NEUROBIOLOGICAL STATISTICS FOR INTERPRETING PROBLEMS OF MUSICAL CONSONANCE PERCEPTION

Jobst P. Fricke

Musicological Institute
University of Cologne, Germany
`alm01@uni-koeln.de`

CONTENT

ABSTRACT

In 1951 Licklider presented a model for pitch perception, based on a periodicity analysis of neural spikes [13]. This process of pitch extraction is similar to an autocorrelation analysis. Langner and Schreiner could prove in 1988 that a periodicity analysis in the auditory organ exists and that the periodicity pitch is neurally represented in the Inferior Colliculus and Cortex [12]. Their dimension is independent of the tonotopic representation and runs in about orthogonal to it. The periodicity of sound signals of consonant intervals could be neuronally proven by means of a periodicity analysis too [15]. The periodicity of the acoustic signals however is imperfect in music performance. Intonation deviations, which are a disturbance of the periodicity, are tolerated in the hearing process to a considerable extent. This can be seen particularly in the judgement of consonant intervals. Depending on the musical context, standard deviations of 13 cents for the optimal intonation were measured for the fourth as well as for the fifth. In total the variation was even at 70 cents. For those on this scale experimentally determined hearing tolerances, the statistical processes of neural coding and processing, in particular the neural integration for the autocorrelation, seem to be responsible.

1. POSSIBILITIES OF INFORMATION CODING

1.1   Local coding

An example for "place representation" (local coding) is given by the so called homunculus, where the neurotopic representation of sensory information in the brain is shown.

The other sensory modalities are organized spatially too. Neighborly position is responsible for similarity relation.
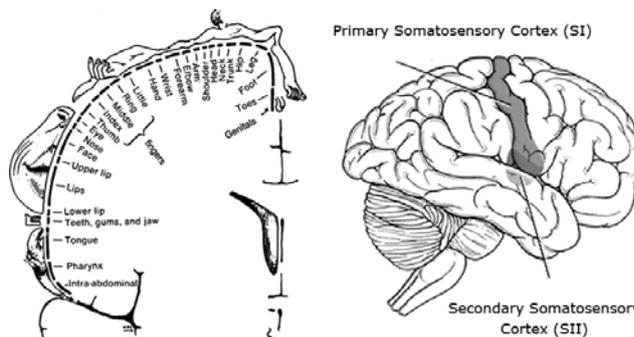

Figure 1: Homunculus: The neurotopic representation of sensory information in the brain.

Tone pitches are situated in the temporal lobe in the brain according to the frequencies next to each other. There we have to distinguish spectral pitch (tonotopic organization) and residual pitch (periodotopic organization). These organizations habe an orthogonal spatial representation [12].

1.2   Temporal coding

The duration of an action potential (spike) lasts about 1 msec. The time dependent course is illustrated in standard publications ([3], p. 24). Neurons generate a series of brief action potentials. The minimum interval between consecutive spikes is determined by the refractory period. The outcome of this is the highest frequency of 0.8 to 1 kHz. In bunches of neurons of the auditory nerve frequencies up to 4 kHz can be observed. This higher frequency is possible because many neurons are activated

during the triggering by the sound signal. Neighboring neurons fill the gap, when other nerves pause firing (population coding, [11], p. 334).

The stimulus intensity is coded by the occurence of spikes per time unit (= spike frequency). This is described as the "temporal coding" of the first kind ([11], p. 333). To overcome the limit of information transmission for a single neuron given by the refractory period (of 0.8 to 1 kHz) neighboring neurons assist in firing: the bundled fire rates are produced by groups of neurons.

The firing behavior of a single neuron (unit) shows a stochastic characteristic. On this background it is possible to simulate the spontaneous activity by stochastic models of single cells. A relevant description is for example the Poisson process, characterized by the rate or mean frequency.
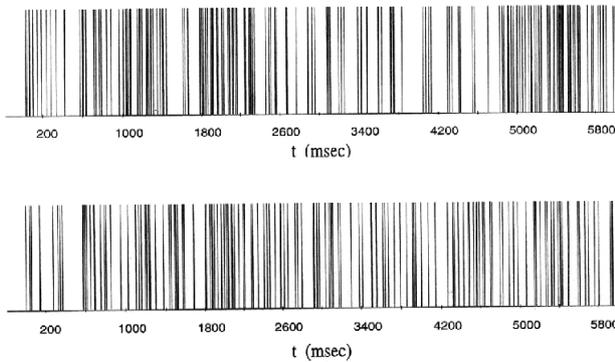

