Development of tempo and precision in semi-spontaneous periodic tapping

Christian GRAFF(1), Nora BIYADI(2), Angélique PIROMALLI(1), Marie CHARMOILLEAUX(1) and Ahmed AHAMI(2)

(1) Laboratoire de Biologie du Comportement, Laboratoire de Psychologie et Neurocognition, UMR CNRS 5015, UFR SHS, Bâtiment Sciences de l'Homme et Mathématiques, Université Pierre Mendès France, BP 47, 38040 Grenoble Cedex 9, France

(2) Laboratoire de Neurosciences et de Nutrition, Composantes cognitives, comportementales et épidémiologiques, UFR BHSP, Faculté des Sciences, Université Ibn Tofail, Kénitra, Maroc

Precision in rhythmic motor production basically satisfies Weber-Fechner principle. However, the performance is better in the preferred motor tempo with regard to faster or slower tempi. In children, the preferred motor tempo is faster, and slows down with age. This change may be due to a change in the choice of the privileged oscillator (replacement), or to a change in the chosen oscillator itself (maturation). Depending on these two mechanisms, a difference between tempi in the improvement of the motor precision with age may point to the preferred tempo. The maturation of the putative oscillator, perhaps privileged among others, was investigated on a finger-tapping task in children aged 6 to 10. We prescribed them to produce different regular sequences at paces that were defined qualitatively: normal, fast, very fast, slow and very slow. The normal one was supposed to approach the preferred tempo. Children's choice of tempi and their precision were estimated through the distribution statistics of inter-tap intervals of the sequences.

Older children actually tapped slower and more regularly than younger ones, for each qualitatively-prescribed pace. However, contrary to adults in a similar situation, children did not perform better (i.e., more regularly or consistently) at the "normal" ones but their tapping was relatively more precise at faster - than at slower paces.

These results show a discrepancy in the temporal registers between children and adults. The development results in a shift towards slower oscillators, or a slowing down of oscillators, with growing age. The higher precision that children displayed in faster tempi or paces may be due to a difference in sequence lengths that were defined by the task requiring a constant number of taps. The resulting bouts were shorter in duration for faster tempi. Younger children may show difficulty to hold a steady tapping activity for more than 10 or 20 seconds due to mnemonic, pre-motor, motor, or muscular restriction. If motor maturation favors a given oscillator among others before age ten, our results did not show it.

Movement timing, Age and Expertise: A practice study

Lavrysen, A.[1], Krampe, R.[1], Swinnen, S.[2].

[1] *Center for Developmental Psychology, K.U.Leuven, Belgium*
[2] *Department of Biomedical Kinesiology, K.U.Leuven, Belgium*

Timing, sequencing and executive control processes are in a more or lesser extent requirements to generate simple or more complex rhythmic movements (Krampe et al., 2005). These processes are found to develop differentially across age and expertise levels, because of the neural changes related to aging and expertise. The present study is part of an ongoing study involving fMRI. The focus of this poster is therefore primarily on methodological issue, describing the approach, presenting first training data and pilot scan results. The goal of the present study is to evaluate the improvement rate and amount in isochronous and rhythmic tapping movements. The isochrounous tasks consisted of producing intervals of different target durations: A (short, 400ms), B (medium, 800ms) and C (long, 1200 ms). In rhythmic tapping, these three target durations were combined to form two sequences: a dominant pattern ABC, and a non-dominant pattern ACB. To maximize executive control demands, these two patterns were combined to form the third sequence: ABC-ACB. Younger (20-30 years) and older (+65) musicians and non-musicians (i.e., novices) produced wrist tapping movements in these six experimental conditions over 8 sessions. Their performance was assessed at pre- (session 2), mid- (Session 5) and post-training (Session 8) evaluation points. No difference was expected between groups for performance in the isochronous tasks, but age and expertise would have a substantial impact on the production of simple and more complex rhythms and the improvement rate and amount. Preliminary results suggest that, for all groups, good performance was reached quite early in practice, whereas the rhythm tasks posed difficulties, specifically for the older novices. The results are discussed in relation to developmental changes in the brain, resulting from age, as well as the impact of long-term training in real-life expertise.

Sensitivity of timing processes for concurrent working memory or executive control tasks

Ralf Th. Krampe, Ann Lavrysen, & Mihalis Doumas

Katholieke Universiteit Leuven

General timing mechanisms as those underlying simple isochronous tapping performance are typically assumed to operate rather autonomously and to be located in subcortical (e.g. cerebellar or basal ganglia). From this perspective, interference from higher cognitive tasks that do not require concurrent output seems unlikely. We show, in contrast, from a several experiments that even simple tapping suffers from simultaneous engagement in working memory or executive control tasks. These effects can be established at the level of realized mean interval durations, variability, and trend. Specifically, under dual-task conditions participants tend to play faster, more variable, and they show stronger trends.

A second issue that can inform our understanding of timing processes are interindividual differences in susceptibility to dual-task interference. In three experiments we investigated the role of age and musical expertise. While tapping as such shows clear developmental improvement from childhood (8, 10, 12 yr olds) to adulthood, performance is similar for young and old adults. At the same time, musical expertise pertains a timing advantage that grows with age from childhood to adulthood and remains stable until old age.
Concurrent working memory performance has a strong effect on tapping, however, if the difficulty of the cognitive task is individually adjusted, dual-task costs for both timing and memory are similar across age groups. In contrast, timing dual-task costs are lower for musically trained subjects. If the concurrent cognitive tasks has a strong executive control component (NBack, Rogers&Monsell type switching task), older adults show higher costs than young adults.

We discuss different timing models (two-level timing, clock-accumulator models) with respect to their potential to explain our findings.

The role of exposure and expertise on making timing judgments in music

Henkjan Honing

Music Cognition Group, ILLC / Universiteit van Amsterdam

This study is concerned with the question whether, and if so to what extent, listeners' previous exposure to music, in everyday life, and expertise, as a result of formal musical training, play a role in making expressive timing judgments in music. This was investigated by using a Web-based listening experiment (Honing, 2006; Honing & Ladinig, 2006) in which listeners, with a wide range of musical backgrounds, were asked to compare two performances of the same composition (fifteen pairs, grouped in three musical genres), one of which was tempo-transformed. The results show that expressive timing judgments are not influenced by expertise levels, as suggested by the expertise hypothesis, but by exposure to a certain musical idiom, as suggested by the exposure hypothesis. Apparently, frequent listening to a certain musical genre allows listeners, with and without formal musical training, to implicitly learn the timing patterns characteristic for that style, and that this (implicit) knowledge facilitates discriminating between a real and a tempo-transformed performance. The current study provides further evidence for the idea that some musical capabilities are acquired through exposure to music, and that these abilities are more likely enhanced by active listening (exposure) than by formal musical training (expertise).

Honing, H. (2006). Evidence for tempo-specific timing in music using a web-based experimental setup. Journal of Experimental Psychology: Human Perception and Performance, 32 (3), 780-786.

Honing, H., & Ladinig, O. (2006). The effect of exposure and expertise on timing judgments: Preliminary results. Proceedings of the International Conference on Music Perception and Cognition (pp. 80-85). Bologna: Italy.

Learning to perform short musical rhythms with expressions

Makiko Sadakata, Alex Brandmeyer, David Hoppe , Renee Timmers & Peter Desain

The concern central to this study is the learning process of expressive musical performances. To our knowledge, there have been no studies done on how quickly expressive music performance can be learned.

A transfer of learning effect, an indication of a learner's ability to abstract and to generalize knowledge of specific structure of the learned materials, was measured in an imitation task of short musical rhythms with expressive variations. Twelve amateur musicians participated in the experiment, which consisted of a pre-test, two training sessions and a post-test. In the training sessions, participants practiced to imitate short rhythms with expressive variations. In the pre and post-test, participants performed an imitation as well as a perceptual discrimination task. Improvement of expressive performance was measured by comparing performance in the pre and post-test, as well as error transitions during the training sessions.

