



Provided by the author(s) and NUI Galway in accordance with publisher policies. Please cite the published version when available.

Title	Process timing and its relation to the perception of tonal harmony
Author(s)	Elliott, Mark
Publication Date	2002
Publication Information	Aksentijevic, A., Elliott, M. A., & Barber, P. J. (2002). Process timing and its relation to the perception of tonal harmony. In J. A. Da Silva, E. H. Matsushima, & N. P. Ribeiro-Filho (Éds.). Annual Meeting of the International Society of Psychophysics, (vol. 18). Rio de Janeiro, RJ, Brazil: Editora Legis Summa Ltda, (pp.161-166).
Publisher	Editora Legis Summa Ltda
Item record	http://hdl.handle.net/10379/1444

Downloaded 2023-03-21T08:13:16Z

Some rights reserved. For more information, please see the item record link above.



Process Timing and its Relation to the Perception of Tonal Harmony

Mark A ELLIOTT^{†‡} and Aleksandar AKSENTIJEVIĆ[#]

[†]School of Psychology, National University of Ireland Galway, Galway, Republic of Ireland

[‡]Center for Applied Perceptual Research, Department of Acoustic Design, Kyushu University
4-9-1, Shiobaru, Minami-ku, Fukuoka 815-0032, Japan

[#]Department of Psychology, Roehampton University, Roehampton Lane, London SW15 5PJ, UK

E-mail: [†]mark.elliott@nuigalway.ie, [#]A.Aksentijevic@roehampton.ac.uk

Abstract Recent advances in auditory research suggest that gamma-band synchronization of frequency-specific cortical loci could be responsible for the integration of pure tones (harmonics) into harmonic complex tones. Thus far, evidence for such a mechanism has been revealed in neurophysiological studies, with little corroborative behavioral data. Using psychophysical techniques, we observed a rate- and time- specific response-time (RT) advantage for a sequence of target pips when the defining frequency of the target was a fractional multiple of an entrainer frequency. The effect was only observed when entrainer and target tone-pip sequences were presented at 33 pips per second (pps) and when the inter-stimulus interval (ISI) was approximately 100 milliseconds (ms). This evidence implicates the oscillatory gamma-band activity in the representation of harmonic complex tones and suggests that synchronization with precise temporal characteristics is important for their integration. A model is presented, which accounts for these findings in terms of fast resynchronization of relevant neuronal assemblies.

Keywords gamma-band synchronization, auditory priming, tonal harmony, harmonic templates

1. Introduction

One of the important issues in auditory research concerns perceptual integration of spectral components of complex tones or harmonics, which represent the basis for speech and music perception. While several spectral components coexist within the harmonic series of a complex tone, the product of their grouping is a distinctive percept defined by a single dominant pitch. This “virtual” pitch is generally unaffected by the absence of the fundamental frequency from the stimulus spectrum [1]. Separate spectral pitches may be “heard out” when conscious effort is made to analyze the complex tone [2]. The dissociation between analytic and synthetic modes of perception also characterizes the processing of “illusory” visual percepts, such as Kanizsa figures [3]. Although the features of a Kanizsa figure are available to conscious inspection, the spontaneous impression is of a square or a triangle, contour discontinuities notwithstanding.

According to recent findings, synchronization of neuronal activity could play a crucial role in the perceptual integration, or binding, of harmonics into complex tones. Synchronization in the auditory cortex manifests itself as the auditory gamma-band response (GBR) that shows peaks in the brain magnetic field between 30 and 33 Hz [4]. Some authors [5] have linked

the GBR with the integration of harmonics into complex tones.