Figure 2: Spike train and the random Poisson process. ([11], p. 354)

The illustrated spike train looks like a bar code. Each bar represents an action potential of a neuron, recorded from a single unit. The spike train shows the statistical distribution of the nerve activity. In the bottom: the spike train from a synthetic Poisson process. It appropriately emulates the nerve activity. This is demonstrated in the next picture.
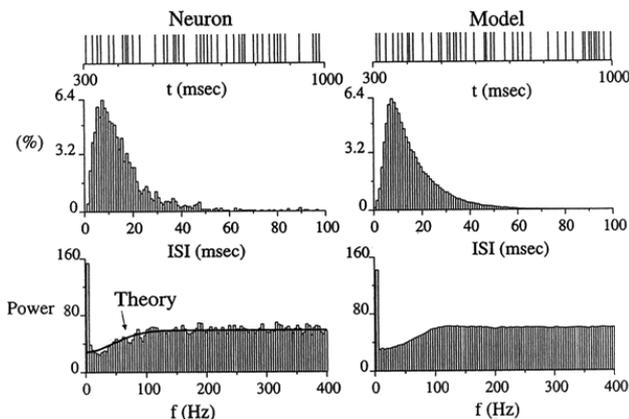

Figure 3: Representative spike train, interspike interval (ISI) histogram, and power spectrum of a neuron recorded from cortical area compared against a modeled point process, Poisson distributed with an absolute refractory period drawn from a Gausssian distribution (of 2 ms standard deviation)(right column). [11], p. 355).

Peri event time histograms and Interspike interval histograms permit a survey of the raster dot displays. Demonstations of these processes are shown in the next picture.
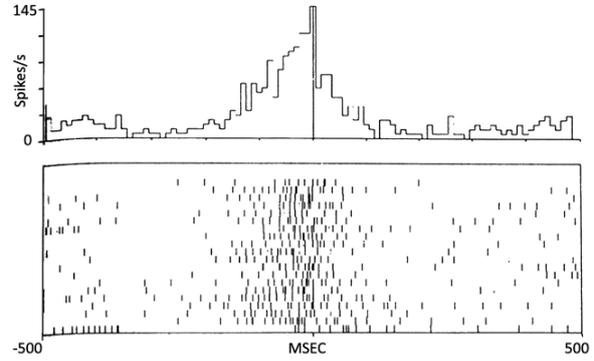

Figure 4: Peri event time histogram (PETH) and a raster dot display showing all the spikes of a single unit around the time of the event. ([1], p. 131). The dots represent the peaks of the registered action potentials.

1.3 Frequency of discharge (rate code)

Compared to other human senses the ear is in a big advantage: sound contains often periodical parts (instead of steady-going stimuli like pressure, smell, temperature etc.). The periodic repeat rate is in the realm of possibilities of the neuronal firing rates.
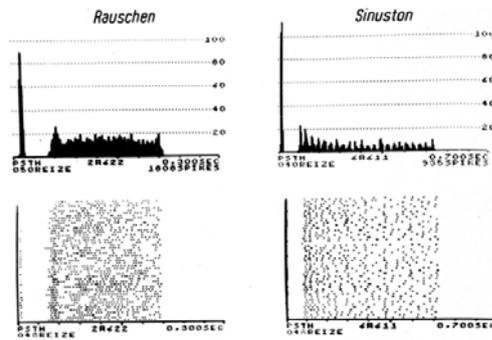

Figure 5: Neuronal activity triggered by sound waves. Left column: noise, right column: sine wave. ([10], p.180).

Peri event time histograms and interspike interval histograms (PET- and ISI-histograms) offer the possibilities to reveal periodicities in the raster dot displays. At the sine wave one can see the periodicity of the sound signal. But one can also see the scattering that appears during the triggering of the waves and during the autocorrelation of the coinciding of the spikes.
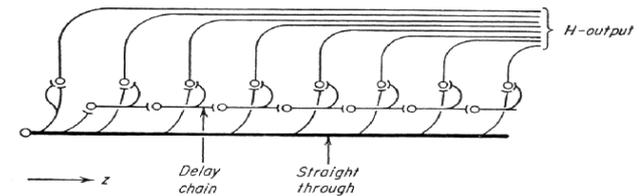

Figure 6: Schematic illustration of neuronal autocorrelation ([13], p. 102, Fig.9). This is a model to detect periodicities.