Analyses revealed that learning took place not only for the imitation of the trained renditions but also for the non-trained ones that were presented only in the tests. The degree of improvement varied with stimuli type and dimension of the responses (timing or loudness). An improvement of perceptual accuracy was not observed.

The results provided insight into the issue of which aspects of expressive music performance share a common basis; certain timing variations are closely associated and show larger transfer of learning. Transfer of learning does not behave in the same way for timing and for loudness, which suggests that these parameters are not likely to be integrated into a single dimension of musical expression. These results are useful, for example, for designing computer-assisted tools to learn expressive music performance, and for training expressive performance skills.

Does real-time visual feedback improve expressive percussion performance?

Alex Brandmeyer, Renee Timmers, Peter Desain

Nijmegen Institute of Cognition and Information, Netherlands

Background: The PracticeSpace project aims to develop a real-time visual feedback system for aiding in the learning of expressive musical performance skills. Recent research has indicated that real-time visual feedback can improve the quality of musical performances by students when compared to those of students who did not have the assistance of visualization methods. However, this research has been mostly limited to singing, and has not looked in depth at the effects of different visual representations. It is possible that visual feedback on the expressive timing and dynamics of performances could provide useful, unambiguous information that would help performers to achieve the desired performance quality. We hypothesize that the real-time visual feedback conditions will improve performance over the control condition, as well as rate of learning. Additionally, we hypothesize that feedback providing reduced, high-level information about expressive style would improve performances more than feedback displaying detailed information about each note of the performance, due to reduced cognitive load.

Methods: In this study, we had 18 conservatory percussion students perform repeated imitations of six performances by a drum teacher, with and without the help of two different real-time visualizations. The target performances consisted of two rhythm patterns (8th note and 16th note), with three expressive variants for each pattern (on-the- beat, rushed, and laid-back). The first visualization we refer to as analytic feedback. It displayed a score-like representation including voice, timing, and dynamics information for each performed note. The second visualization was called holistic, and showed a simple shape representing the accuracy of a performance in imitating the target's expressive style. This shape was updated in real-time based on measurements of expressive timing and dynamics from the imitation. Following the experiment we had participants complete a questionnaire to gather qualitative data about the usefulness and ease of understanding of the two visual representations.

Results: Students indicated that they felt positively about the visual feedback, and felt that it would be useful for their study. They were more enthusiastic about analytic feedback, indicating that it was easier to interpret. The results of analysis indicate that the accuracy of the imitations significantly improved across trials. For the 16th note patterns, which was the more difficult of the two patterns, overall performance was significantly higher in the holistic condition, with no significant difference between the analytic and control conditions. No ceiling effect on accuracy of imitation was observed across trials, indicating that longer training periods could lead to additional improvements. Therefore, we conclude that further research involving longer periods of training with visual feedback and additional materials is needed to fully explore the effects of visual feedback on musical performance quality.

Temporally and dynamic features of ID singing

Takayuki Nakata (Future University _ Hakodate, Department of Media Architecture)
Sandra E. Trehub (University of Toronto at Mississauga, Department of Psychology)

We compared the temporal and dynamic features of mothers' performances ($n = 10$) of Twinkle Twinkle for their infants (i.e., ID singing) with non-mothers' performances ($n = 10$) of the same song while singing alone (solo singing). ID singing was temporally more stable and slower than solo singing. Greater stability of ID singing was reflected in smaller average absolute Z-scores of note duration for ID singing ($M = .75$, $SD = .08$) than for solo singing ($M = 1.10$, $SD = .46$), $t(18) = -2.38$, $p = .029$. Slower tempo of ID singing was indicated by longer average note durations (1/4 note as a unit) for ID singing ($M = .479$ s, $SD = .077$) than for solo singing ($M = .385$ s, $SD = .080$), $t(18) = 2.66$, $p = .016$.

Enhanced expressivity of ID singing was revealed by greater dynamic range for the first two bars ($M = 8.51$ dB, $SD = .65$) than for solo singing ($M = 6.33$ dB, $SD = 1.95$), $t(18) = 2.43$, $p = .026$. Dynamics also changed more smoothly for ID singing than for solo singing. When the first four notes of the song were submitted to a mixed-model analysis of variance, there was an interaction between note order (1st, 2nd, 3rd, or 4th note) and condition (ID singing, solo singing), $F(3, 54) = 5.50$, $p < .005$. Subsequent analyses of simple effects of note order for each condition revealed that dynamics (i.e., intensity) increased linearly only for ID singing, $F(1,9) = 65.35$, $p < .001$.

In musical performances, singers convey the emotional message of a song by means of temporal changes and dynamic modulation that is often correlated with the temporal changes. We found, however, that rich dynamic modulation in ID singing was based on an impressively stable tempo. Our findings imply that temporal stability may be an important means of enhancing the predictability of the music for infants, which may, in turn, enhance their attention.

A review of P-centre models

Rudi Villing, Joe Timoney and Tomas Ward

National University of Ireland, Maynooth

The perceptual centre (P-centre) is hypothesised to be the unique moment of occurrence of an auditory event. It is furthermore hypothesised that the P-centres induced by an acoustic signal form the temporal pattern which is perceived as rhythm. For this reason researchers of rhythm need a good P-centre model which can be used to prepare rhythmic stimuli or determine the rhythmic properties of produced speech and musical performance. In order to be widely utilised, a good model should be data independent (subject to level and perhaps frequency range calibration), robust, straightforward to implement and in good agreement with all known P-centre data. For natural speech rhythm research, it is desirable that the model could be directly applied to continuous acoustic signals perceived as sequence of auditory events. This work reviews the existing P-centre models in detail. With varying degrees of sophistication, these models are shown to use one of just two P-centre identification strategies: an onset threshold crossing (requiring only short term temporal information) or a weighted sum of features drawn from the entire acoustic signal corresponding to a single auditory event. Each approach is considered in detail and notable differences of prediction between the models are identified for further empirical investigation.

A problem which affects the validity of all existing P-centre models is that they have been trained/fitted and tested with fairly sparse data. A standard corpus of sounds, labelled with experimentally determined P-centres, would be enormously useful to all future research in the field, but is time-consuming to create. In the absence of a large P-centre labelled corpus, a simple test is proposed for coarse subjective evaluation of the perceptual accuracy of any P-centre model. Finally, it is noted that some published P-centre models are underspecified so that they cannot be implemented without making significant assumptions. This prevents replication and validation of earlier results and reduces confidence in the model. For this reason, it is proposed that any future P-centre model should be accompanied by reference implementation code.

The Continuous Wavelet Transform as a Multiresolution Representation of Musical Rhythm

Leigh M. Smith

Music Cognition Group, ILLC / Universiteit van Amsterdam

A computational method is described that exhaustively represents the periodicities created by a musical rhythm. The continuous wavelet transform (CWT) is used to decompose a rhythm into a hierarchy of time varying frequency components. By using Morlet and Grossman's Gabor wavelet (Grossman et. al 1989), the CWT possesses the best simultaneous time/frequency resolution of multi-resolution analysis techniques, close to the Heisenberg uncertainty relation. This reveals implicit temporal relationships between events over multiple time-scales within the rhythmic frequency range (sample rate of 200Hz), including metrical structure and expressive timing (Smith 1996, Smith & Kovesi 1996, Smith & Honing 2007a). The invertibility of the CWT, like the Fourier transform, enables it as a rhythm visualisation tool and as a representation usable for computational models of musical rhythm cognition.