Generally, auditory gamma-band synchronization occurs at frequencies lower than the universally cited 40 Hz [6, 7]. Psychophysical research suggests that temporal quantization of perceptual events occurs in packets of around 30 msec duration (i.e., 33 Hz [8, 9]). In addition, task-specific 33-Hz oscillations have been observed in auditory reaction times [10]. On these grounds, it was hypothesized that the entrainment of frequency-specific neuronal assemblies in the lower gamma range (around 33 Hz) would co-activate the harmonically related assemblies [11]. Recently, anatomical evidence has been obtained for long-range plastic connections between harmonically related loci in the mammalian primary auditory cortex [12].

Consequently, detection mechanisms would differentiate between targets carrying fractional (inharmonic) and integral (harmonic) multiples of tones

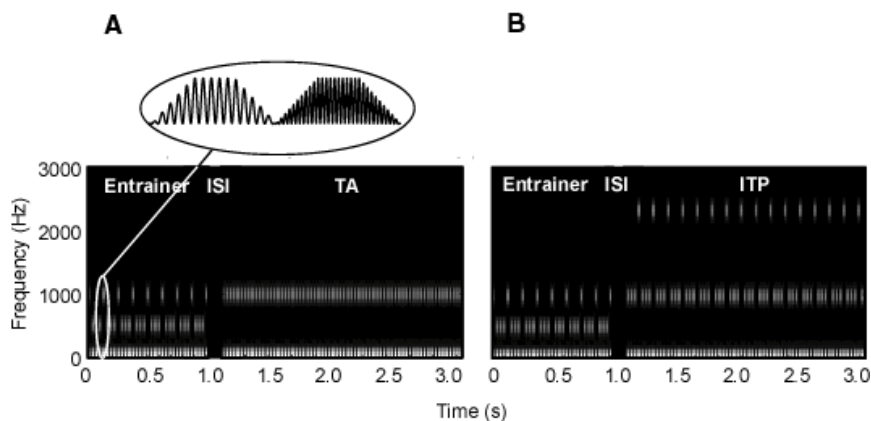


Figure 1. The spectral decomposition of stimuli used in Experiment 1. Illustrated stimuli presented at 33 pps. Both panels, left: An entrainer stimulus comprising a repeating sequence of 4 pips which carry 500 Hz, with the fourth pip (deviant) carrying 1000 Hz (see inset). In panel A, the entrainer is followed by an inter-stimulus interval (ISI) of 100 ms and a target-absent (TA) stimulus comprising only 1000-Hz pips. Panel B shows the entrainer followed by the same ISI and an inharmonic target-present (ITP) stimulus, which consists of a repeating 4-pip sequence carrying 1000 Hz, with every fourth pip carrying 2400 Hz. In the harmonic target-present condition (HTP; not shown) every fourth pip carried 2000 Hz. Inset: Two tone pips carrying 500 Hz and 1000 Hz respectively. The duration of the pips was 30.3 ms and the plateau was 33% of the overall duration.

presented in an entrainer stimulus (Figure 1). If the time required for stimulus identification depended on the coherence of the competing stimulus, inharmonic targets should be identified rapidly relative to harmonic targets, since the latter would require additional analysis and/or extraction from the persisting neuronal representation of the harmonic series induced by the entrainer. This processing cost could be described as analogous to the cost incurred in the extraction of features embedded in a visual Gestalt.

2. Experimental Evidence

Three principal patterns of effects emerged from the results of Experiment 1. Firstly, compared to the inharmonic target (ITP) stimuli, the harmonic target (HTP) stimuli were significantly more often confused with target-absent (TA) stimuli. The high error rate associated with HTP responses was consistent with the well-documented difficulty in discriminating octaves (the octave equivalence phenomenon; See e.g. [13]). Secondly, a quasi-sinusoidal pattern in the HTP RTs (with maxima at 32 and 36 Hz) suggested that detection of the HTP stimuli was sensitive to the position of the deviant pip in the repeating 4 tone-pip sequence relative to the offset of the entrainer.