The straight through neuron is one of the ascending neurons, connected with others in parallel having a delay line inserted. If you imagin that they are running in parallel to the cochlear duct the delay chain is realized by the traveling wave in the duct. The second condition realizing an autocorrelation process is the summation of coinciding spikes. This is veryfied by synaptic integration.

One can find 10.000 synapses from presynaptic cells on every single neuron [3]. In the brain every neuron has contact to about 20.000 other neurons and receives signals from just as much neurons ([14], p. 64).
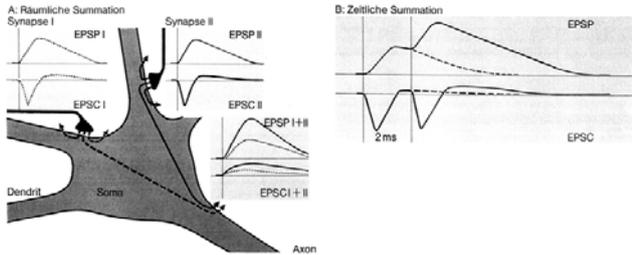

Figure 7: Synaptic integration. Spatial and temporal summation of synaptic potentials.

Note the span of time where a summation still succeeds. After the attenuation of the excitation because of the first synaptic impulse – at a larger time difference – no summation takes place any more. The time span between the synaptic excitations where an addition takes place provides an indication on the imprecision at the coincidence of spikes where a correlation still proceeds. When calculating the correlation that span of time has to be considered as uncertainty.

2. ORIGINS OF UNCERTAINTY

2.1 Uncertainty of localisation

The uncertainty of local representation in the brain corresponds to the local uncertainty of the receptors in the cochlea (tonotopic representation).

2.2 Uncertainty of/by pulse width and uncertainty of time to trigger

2.2.1 Temporal scattering in the triggering process of adjoining neurons

The first source is inside the volley: Here the single spikes are distributed statistically. The volleys appear in a sequence where the distances between single volleys conform to the cycle duration of the sound wave (stimulus). The effect of scattering is even stronger in the distance between the spikes of two successive volleys, corresponding to the cycle duration of the stimulus: The distances between two arbitrary spikes taken from two succeeding volleys are much wider scattered than the width of one single volley. This is a second effect.

2.2.2 Uncertainty of coincidence in the correlation process

In addition (as a third source) there is an uncertainty at the coincidence during the integration process (autocorrelation).
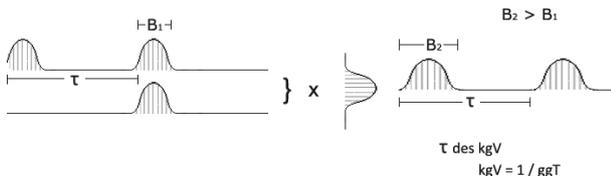

Figure 8: The correlation of spike volleys (containing statistically distributed spikes) with a fuzzy coincidence

pattern results in an even greater variance of periodotopic representation..

The spikes that are triggered by a periodically moved basilar membrane show a random distribution. They are collected in volleys. The periodicity information therefore is coded in the volley distances. By time delay the preceding volley of two succeeding volleys is time shifted. The amount of this time shifting corresponds to the period time. But coincidence detection takes place between two spikes out of two different volleys. This is the source of a much bigger scattering than that in a single volley.

The coincidence detection itself is impaired by the imprecision of the summation process. This is a third source of uncertainty.
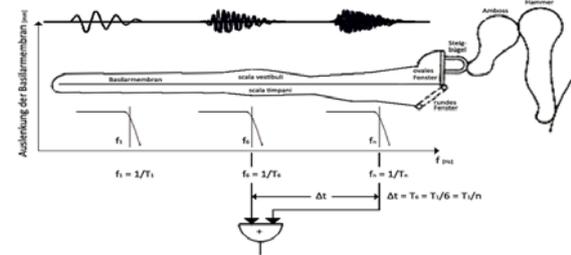

Figure 9: Spatially organized autocorrelation. A filterbank of parallely organized correlators is arranged along the cochlear duct. Each correlator has a fixed time delay. A sequence of low pass filters prevents the higher frequencies to arrive the higher cochlea helix where lower frequencies are processed.

Each filter that works as a correlator with fixed delay time reacts on the fundamental and its partials. But the harmonic partials are suppressed due to the mechanical properties of the cochlear duct [2][5]. These mechanical properties are comparable to a series of low pass filters with decreasing edge frequencies.