The validity of the representation is evaluated by applying the CWT on datasets of inter-onset intervals of musical rhythms, including 105 national anthems. Evidence is presented that periods corresponding to the beat (typically a quarter-note period) and the bar (the music theoretic measure of metrical group) are revealed by this decomposition (Smith & Honing 2007b). The use of the CWT in the development of a pulse tracker is an ongoing project which will be reported. Using the CWT, pulse tracking is modelled as a selection process, choosing the most salient time-frequency ridges (or "rhythmic strata") to attend to, at each moment in time (Smith & Honing 2007a).

Of note is that the CWT does not explicitly model the human auditory system. In doing so, it attempts to establish a baseline of how well a data-oriented approach (or a non-cognitive model) can perform. This clarifies the distinction between what mechanisms a perceptual model of beat or meter perception must represent, compared to what information any pulse induction process is capable of revealing directly from the signal representation of the rhythm.

A. Grossmann, R. Kronland-Martinet, and J. Morlet. Reading and understanding continuous wavelet transforms. In J. Combes, A. Grossmann, and P. Tchamitchian, editors, Wavelets, pages 2-20. Springer-Verlag, Berlin, 1989.

L. M. Smith and H. Honing. Time-frequency representation of musical rhythm by continuous wavelets. (submitted), 2007a.

L. M. Smith and H. Honing. Evaluation of multiresolution representations of musical rhythm. (submitted) 2007b.

L. M. Smith. Modelling rhythm perception by continuous time-frequency analysis. In Proceedings of the International Computer Music Conference, pages 392-5. International Computer Music Association, 1996.

L. M. Smith and P. Kovesi. A continuous time-frequency approach to representing rhythmic strata. In Proceedings of the Fourth International Conference on Music Perception and Cognition, pages 197-202, Montreal, Quebec, August 1996. Faculty of Music, McGill University.

Syncopation as a measure for rhythmic expectation

Olivia Ladinig (PhD student)

Music Cognition Group, ILLC/Universiteit van Amsterdam

We aim to contribute to the study of rhythmic expectation by examining the explanatory power of different models of event salience (in the literature sometimes referred to as metric salience) for different listeners. Appreciation of meter (as two or more hierarchically related levels of pulse) versus simply perceiving a beat (as one level of pulse) is often attributed to a difference in skills between musicians and non-musicians (Palmer & Krumhansl, 1990; Jongsma, Desain, & Honing, 2004), arguing that the former being explicitly trained to recognise more complex metrical relationships. However, the observed differences could well just be due to the nature of some experimental tasks and the type of auditory stimuli used (Honing & Ladinig, in revision). For example, the probe-tone method - numerously used to arrive at event salience profiles (Palmer & Krumhansl, 1990; Jongsma et al., 2004) - applied in the temporal domain is confounded with the actual placement of an event in the bar, with events at the end of a bar requiring more effort, to the point of counting along, which is a typical musicians task.

We investigate event salience through studying syncopated rhythms, using syncopation as an indirect measure to probe listeners event expectation (a syncopation being a violation of that expectation). Maintaining rhythmical frames and leaving out single tones - the inverse of a probe-tone approach - provides a more ecologically setup, and allows the testing of mental constructs independent of actual auditory stimulation. We will evaluate four different models of event salience (fully hierarchical, hierarchical only at the highest level, hierarchical only at the lowest level, purely beat- based). In order to test these models and make predictions, we apply the method of calculating syncopations as proposed by Longuet-Higgins and Lee (1984) to the theoretical derived event salience for each model.

Listeners judgements of syncopatedness are collected, using a rating task within a web-based setup. Stimuli are generated of rhythms of the length 16, always providing a natural rhythmical frame, but with certain events systematically left out. In order to make the rating task less difficult and arbitrary, the stimuli are clustered in groups of maximal four rhythms, and the participants have to make a judgement of the syncopatedness within this small context of rhythms. In addition, information about musical training and listening experience is collected.

The data will be used for two things: First, listeners judgements about syncopatedness of rhythmical pattern are used to decide between the four different above mentioned models, and the hypothesis about expert-musicians and non-musicians using different underlying concepts in their rhythmic expectation will be evaluated. Furthermore, individual event salience profiles for each of these listener groups will be generated, and it will be tested if these profiles differ significantly.

Honing, H., & Ladinig, O. (in revision). Exposure influences expressive timing judgments in music. Cognition.

Jongsma, M. L. A., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. Biological Psychology, 66, 129–152.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. Music Perception, 1 (4), 424–441.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. Journal of Experimental Psychology: Human Perception and Performance, 16 (4), 728–741.

Rhythmic similarity in the oral tradition of folk songs

Anja Volk, Peter van Kranenburg, Jörg Garbers, Frans Wiering, Louis P. Grijp, Remco C. Veltkamp

In this talk we present our findings about the rhythmic stability of orally transmitted folk songs using a computational approach to rhythmic similarity.

Folk songs are part of the oral culture and hence are learned by imitation. The transmission of these songs is therefore determined by the capabilities and limitation of the human perception and memory. Since the identity of a song is not determined by a score, the melody is recreated at the moment of the performance from the abstract representation in the mind. During this process variation may occur. As a result, collections of orally transmitted melodies contain a considerable number of variations of one song. These variations are the traces of the original process of passing the songs from one person to the other. Hence the study of these variations and the stable elements may help to understand the cognitive process involved in oral tradition.

To support the study of folk song variation with computational methods is the goal of the interdisciplinary WITCHCRAFT project (What is Topical in Cultural Heritage: Content-based Retrieval Among Folksong Tunes) at the Utrecht University and the Meertens Institute Amsterdam. The aim of this project is to develop a content-based retrieval system for the collection “Onder de groene linde” of Dutch folk songs in order to give the public access to the collection and to support researchers in the study of oral transmission. In this talk we present our investigation into the stability of rhythm among melodies that are considered as variants. Applying a rhythmic similarity measure [Volk et. al. 2007] that is based on Inner Metric Analysis [Fleischer (Volk), 2003] we find that rhythmic information is an important stable element within similar melodies. Hence our results show that rhythmic similarity is a useful characteristic for the classification of folksongs. Therefore the results support the findings of cognitive studies that the rhythmic structure plays a central role in the perception of melodic similarity. Moreover, it confirms the conclusions from an experiment about passing a melody from one person to another in [Klusen et al, 1978] that rhythm is the less varied component than the pitch component in the process of oral transmission.

Kranenburg, P. van, Garbers, J., Volk, A., Wiering, F., Grijp, L. P., Veltkamp, R. C. “Towards a Computer Assisted Folk Song Research System”, submitted to ISMIR 2007.

Fleischer (Volk), A. Die analytische Interpretation. Schritte zur Erschliessung eines Forschungsfeldes am Beispiel der Metrik. dissertation.de - Verlag im Internet Gmbh, Berlin, 2003.

Klusen, E., Moog, H., and Piel, W., 'Experimente zur mündlichen Tradition von Melodien', in: Jahrbuch für Volksliedforschung 23 (1978), 11-32.

Volk, A., Garbers, J., Kranenburg, P. van, Wiering, F., Grijp, L. P., Veltkamp, R. C. "Comparing Computational Approaches to Rhythmic and Melodic Similarity in Folksong Research", Proceedings of the First International Conference of the Society for Mathematics and Computation in Music, Berlin, to appear 2007.

Arom's aksak rhythm: A basis for frequency modulation in rhythm cognition.