RTs were slowed down at rates at which the offset of the entrainer was in phase with the slow rhythm imposed by the deviant (one quarter of the stimulus presentation rate). The difference in the height of the peaks was consistent with the interaction between the deviant-entrainer offset and deviant-target onset functions. Elsewhere, it has been suggested that this type of RT effect could reflect the resonance in processing mechanisms induced by periodic stimuli [14]. By contrast, the pattern of ITP responses showed no dependence on the periodic structure of the entrainer. Thirdly and most importantly, a significant advantage for ITP stimuli was observed at 33 pps but not at other rates (Figure 2A; Asterisk).

Despite the observed rate-specific advantage for ITP stimuli, the subjects could have based their responses on an audible interaction between presentation rate and target frequency. In other words, it was not clear whether the entrainer was responsible for the observed effect. In a second experiment, the temporal relationship between the entrainer and the target was examined by focusing on the ISIs from 0 – 106 ms, varied in 15-ms increments. If the inharmonic RT advantage were due to the perceptual characteristics of the target, no particular ISI should be favored. By contrast, an interval-specific effect would indicate that the entrainer was necessary for the effect to occur. The main finding was a significant inharmonic

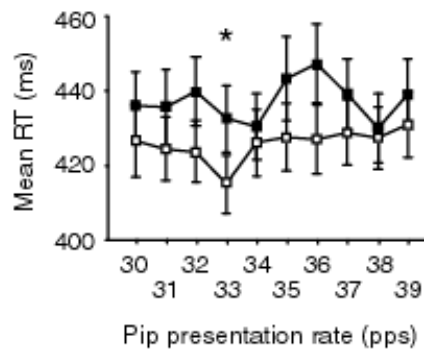
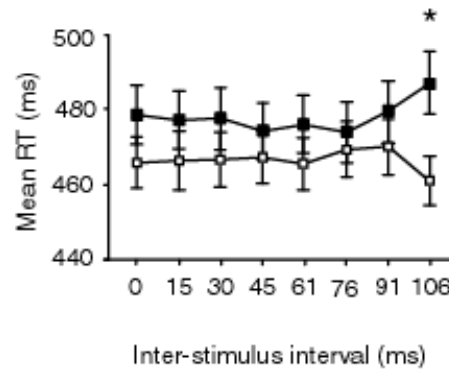
A**B**

Figure 2. (A) Results of Experiment 1. Mean HTP (filled) and ITP RTs and 99% confidence intervals plotted against presentation rate. A significant difference was observed only at 33 pps ($p < .0001$; asterisk). (B) Results of Experiment 2. Mean HTP (filled) and ITP RTs and 95% confidence intervals plotted against ISI. A significant difference was observed only at 106 ms ISI ($p < .0001$; asterisk).

advantage at 106 ms ISI (Figure 2B; Asterisk) with no effects at other intervals. Consequently, the rate-specific inharmonic advantage observed in Experiment 1 was due to the presence of the entrainer.

3. Model

The results of Experiment 2 allowed a more stringent specification of the underlying mechanism(s). The rhythmic structure of the entrainer induced a global steady state deriving from temporal superposition of consecutive transient oscillations. However, the synchrony pattern was modulated by the presence of two similar but distinct harmonic series involved in the encoding of the entrainer. The 500-Hz harmonic series includes most of the harmonics of the 1000-Hz series, but the opposite is not the case (Figure 3). Whereas the activation of the 500-Hz harmonic assembly leads to the co-activation of at least some parts of the 1000-Hz template, the activation of the 1000-Hz template by the deviant must lead to the desynchronization of the 500-Hz activity as a unique 1000-Hz representation needs to be extracted from the active 500-Hz template. The disruption of the 500-Hz synchrony results in a re-synchronization cycle that ends with the onset of the next deviant. There is reliable electrophysiological evidence for a mechanism for early change detection in audition. According to some research [15], mismatch negativity (MMN) is evoked by an infrequently presented stimulus ("deviant"), differing from the frequently occurring stimuli

("standards") in one or several physical parameters such as duration, intensity, or frequency. The MMN could be related to the desynchronization of active representations in the presence of novel stimuli.