The fundamental T_1 is corresponding to τ.
The partials $T_n = T_1/n$ and consequently $f_n = nf_1$ are suppressed.

2.3 Uncertainty of discharge frequency combining the uncertainty of "which" and "when"

The consequence of a filter bank with these characteristics (spatially distributed delays) is a spatial organized autocorrelation and a spatial organized representation i.e. a periodotopic representation in the brain, ("correlation code", [11], p. 333).

3. CORRESPONDING EMPIRICAL /EXPERIMENTAL DATA.
Judgements of musical interval distances. evidence of statistical distribution.

3.1 Experimental setting and results. Evidence of statistical distribution of judgements

Pairs of successive chords played by either harpsichord or a group of violins were presented to music experts (choir directors, music theory teachers and conservatory students). In most cases, a dissonant chord (containing a diminished or augmented fourth, an augmented fifth and a minor sixth, respectively) resolves into a consonant chord following the rules of classical harmony. When hearing these successions of chords in which one note was varied in its frequency, the subjects were asked to decide between a) optimal, perfectly in-tune intonation and b) tolerated, 'still barely acceptable' divergence from

optimal intonation. (For more on experimental setting see [6][7]).

Consonant intervals (fifth, fourth and thirds) included in both the dissonant and the consonant chords were to be judged. These intervals were evaluated separately according to their musical context. Some examples of fifths (Figs. 11, 12) and fourths (Fig. 13) are presented here to demonstrate the extent of variation and statistical distribution.

3.2 Judgements of the fifth in two different musical contexts

In the first case, Fig. 10, optimal, perfectly in-tune intonation of the fifth is presented in different musical contexts. The mean 703,47 Cents (N = 237) has nearly the same value as the theoretical one (702 Cents). But the perfectly in-tune intonations show a standard deviation of nearly 13 Cents. The whole range of scattering compounds about 70 Cents. If you have in mind that a semitone has the extension of 100 Cents this makes about three quarters of a semitone.
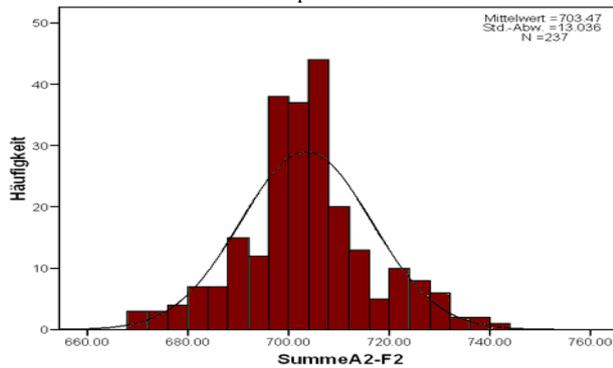

Figure 10: The statistical distribution of judgements on tone intervals is a result of the (at least threefold) statistical distribution in the information processing of the neuronal network.

The next picture (Fig. 11) shows the distributions in case of a differentiation of the musical context.
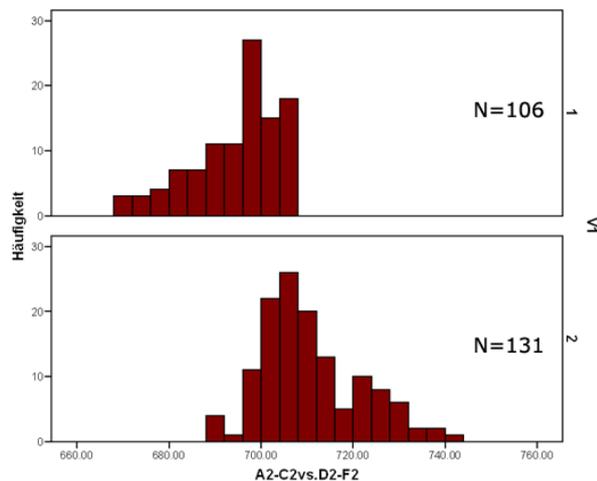

Figure 11: Perfectly in-tune intonation of the fifth in two different contexts in comparison with each other.

If the judgements are separated according to their contexts then the evaluation shows 2 different distributions with two different means. In the first case the mean is positioned at 694.5 Cents (N = 106), in the second case 710.7 Cents (N = 131). The standard deviations amount to 9.22 Cents and 11.03 Cents respectively.