John Neelin,

Department of Music, University of Saskatchewan, Saskatoon, Canada

Simha Arom's (2006) insightful examination of aksak rhythm patterns suggests that multiple interpretations of common temporal groupings in music may exist, and that cultural context may critically influence the cognitive representations from which interpretations are made. The perception of the common aksak pattern 3+3+2, for example, is interpreted in some cultures as a syncopated pattern of three onsets within a binary metric structure of two or four beats. If the tempo is sufficiently fast, it may be alternatively interpreted as an unsyncopated pattern of three onsets that "stumbles along" in irregular fashion. In the former interpretation, the pulse structure is isometric, and the cognitive representation is divisive. In the latter, unequal interonset intervals (IOI's) reinforce a heterometric pulse structure, and the cognitive representation is additive (Kolinski, 1973).

Previous studies of categorical rhythm perception have investigated the differentiation of rhythm patterns in terms of duration (Pressing, 1987; Sloboda, 1985). Aksak rhythm patterns may provide a useful basis for the investigation of categorical rhythm perception in terms of frequency and cyclic motion. In this scenario, note onsets may be collectively differentiated according to their frequency of occurrence relative to the pulse, and individually differentiated according to their relative perceptual weight within the metric cycle. The accompanying cognitive representation is neither additive or divisive but, instead, organizes the onsets holistically. The entire pattern may be perceived and conceived as a temporal Gestalt, and while the total number of onsets may be quantified, the characteristic irregularity of the pattern is experienced qualitatively as microstructure. There is evidence for this type of holistic interpretation of rhythm in Cuban music practices, where the aksak pattern 3+3+2 is conceived of as a modulated triplet within a binary metric structure (Mauleon, 1993).

This paper aims to address two principal questions. Firstly, how can the reiteration of binary and ternary cells in aksak rhythm patterns be categorically interpreted in terms of cyclic-harmonic structures? Secondly, what possible roles can cognitive representations of modulated pulse frequencies perform in the perception and performance of microrhythmic music? The systematic modulation of pulse frequencies during music perception and performance varies greatly according to stylistic and cultural norms. However, all frequency modulation practices may be underpinned by common cyclic-harmonic cognitive representations of rhythm structure, and the rich diversity of stylistic and expressive microrhythmic structure may be largely determined by the cultural context of these musical practices.

Walking on Music

Frederik Styns, Leon van Noorden, Dirk Moelants, Marc Leman

IPEM – Dept. of Musicology, Ghent University, Belgium

There is a close relationship between music and body movement. Not only music production, but also music perception is linked to bodily articulations. This experimental study focuses on two aspects of moving the body while listenening to music, synchronization and spatialization i.e. the spatial movement trajectories of human body parts between successive synchronization points. We focused on a basic movement pattern that everyone performs in daily life, namely walking. First, 20 subjects were asked to walk while listening to music and simple metronome fragments and to synchronize their steps with the perceived musical pulse. Second, in a follow up session, the same 20 subjects were asked to tap along with the same fragments. The degree of synchronization and the metrical level at which subjects synchronized was analyzed for both experimental conditions. The metrical levels at which subjects synchronized both their walking and tapping movements seemed to be more or less the same. The spatialization aspect of walking was operationalized as step size or walking speed (walking speed = step size * step frequency). It was studied whether there were differences in walking speed between musical and metronome fragments of the same tempo and between musical fragments of the same tempo. There was a clear difference in walking speed while synchronizing with musical stimuli and with metronome stimuli, but differences in spatialization between different musical fragments of the same tempo could not (yet) be established. Finally, step size was analyzed in terms of a resonance model. The resonance curve obtained from the walking data was compared to the perceptual resonance curve as found by Van Noorden and Moelants (1999). The apparent close relationship between the walking resonance and the perceptual resonance may indicate that the perceptual resonance is based upon an awareness of the locomotion resonance while listening to music. Walking on music turns out to be a rich and multidisciplinary research area for which surprisingly little knowledge has been assembled yet. It combines relevant insights for music and rhythm perception research, music education, sports and handicapped training and walking dynamics.

Tracking the path of rhythms in a virtual environment

João M. Martins and Eduardo R. Miranda

Computer Music Research - Future Music Lab, School of Computing Communications and Electronics, University of Plymouth

Evolutionary computation and Artificial-life provide a path to understand large scale phenomena such as the evolution of what is broadly defined as human culture. Although A-life simulations are in many ways highly simplified models of reality, these models benefit from reduction of scale and create possible avenue for understanding complex phenomena such as music evolution. These tools have already been successfully applied to the evolution of language, namely the emergence of vowel systems, grounding word meaning, etc. These typically involve agents playing games and adapting internal representations after the interaction process. We are currently looking into ways of exploring similarities between language and music regarding their structure by using the same methodology of analysis.

In this study, we consider that rhythm structure develops in time towards higher complexity through a process of cumulative culture evolution. This approach allows for the creative process to be rooted within the social environment as a result of imitation games, incremental transformations and recurrent use of rhythm sequences acquired during previous interactions.

Virtual agents' societies were developed, considering geographical position of the agents, body related constrains, and interaction mechanisms based on imitation that enable the emergence of rhythmic patterns, ranging from simple isochronous pulses to more complex structures. We also developed three algorithms, referred to as popularity, transformation and complexity algorithms, respectively. For modelling purposes we also found it useful to develop measurements of similarity and complexity of rhythmic sequences.

In this system the user can explore the potential of these algorithms to generate rhythmic sequences and also monitor the behaviour of the system. The system offers the ability to extract information about its behaviour, providing researchers and composers the means to explore the outcomes systematically. We look into data from the perspective of the sustainability of particular rhythmic sequences during the simulation, from the perspective of repertoires of individual agents and, in a larger scale, to clustering phenomena within the society.

While the system is able to produce a great variety of rhythms and coherent rhythmic variations, which is what we had expected to observe in the first instance, the system also displayed a number of interesting and surprising behaviours which we wish to present.

Attentional load discontinuity across ITI or activating effect of the pace?

Jacques LaRue

Laboratoire AMCO, Université d'Orléans, France

Subjects performed a finger tapping task at different ITI's (250, 300, 350, 400, 450, 500, 550, 600, 650, 700 ms) according to a synchronization/continuation paradigm. ITIs were tested in random order. Synchronization stimuli were provided by auditory beeps (buzzer, 1000 Hz, 50 ms duration). During the continuation phase, the subjects had to verbally respond (as fast as possible) to visual stimuli. They were instructed that the most important task was to maintain a stable pace. Visual stimuli occurred randomly, but were separated by at least four taps. Results show that mean voice reaction times are faster for the fast paces and longer for the slow paces. Most interestingly, the increase in reaction times occurs suddenly around the 500 ms ITI. According to the double task paradigm, this result suggests that long ITIs request more attentional resources. However, many subjects spontaneously reported a subjective impression of being slower on longer ITI's. Therefore, the faster reaction times might be due to an increased activation induced by fast pace. We are currently testing this hypothesis.

The Role of Distal Action Effects in Simultaneous Intra-personal and Inter-agent Coordination

Peter E. Keller[1] and Bruno H. Repp[2]

[1] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[2] Haskins Laboratories, New Haven, Connecticut

Skilled rhythmic activity (e.g., music performance) often involves precise temporal coordination at two levels: (1) intra-personal – between an individual's own body parts (e.g., two hands), and (2) inter-agent – between one's own actions and externally controlled events (e.g., another's actions). Actions in such contexts often have salient, discrete distal effects (e.g., musical tones). The current study addresses how the stability of simultaneous intra-personal and inter-agent coordination is affected by the degree to which one's own action effects and externally controlled events can be integrated in a simple Gestalt.

This issue was investigated using a challenging task that required participants to tap on a MIDI percussion pad in antiphase with computer-controlled tone sequences (inter-agent coordination) while alternating between the two hands (intra-personal coordination). The maximum rate at which musicians could perform this task was measured (using an adaptive staircase procedure) under conditions in which taps did or did not trigger tones. The pitches produced by the two hands (very low, low, medium, high, very high) could be the same as, or different from, one another and the (medium-pitched) metronome tones.