The virtual 1000-Hz representation was cyclically refreshed by the deviant pips and maintained by the 500-Hz standards. The sinusoid-like modulation of HTP RTs observed in Experiment 1 suggested that the virtual harmonic representation of the 1000-Hz series was quantized into cycles delimited by successive deviants. If a harmonic target (HTP) was locked into the rhythm imposed by the entrainer, the discrimination between harmonically related target present/absent (i.e. TA and HTP) stimuli became more difficult. This effect should be contrasted with the fast re-synchronization cycle of the 1000-Hz harmonic representation, which was regularly reset by the deviant. Experiment 1 showed that the discrimination between harmonic and inharmonic targets was facilitated at 33 pps and that the effect reached significance at ~ 100 ms ISI. The inspection of individual data showed that out of 16 subjects, two showed a maximum enhancement at 30 pps, five at 32, six at 33 pps and three at 34 pps. One subject's maximum enhancement was at 37 pps. Figure 4 shows a possible model of fast cumulative entrainment of the 1000-Hz representation, reset by the deviants. The successive 500-Hz pips increase the synchrony between the loci coding the 1000-Hz representation. The maximum coherence value is

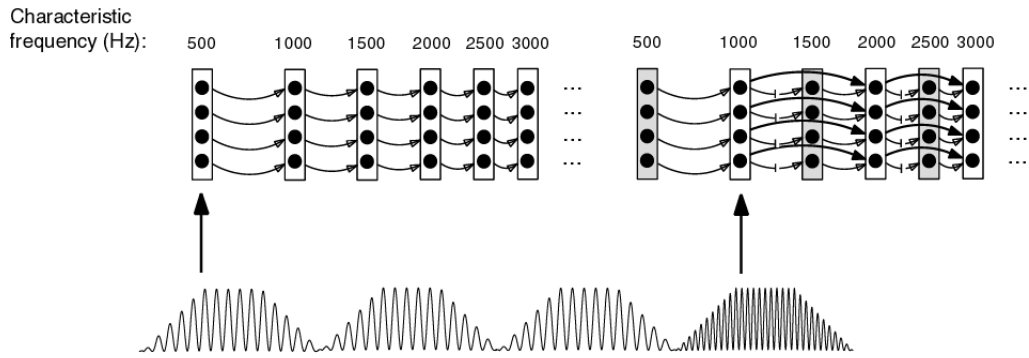


Figure 3. The asymmetric activation pattern for two overlapping harmonic templates. Left, a 500-Hz pip activates the 500-Hz template, which includes components of the 1000-Hz series. Right: A 1000-Hz deviant activates only the 1000-Hz template. To be encoded as a distinct percept, a 1000-Hz tone must inhibit the 500-Hz template-specific activity (shaded).