This should not be possible according to musical theory, because the fifth, defined by the frequency relation 3/2, has the value 702 Cents independently from the context.
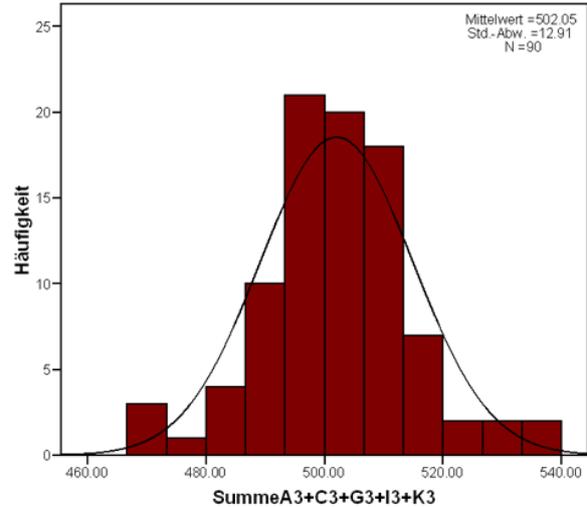
3.3 Judgements of the fourth

Figure 12: The statistical distribution of judgements of optimal, perfectly in-tune intonation of the fourth in different musical contexts. The mean has the value of about 502 Cents (N = 90).

The standard deviation of the statistical distribution amounts in this case 13 Cents as well. The whole range of scattering compounds about 70 Cents too. These findings are in contradiction to music theory and up to now to known consonance theories.

5. CONCLUSIONS

Consonant musical intervals are characterized by periodicity of the sound signal. The limited precision of neuronal information processing in the adjusted Licklider model is responsible for a better fit of the model to the psychophysics of the human ear. It is able now to react on consonant musical intervals that are out of tune in a way we are accustomed to in music hearing. As of yet the measured hearing tolerances have not been able to be derived from the neural processes quantitatively. But they can be explained by the statistical behavior of the neurons.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] Abeles, M. (1994) Firing rates and well-timed events in the cerebral cortex. In *Models of Neural Networks II.* New York et al., 121-140.

[2] Békésy, G. von (1947) The variation of phase along the basilar membrane with sinusoidal vibrations. In *Journal of the Acoustical Society of America* 19, 452-460.

[3] Dudel, J. (1990). Erregungsübertragung von Zelle zu Zelle. In *Physiologie des Menschen.* Berlin, 45-63.

[4] Ebeling, M. (2007). *Verschmelzung und neuronale Autokorrelation als Grundlage einer Konsonanztheorie*. Frankfurt/Main.

[5] Eldredge, D. H. (1974). Inner ear-cochlear mechanics and cochlear potentials. In *Handbook of Sensory Physiology*, Vol. V/1. Berlin, 549-584.

[6] Fricke, J. P. (1988). Klangbreite und Tonempfindung. Bedingungen kategorialer Wahrnehmung aufgrund experimenteller Untersuchung der Intonation. In *Musikpsychologie. Jahrbuch der Deutschen Gesellschaft für Musikpsychologie. Wilhelmshaven,* 67-87.

[7] Fricke, J. P. (2005). Classification of perceived musical intervals. In *Classification -The Ubiquitous Challenge, Proceedings of the 28th Annual Conference of the Gesellschaft für Klassifikation e.V.*, Berlin: Springer, 585-592.

[8] Fricke, J. P. (2010). The recognition of consonance is not impaired by intonation deviations. A revised theory. In *Proceedings of the 34th Annual Conference of the Gesellschaft für Klassifikation e.V., Karlsruhe,* Berlin: Springer. *Im Druck.*

[9] Kandel, E. (2006). *Auf der Suche nach dem Gedächtnis.* München.

[10] Keidel, W. D. (1989). *Biokybernetik des Menschen.* Darmstadt.

[11] Koch, Chr. (1999). *Biophysics of Computation. Information Processing in Single Neurons.* New York et al.: Oxford University Press. (Computational Neuroscience)

[12] Langner, G., & Schreiner, Chr. E. (1988). Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. In *Journal of Neurophysiology* 60 (No 6),1823-1840.

[13] Licklider, J. C. R. (1951) A Duplex Theory of Pitch Perception. In *Experientia* 7, 128-134.

[14] Singer, W. (2002). *Der Beobachter im Gehirn.* Frankfurt/Main.

[15] Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological foundations for the theory of harmony in western tonal music. In *The Biological Foundations of Music, Annals of the New York Academy of Sciences* 930, 92-116.