Faster movement rates were achieved with than without action-effect tones. This benefit was greatest when action effects were close in pitch to metronome tones and the left hand triggered low tones while the right hand triggered high tones. Thus, coordination was facilitated by action-effect tones that were perceptually distinct from – but easy to integrate with – each other and the metronome tones, especially when movements and action-effect pitches were mapped compatibly as on a piano (left = low, right = high).

Readily integrable action-effects may allow a common (perceptual) locus of control to be used for intra-personal and inter-agent levels of coordination. This may (1) redress the asymmetry in coupling strength between the two levels (inter-agent coupling is weaker than intra-personal coupling for antiphase coordination), and (2) allow the dual task (two levels of antiphase coordination) to be conceptualized as a simpler single task (producing a melody).

Distinctiveness in action-effect tones may allow relatively strong, unambiguous associations to be formed between movements and their distal effects (especially when compatible). Such associations may facilitate (1) the planning (selection and programming) of upcoming movements, and (2) the evaluation of feedback from previous movements. Rapid planning and feedback evaluation would permit stable coordination at fast rates.

The present findings suggest that the stability of simultaneous intra-personal and inter-agent coordination in music-like contexts may be influenced jointly by the ability to perceive a coherent whole Gestalt and the ability to keep track of the separate parts comprising this whole.

Reuniting Speech and Rhythm

Fred Cummins, University College Dublin

To be written

Perceived speech tempo and articulatory effort

Hugo Quené

Utrecht inst of Linguistics OTS, Utrecht University

Speech tempo or speaking rate is important in speech perception, because listeners use tempo to normalize other phonetic distinctions (e.g. voiced/unvoiced consonants, long/short vowels, etc) and to estimate the relative importance of the spoken information. Listeners' judgements of tempo are based not only on the number of realized speech sounds per second, but also on the number of intended speech sounds per second (J. Koreman, 2006, JASA 113(1), 582-596). This suggests the main hypothesis for this study, viz. that perceived speech tempo is affected by listeners' awareness of the articulatory effort required for the stimulus speech. For example, the sequence "titititi" spoken at 5 syllables/sec will be judged as *slower* than the sequence "papapapa" spoken at exactly the same tempo. The sequence "titititi" requires only little articulatory effort (short trajectory between consonants and vowels), whereas the sequence "papapapa" requires more articulatory effort (long trajectory). Presumably, listeners will compensate for this articulatory difference when judging speech tempo. This hypothesized effect is similar to listeners' compensation for coarticulation between adjacent speech segments, where listeners "undo" the low-level acoustic effects of coarticulation that arise during speech production.

The above hypothesis was investigated in a magnitude estimation experiment. Stimuli consisted of nonsense sequences of 5 disyllabic words of low articulatory effort (easy, e.g. "tisi"), or of high articulatory effort (hard, e.g. "soke"). These speech sequences were constructed by cutting the target words out of reiterant speech (which was produced in sync with a metronome), and aligning these words with a click track at the target tempo. This yielded 6 easy and 6 hard sequences. Similar stimuli were constructed for target speaking rates of 3.0, 3.5, 4.0 and 4.5 syll/sec. In the magnitude estimation experiment, listeners first heard a baseline stimulus of mixed easy and hard words, to which an arbitrary subjective tempo value of 100 was preassigned. They then assigned subjective tempo values to the following stimuli, with the instruction to give higher values to stimuli perceived to be faster (e.g. if the stimulus sounds twice as fast, it should be given a subjective tempo value of 200). Presentations were blocked by stimulus speech tempo, with varying order between tempo blocks.

Although the experiment is not yet concluded, preliminary results suggest that hard sequences (of higher articulatory effort) receive somewhat higher tempo values than easy sequences, as predicted by the main hypothesis under investigation. This predicted effect of articulatory effort on perceived speech tempo seems to be limited, however, to the faster stimuli. Listeners seem to compensate for articulatory effort, but only if the speech tempo is so fast that it may indeed have been constrained by articulatory limitations during speech production. If confirmed by subsequent experimentation, these findings support the general idea that speech perception is affected by listeners' implicit knowledge about speech production.

Variability in articulation across speakers and tasks

Peter Howell and Andrew Anderson
University College London

A measure of speech performance, used in many studies, is mean speech rate. Recently work has estimated performance variability over repeated attempts of the same utterance. The phrase most often used is “buy Bobby a puppy” and lip movement records are usually obtained. The utterances are amplitude-normalized and time-normalized (the latter was originally done linearly, but is being done non-linearly more recently), variability across the normalized records at discrete time points are obtained and variance across all time points obtained to provide the so-called spatio-temporal index (STI). The STI has been shown to be sensitive for detecting differences between speakers and tasks.

A drawback to the technique is that it has mainly been limited to lip movement. Equipment for lip movement is bulky, expensive, scary to children and patients and the equipment cannot be used in scanners so concurrent CNS activity in tasks cannot be established. Also, as measurements are restricted to lip measures, there are few opportunities for measuring inter-articulator coupling. Two experiments are reported that address how STI can be extended to other articulations that do not need specialized equipment (audio records using just a microphone).

STI indices of variability, except those used for lip tracking, were obtained from the audio signal (energy, pitch and formants). In study one fluent speakers and speakers who stutter repeated “buy Bobby a puppy” 20 times. The second involved fluent speakers saying the phrase “Well we’ll will them” 25 times each at fast and slow rates.

In experiment one, lip and audio STI scores were shown to discriminate successfully between fluent speakers and speakers who stutter. Experiment two showed that audio STI scores were greater at normal rate than at fast rate. The reduced variability at the fast rate was consistent with undershoot of articulator positioning happening.

In conclusion, the main points are 1) audio STI scores are a useful adjunct to lip movement scores and 2) the audio scores can also be used when participants have to be scanned. We discuss ways in which these ideas can be applied to a range of other study areas.

EEG signatures of Subjective Rhythmisation

Rebecca Schaefer, Bas de Kruif & Peter Desain

"Music, Mind, Machine" group, NICI, Radboud University Nijmegen

In the so-called clock illusion, isochronous stimulus trains are subjectively grouped into a binary percept (tick-tock-tick-tock instead of tick-tick-tick-tick). To use the manifestation of subjective accenting in EEG (cf. Brochard et al. 2003, Snyder & Large 2005) for realisation of a Brain-Computer Interface (BCI), we measured EEG after instructing participants to imagine different groupings superimposed on an isochronous train of stimuli, thus producing accented and nonaccented beats in identical metronome ticks. Binary, ternary and quaternary metric patterns were used, including a perception part and an imagery part in each trial. For data-analysis, ICA was used to determine localized dipoles, and ERP's were calculated based on single ICA-components. Preliminary results show a marked effect of inter-trial coherence as an effect of the metronome, and effects of subjective accenting, though varying somewhat over participants, in the beta band and, surprisingly, less so in the gamma band, as would be expected. As realisation of a BCI system only requires robust within-subject consistency, the results appear promising. Next steps will focus on single-trial classification of single beats.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. Psychological Science, 14(4), 362-366.

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. Cognitive Brain Research, 24(1), 117-126.

