of temporal patterns in the central auditory nervous system. In G. M. Edelman, W. Gall, & W. Cowan (Eds.), *Auditory function: neurobiological bases of hearing*. New York: John Wiley and Sons.

[Schreiner & Langner, 1988b] Schreiner, C. E. & Langner, G. (1988b). Periodicity coding in the inferior colliculus of the cat. ii. topographical organization. *Journal of Neurophysiology*, 60(6), 1823–1840.

[Schreiner & Merzenich, 1988] Schreiner, C. E. & Merzenich, M. M. (1988). Elements of signal coding in the auditory nervous system. In W. von Seelen, G. Shaw, & U. M. Leinhos (Eds.), *Organization of neural networks: structures and models*. Weinheim: VCH Verlagsgesellschaft.

[Tervaniemi et al., 1993] Tervaniemi, M., Alho, K., Paavilainen, P., Sams, M., & Näätänen, R. (1993). Absolute pitch and event-related brain potentials. *Music Perception*, 10(3), 305–316.

[Toiviainen, 1996] Toiviainen, P. (1996). Timbre maps, auditory images, and distance metrics. *Journal of New Music Research*, 25(1), 1–30.

[VanImmerseel & Martens, 1992] VanImmerseel, L. & Martens, J. P. (1992). Pitch and voiced/unvoiced determination with an auditory model. *The Journal of the Acoustical Society of America*, 91(6), 3511–3526.