Dynamic Attending to Syncopated Rhythms: Behavioral and Neural correlates

Theodore P. Zanto[2], Heather L. Chapin[1], Edward W. Large[1], J. A. Scott Kelso[1], Fred Steinberg[2]

[1]Center for Complex Systems and Brain Sciences, Florida Atlantic University
[2]Department of Neurology, University of California at San Francisco
[3]University MRI, Boca Raton, FL

We examined the role of attending in the perception of syncopated rhythms. Participants performed a rhythm reproduction task in which the stimuli were ten highly syncopated patterns. During the attend auditory condition, participants were instructed to attend to a syncopated rhythmic pattern, mentally rehearse the pattern during a retention interval, and then reproduce the pattern. During the attend visual condition, participants heard the rhythmic patterns, however, they were instructed to study a list of words, remember the words during the retention interval, and then recall as many words as possible. Participants performed the task twice, once during an EEG recording session, and once during an fMRI recording session. We measured evoked (phase-locked) beta-band and gamma-band responses to rhythmic patterns in the attend auditory and attend visual conditions. Evoked activity anticipated the beat of the rhythm in both conditions, and power was attenuated when ignoring the rhythmic stimulus. Functional brain activation was observed in medial frontal gyrus and anterior cingulate gyrus when attending to the auditory versus visual stimuli. Implications for dynamic allocation of attention are discussed.

Evidence of metrical accenting in physically identical tones: a MEG study

Donna Abecasis[1], Renaud Brochard[2], David del Río[3] and Tomás Ortiz[3]

[1] Laboratoire de Psychologie Expérimentale, CNRS UMR 8581, Institut de Psychologie, Université René Descartes (Paris 5), Boulogne-Billancourt, FRANCE
[2] Université de Bourgogne, Pôle AAFE- Esplanade Erasme, Département de Psychologie, BP 26513, 21065 Dijon Cedex, FRANCE
[3] Centro de Magnetoencefalografía Dr. Pérez Modrego, Pabellón no. 8 Facultad de Medicina, Universidad Complutense de Madrid (UCM), Avda Complutense s/n, 28040 Madrid, SPAIN

The perception of a meter, or the alternation of strong and weak beats, was assessed in musically trained listeners and possible lateralisation effects were explored through magnetoencephalography (MEG). Metrical accenting was examined directly, with no disruption of the temporal structure of the sequence, thus not affecting the serial grouping of temporally adjacent events. Furthermore, no auditory task was given to the participants. Results showed an effect of metrical processing in standard, identical tones in the left hemisphere of musicians at around 170ms post-stimulus onset, where tones on strong beats showed larger dipole strength than on weak beats. Occasional intensity deviant tones (8 dB louder than standards) disrupted this pattern, with weak beats tending to elicit larger responses. These results seem to support the view of a relatively early, left-hemispheric effect of metrical processing in musicians, perhaps related to covert musical expectancies.

Binary-ternary ambiguity in tempo perception

Dirk Moelants
IPEM-Dept of Musicology, Ghent University, Belgium

When listeners are asked to tap along with the tempo of the music, there appears to be a choose between a limited number of metric levels. Usually these levels stand in a 2:1 or 3:1 relationship to each other. However, in previous experiments it was observed that some pieces yield responses that indicate conflicting metric interpretations, mostly with tap rates in a 3:2 ratio. In those pieces, metric levels that were considered to have a ternary structure from a music theoretical viewpoint were interpreted as a binary level by a number of listeners. In order to gain more insight in this phenomenon a set of 60 excerpts containing a 'ternary' metric level was selected, assuring a spread of musical styles and a wide range of tempi (68-336 at the level below the 'theoretically ternary' level). 60 subjects participated in the experiment. The results show that only in 2 of the pieces there was no trace of 2:3 ambiguity. In the others it varied between a single 'dissident' participant and 'perfect confusion'. Clearly some music more easily gives rise to ambiguity than other. Analysis of the data of the individual participants shows that there are huge differences between participants, both in the mean tapped tempo (47 - 157 bpm) and in the ternary:binary ratio (35.3 - 0.38). However, factors as musicianship, age, gender and ethnic background could not give any explanation for these differences. Thus perception of metrical structure seems to be highly individual. These results have a number of implications in different domains. It challenges music theoretical approaches that take the score and the personal vision of the music theorist as ground-truth and similarly challenges automatic beat tracking systems which normally don't allow a personal interpretation of individual listeners. Within the field of signal processing it would be interesting to take a look at which factors in the music influence the amount of ambiguity. The results also raises questions about musical communication: if different subjects give different body responses when interpreting the same musical structures, it indicates that they attibute a different meaning to the musical signal. How can a theory of embodied music cognition, founded on musical communication through gestures, deal with this?

Modality effects in time discrimination in human and nonhuman primates: Further support for relations between speech and time perception.

Dietmar Grube [1], Robert Niebergall [2], Stefan Treue [1,2] & Kristian Folta [2]

[1] University of Göttingen, Georg-Elias-Mller-Institute for Psychology
[2] German Primate Center, Göttingen

In human subjects, processing of auditorily presented temporal information is superior to that of visually presented information. One explanation of this phenomenon is that, within the internal clock, processing of visual signals requires more attentional resources than processing of auditory signals.

Another explanation assumes that modality effects result from the differential involvement of modality-specific cognitive resources in temporal processing of auditory and visual information. There are different lines of evidence supporting the position that time and rhythm processing is linked to resources that are known to be involved in language related processing: Immediate reproduction of time patterns is substantially and specifically related to immediate reproduction of verbal series; in dual-task experiments, temporal and verbal processing specifically interfere with each other; there are overlaps in brain areas involved in temporal and verbal processing; and musical training was shown to specifically improve verbal memory - to show only some examples.

The present study examines the relation between language and time processing by comparing the modality effect from data of (a) language experienced humans and (b) nonhuman primates (rhesus monkey) who do not dispose of the faculty of language. Series of six auditory (tone) and visual (rectangle on a computer screen) signals were presented. Subjects had to identify a prolonged signal within a series of six stimuli (the time-difference between standard and target was manipulated). In a first step subjects were trained under conditions of combined presentation modalities until they reached a performance criterion. Afterwards series were presented only auditorily, only visually or auditorily and visually in a combined condition.

The hypothesis of a relation between language and time processing predicts a stronger modality effect in humans than in nonhuman primates. What we found was a double dissociation between subject groups and modality, representing strong support for the hypothesis.

The role of the target position in a sequential paradigm of time discrimination

Kristian Folta[1], Robert Niebergall[1], Stefan Treue[1,2] & Dietmar Grube[2]

[1] German Primate Center, Göttingen
[2] University of Göttingen, Georg-Elias-Mller-Institute for Psychology

Time discrimination is a widely used paradigm to investigate models of time perception. We used a sequential presentation paradigm in order to investigate time discrimination: Series of six auditory (tone) and visual (rectangle on a computer screen) signals were presented within the range of few seconds. Subjects had to identify a prolonged signal within a series of six stimuli. The longer signal could occur at the fourth, fifth, or sixth position. The time-difference between standard and target was manipulated systematically. In a first step subjects were trained under conditions of combined presentation modalities until they reached a performance criterion. Afterwards series were presented only auditorily, only visually or auditorily and visually in a combined condition.

This study was aimed to explore the dependency of discrimination performance of the target position that is related to changes of attention and/or predictability within the course of the sequence. The manipulation of subject group (human and nonhuman primates) and modality (auditory, visual and combined) allows detailed analyses of interdependencies between the factors. Results are discussed within the contexts of models of time processing and of the multimodal working memory framework.

EFFECTS OF METRICAL SUBDIVISION ON SUBJECTIVE BEAT RATE

Bruno H. Repp
Haskins Laboratories, New Haven, Connecticut

One explanation of the anticipation tendency in synchronized tapping is that the empty inter-onset intervals (IOIs) of the pacing sequence are systematically underestimated in perception (Wohlschläger & Koch, 2000). Wohlschläger and colleagues have conducted experiments in which the anticipation tendency, and hence presumably the underestimation, disappeared when soft random tones ("raindrops") or additional finger movements were inserted into the IOIs of the pacing sequence. This is consistent with psychophysical studies on the "filled interval illusion": Intervals appear to be longer when they are not empty.

The present research investigated whether metrical (i.e., regular) subdivision of an auditory beat slows its perceived rate, as this might have some interesting implications for music perception and performance. Experiment 1 used a synchronization-continuation tapping paradigm: Musically trained participants tapped in synchrony with the beat of an isochronous auditory sequence that was either a simple beat or was subdivided metrically (duple, triple, or quadruple) by additional tones, with beat IOIs ranging from 600 to 1000 ms. After the sequence stopped, participants were required to continue tapping the beat. All participants tapped slower following a subdivided beat than following a simple beat, more so with triple and quadruple than with duple subdivision. Experiment 2 explored several variants of the task. A similar, though somewhat attenuated effect of subdivision was obtained when participants produced the subdivisions themselves by tapping with the other hand during synchronization. By contrast, the effect largely disappeared when participants produced both beats and subdivisions during continuation tapping. Experiment 3 demonstrated a corresponding effect in a purely perceptual task requiring judgments of relative tempo: Most participants judged a continuation beat to be faster than a preceding subdivided beat when their IOIs were in fact the same. Results from a fourth experiment, in which participants will carry out matched production and perception tasks, may be available at the time of the meeting.

On the whole, the results suggest that "busy" music should sound slower than "sparse" music if the beat period is the same, and should also tend to be played slower when the same tempo is intended. These predictions remain to be tested with real music.

On the amazing robustness of the Two-Level timing model.

Dirk Vorberg

Institut für Psychologie, Technische Universität Braunschweig
Spielmannstr. 19, D-38106 Braunschweig

Finger tapping at a given tempo is trivially simple, yet on closer look the task offers fascinating possibilities for studying the temporal dynamics of movement planning and control. This is due to the elegant framework introduced by Wing and Kristofferson (1), which explains why temporal precision in tapping is limited, and at the same time offers a decomposition of observed variability into central and peripheral sources. Its simplicity notwithstanding, the model has successfully passed numerous empirical tests, and serves as the basic building block in more complex models for tasks like, e.g., rhythmic, bimanual, and synchronized tapping. Recent tests of the model's open-loop assumption, however, suggest a reinterpretation in terms of reafference within a framework in which anticipated action-effects are timed, rather than movement initiations. The revised model dynamically adapts to feedback perturbations, but at steady-state is virtually indistinguishable from the Two-Level timing model. This hints at an explanation for why the original model has turned out so successful.

(1) Wing, A. & Kristofferson, A.B. (1973). Response delays and the tim-ing of discrete motor responses. /Perception & Psy-chophysics/, *14*, 5-12.

Avoiding the Ungrammatical: Supporting Evidence for Auditory Grammar and the Event Construction Model

Perez, E. & Nakajima, Y.

Dept. of Acoustic Design, Kyushu University

It is well established that what we hear is not necessarily what is present. The auditory system seems to reorganize the incoming sounds in order to create plausible auditory events and streams. This reorganization is clearly illustrated with new phenomena such as the gap transfer illusion (Nakajima et al., 2000) or the classic restoration effect (Miller & Licklider, 1950; Warren et al., 1972). According to Auditory Grammar (Nakajima & Sasaki, 1996), the occurrence of the phenomenon is due to the auditory system imposing a set of rules and temporal constraints to avoid any `ungrammatical' percept. To test this hypothesis, we measured the listeners' ability to resolve ungrammatical patterns by creating complex with synchronous onset time, but with asynchronous offset time. The `unusual' offset asynchrony was created by ending one of the components before the other components in the complex. In some conditions, two consecutive tones were perceived. Similar to the time-swelling effect (Sasaki et al., 1993), the duration of the second tone, which should have been the rest of the complex after one of the components was ended, was overestimated. The stimuli used were two, four or six-tone cluster of steady-state tones of different durations (160, 320, or 800 ms). The parameters investigated were temporal offset asynchrony (20, 50, 100, 240 ms offset asynchrony for 320 ms and 800 ms complex, or 20, 40, 80, 120 for 160 ms complex), the frequency positions of the displaced components in the complex (higher vs. lower frequencies), frequency separation between components (large, 500 Hz, vs. small, 200 Hz, for harmonic components and equivalent ratio for non-harmonic components) and the harmonic relationship between components (harmonic or linear vs. non-harmonic or logarithmic). Listeners were able to describe the percept of the stimuli presented and judge the duration of both tones. The appearance of this new auditory phenomenon suggests that offset cues can behave as if they were independent and dynamic subevents (i.e., onsets, offsets, fillings and silences) that can change their functional role (e.g., from offset to onset) to avoid ungrammatical percepts or ill-formed auditory events. Moreover, the auditory system seems to reorganize subevents in order to create equal temporal events of approximately 150-300 ms of duration by overestimating the duration of second tone. The results derived from these auditory illusions illustrate that time perception, with regard to audition, do not correspond to the immediately present acoustic reality, but rather to the outcome of temporal integration during a sliding window of approximately 150-300 ms. Auditory Grammar takes in account these temporal constrains by proposing a two stage analysis, one associated with approximately 30 } 10 ms interval for subevent detection and processing, and a longer 150-300 ms interval (Cowan, 1984) required for the integration of those subevents into a coherent representation for higher-level processing.

Cowan, N. (1984). On short and long auditory stores. Psychological Bulletin, 96, 341-370.

Miller, G. A., & Licklider, J. C. (1950). Intelligibility of interrupted speech. Journal of the Acoustical Society of America, 22, 167-173.

Nakajima, Y. & Sasaki, T. (1996). A simple grammar of auditory stream formation. Journal of the acoustical society of America, 100, 2681.

Nakajima, Y., Sasaki, Y., Kanafuka, K., Miyamoto, A., Remijn, G., & ten Hoopen, G. (2000). Illusory recouplings of onsets and terminations of glide tone components. Perception & Psychophysics, 62, 1413-1425.

Sasaki, T., Nakajima, Y., ten Hoopen, G. (1993). The effect of preceding neighbor-tone on the perception of filled durations. Proceedings of the spring meeting of the Acoustical Society of Japan, 347-348

Warren R.M., Obusek C.J., Ackroff J.M. (1972). Auditory induction: perceptual synthesis of absent sounds. Science, 176, 1149-1151

About temporal information processing limitations

Simon Grondin,

Université Laval, Québec, Canada

If you are asked to estimate a 15-s interval, it is very likely that you will spontaneously adopt some rhythmic activity, like counting steamboats or numbers, or tapping with your foot, to keep track of time. Indeed, humans do segment temporal information like they chunk pieces of cognitive information to remember it. However, if you are asked to estimate (or produce) a 1-s interval, it is less certain that segmenting time will increase the accuracy of your estimates. One perspective on the meaning of this need to segment temporal information is that it reflects the limitation of some temporal span, i.e. of the capacity to process temporal information. One way of defining this limitation is to look at Weber's law for temporal tasks. This law states that the variability of estimates (or difference threshold) is proportional to the magnitude of the sensory continuum under investigation. In other words, in the context of temporal estimates, the ratio of variability to time should be constant—the Weber's fraction is supposed to be constant—and increases of this fraction with longer durations would reveal a decline of the capacity to process temporal information.

After a brief review of the time perception literature indicating that, as early as for values below 1.5 s, there is a critical transition point in the capacity to process time, two experiments involving two methods confirmed the disruption of the Weber's law as intervals get longer. In the first experiment, participants used a finger-tapping method to produce series of .8- to 3-s target intervals. The dependent variables of interest were the mean produced intervals (MPI), their variability (SD) and the coefficient of variation (CV=SD/MPI), i.e., a variant of the Weber fraction. Data showed that the CV became much higher at 1.4 s than at 1.2 s; and also showed that clear benefits of using sub-divisions occurred at 1.4 s. In the second experiment, time intervals were presented in sequences marked by brief auditory signals (standard values = 1, 1.3, 1.6 and 1.9 s). Participants were asked to indicate whether, in a presentation of two series of 1, 3 or 5 intervals, with each series being marked by 2, 4, or 6 signals, the intervals of the second sequence were shorter or longer than those of the first. The results showed that there is a gradual increase of the Weber fraction with longer intervals. Moreover, discrimination was poorer when only one interval was presented, but remained about the same whether 3 or 5 intervals were presented. The increased CV or Weber fraction, when target/standard intervals lasted more than around 1.2 - 1.3 s, is interpreted as a fundamental temporal limitation of the information processing system.

Metricality-enhanced encoding accuracy and subjective perception of rhythmic sequences.

Manon Grube, Tim D. Griffiths

Newcastle Auditory Group, Medical School, University of Newcastle-upon-Tyne

We tested the hypothesis that higher-order metrical structure enhances temporal encoding accuracy alongside with the subjective perception of rhythmic sequences. Behavioral correlates of encoding accuracy and subjective perception were assessed in two experiments. Sequences were composed of a number (6-9) of identical tones (200 Hz, 100 ms) occurring over 16 units (200 ms on average), mimicking a minimum hierarchical structure of four 4/4 bars. Higher-order metrical structure of sequences was manipulated in a twofold manner. Firstly, sequences either contained regularly spaced purely temporally accents, inducing a strongly metrical beat or not (weakly metrical), based on the model by Povel & Essens (1985); secondly, sequences ended at a metrically plausible point on the last beat (compact) or not (open), resembling a final lengthening of a musical piece or not (Paler & Krumhansl, 1990). The first experiment investigated encoding accuracy and subjective ratings in a discrimination task using an adaptive 3-alternative-forced-choice paradigm for 12 sequences in total. Previous studies on the influence of metricality on temporal encoding looked at changes in interval length at one or two specific points of a given sequence and yielded inconsistent results (pro: Hebert & Cuddy, 2002; Ross & Houtsma, 1994; contra: Handel, 1998). In the discrimination task here, changes concerned a number of intervals at multiple points in the sequences; a tempo roving between sequences was applied to prevent the use cues of absolute durations. The critical cue for the detection of change was the temporal relation between long and short intervals across the sequences as a whole. Change detection thresholds were significantly lower for strongly compared to weakly metrical beat induction and also for compact compared to open endings (n=27; main effects of beat strength, $p<.001$, and ending, $p<.001$; Repeated Measures ANOVA). An additionally administered ratings task yielded lower difficulty ratings for strongly compared to weakly metrical beats and compact compared to open endings (main effects of beat strength, $p<.001$, and ending, $p<.01$), paralleling the effects of metricality on temporal encoding accuracy. The second experiment was designed to replicate the effects of metricality on subjective ratings for the same kinds of sequences in a single trial rating paradigm, where subjects heard every sequence only once, using a total of 740 sequences. Difficulty ratings were again significantly lower for sequences featuring a strongly metrical beat (N = 4, $p<.001$, Repeated Measures ANOVA on Ranks, Man-Whitney U-Test); the effect of ending was not significant. Independently of the effect of metricality, ratings decreased with increasing number of tones ($p<.05$, Jonckheere Test for Ordered Alternatives). The behavioral findings reliably demonstrated how the presence of higher-order metrical structure improves temporal representations and perceived salience of rhythmic sequences, and now lead on to functional approaches including EEG, fMRI and patient work to find out how and where the rhythm network of the brain “feels” that beat.

References:

Handel, S. (1998). "The Interplay Between Metric and Figural Rhythmic Organization." Journal of Experimental Psychology: Human Perception and Performance 24(5): 1546-1561.

Hebert, S. & L. L. Cuddy (2002). "Detection of metric structure in auditory figural patterns." Percept Psychophys 64(6): 909-18.

Povel, D. J. & P. Essens (1985). "Perception of Temporal Patterns." Music Perception 2: 411-40.

Ross, J. & A. J. Houtsma (1994). "Discrimination of auditory temporal patterns." Percept Psychophys 56(1): 19-26.

Tracking and reproduction of non-metric interval sequences: How training affects accuracy and representation

Guy Madison

Department of Psychology, Umeå University, Sweden

People are better at reproducing time intervals with 1:2 ratios than non-integer ratios. The question is why. Inasmuch as music as a human product reflects our capacities, it suggests a number of explanations. For example that we have an innate predisposition for using only two categories of time intervals (Fraisse, 1946; 1956), that integer ratios are easily combined to fit an underlying isochronous pulse (Essens & Povel, 1985), or that categorisation as such reduces information and thus facilitates processing. These and other alternative hypotheses can to some extent be tested. The present experiment addresses how the representation of sequences with non-integer ratios is affected by training. The task is to listen and try to “understand” a pattern consisting of 6 intervals that is continuously repeated, then to synchronise with the sequence, and finally to reproduce the pattern without any stimulus. First, the representation of the pattern is described in terms of the differences between the stimulus and the two response modes. Second, the accuracy of the responses is considered from an ecological perspective of what constitutes a perceptually acceptable reproduction. Third, these aspects are related to the amount of training. The main hypothesis is that non-integer ratios can be accurately reproduced with sufficient training, which would speak against an innate predisposition.

Keeping synchrony with complex rhythms while tempo changes.

Hans-Henning Schulze
Philipps Universität Marburg

Dirk Vorberg
Technische Universität Braunschweig

In a previous paper (1) we have studied synchronization with a metronome that changes tempo. Detailed analysis of the behavioral data within the framework of a two-level synchronization model revealed evidence for both both phase and period correction. In this paper we use the same paradigm, but instead of synchronizing with a metronome that was isochronous except for the tempo change, subjects tapped complex repeating rhythms while tempo changed. The question at which rhythmic level period correction occurs will be analyzed comparing the fit of different variants of the general model.

(1) Schulze, H.H., Cordes, A., and Vorberg, D. Keeping synchrony while tempo changes: Accelerando and ritardando. Music Perception, 22, 2005

Interval timing and trajectory formation in unequal amplitude movements

Michail Doumas[1, 2], Alan M. Wing[1] and Kelly Wood[1]

[1] Behavioural Brain Sciences Center, School of Psychology, The University of Birmingham, UK
[2] Department of Psychology, K.U. Leuven, Belgium

The Wing Kristofferson (WK) model of movement timing emphasises the separation of central timer and motor processes. Several studies of repetitive timing have shown that increase in variability at longer intervals is attributable to timer processes, however, relatively little is known about the way motor aspects of timing are affected by task constraints. In the present study, first we examined timing variability in finger tapping with differences in interval to assess central timer effects, and with differences in amplitude to assess motor effects. Then, we asked whether and in what way changes in kinematic aspects of performance have differential effects in timing of the two phases of the movement (flexion, extension). Eleven participants performed bimanual synchronous tapping, at two target intervals (400, 600 ms) with the index finger of each hand performing movements of equal (3 or 6 cm) or unequal amplitudes (Left Hand 3, Right Hand 6 cm, and vice versa). Timer variability increased linearly with the mean interval and showed small, non-systematic effects with changes in amplitude. Conversely, asynchrony (motor) variability was greater in unequal amplitude conditions, both at the point of response (flexion) but also in the extension phase. These results suggest that interval timing is triggered centrally, and that motor output timing is sensitive to changes in the manner of movement. A further analysis at the level of motor output timing revealed that asynchronies between flexion and the next extension were correlated, suggesting that the two phases are organised as a unit around the response. On the basis of these findings we propose that movement timing is organised hierarchically, with a higher level timer triggering responses at flexion, initiating a series of actions within a unit at the peripheral level, including flexion and the following extension.