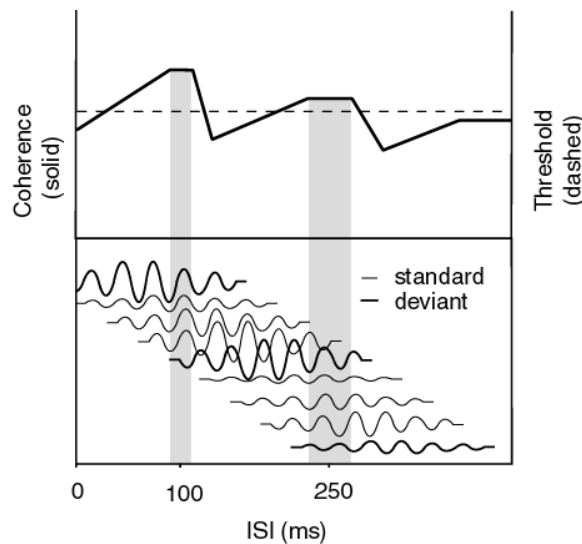


Figure 4. Temporal effects of fast re-synchronization: lower panel: The horizontal axis represents the entrainer-target ISI. Waveforms represent the residual oscillatory activity following an entrainer presented at 33 pps. The straight segment at the beginning of each waveform represents a fixed onset delay of 10-15 ms. The deviant-induced oscillation (thick curve) is out of phase with the preceding standard-induced oscillations, reflecting a temporary disruption of the fast cycle by the deviant. Upper panel: Oscillatory coherence of the residual oscillations as a function of ISI. The phase interaction of successive oscillations creates coherence windows centered at approximately 100 ms and 250 ms ISI. (The presence of the second window was confirmed by three experiments not reported here). Synchronous activity induced by sub-optimal presentation rates fails to reach the threshold.

reached at the time of the onset of the deviant. The coherence pattern persists after the offset of the entrainer. If the inharmonic target falls within the coherence window created by the structure of the entrainer, the harmonic gestalt is strengthened to the extent that discrimination between harmonic and inharmonic targets is facilitated beyond the degree achieved through

perceptual confusability of TA and HTP stimuli. The novel factor is the increased salience of the inharmonic target (as opposed to the reduced salience of the harmonic target, responsible for the modulation in HTP RTs). The upper panel of Figure 4 demonstrates the decay of the synchrony pattern and the accompanying change in the shape of the coherence window. The optimal

synchronization frequency appears to lie around 33 Hz, supporting the available physiological and psychophysical evidence.

4. Conclusions

These experiments provide psychophysical evidence for rate-specific discrimination between tones according to their harmonic structure. They agree with the available neurophysiological hypotheses linking synchronized gamma-band activity and the integration of complex tones. The mechanism involved in the integration of harmonic complex tones appears to involve fast synchronization of harmonic templates at approximately 33 Hz. It takes approximately 100 ms (or longer) for the oscillatory representation to reach the suprathreshold level of activation at which it can be discriminated from a distinct harmonic representation. More generally, psychophysical evidence for a similar mechanism in visual binding supports the notion of gamma-band synchronization as a supra-modal correlate of feature integration [16].

References

- [1] E. Terhardt. Warum hören wir Sinustöne? *Naturwissenschaften*, 76, 496-504 (1989).
- [2] H. L. F. von Helmholtz. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. English translation A. J. Ellis. New York: Dover (1954).
- [3] G. Kanizsa. *Organisation in Vision*. New York: Praeger (1979).
- [4] C. Pantev, & T. Elbert. The transient auditory evoked gamma-band field. In C. Pantev, T. Elbert, & B. Lütkenhöner (Eds.) *Oscillatory Event-Related Brain Dynamics*. New York: Plenum Press, 219-229 (1984).
- [5] C. Pantev, S. Makeig, M. Hoke, R. Galambos, S. Hampson, & S. Gallen. Human auditory evoked gamma-band magnetic fields. *Proc. Natl. Acad. Sci. USA*, 88, 8996-9000 (1991)
- [6] M. Joliot, U. Ribary, & R. Llinas. Human oscillatory brain activity near 40 Hz coexists with cognitive temporal binding. *Proc. Natl. Acad. Sci. USA*, 91, 11748-11751 (1994).
- [7] A. Knief, M. Schulte, O. Bertrand, & C. Pantev. The perception of coherent and non-coherent auditory objects: a signature in gamma frequency band. *Hear. Res.*, 145, 161-168 (2000).
- [8] D. Schwender, C. Madler, S. Klasing, K. Peter, & E. Pöppel. Anesthetic control of 40-Hz brain activity and implicit memory. *Conc. Cogn.*, 3, 129-147 (1994).
- [9] E. Pöppel. A hierarchical model of temporal perception. *Trends Cognit. Sci.*, 1, 56-61 (1997).
- [10] S. Dehaene. Temporal oscillations in human perception. *Psych. Sci.*, 4, 264-270 (1993).
- [11] J.P. Makela, & R. Hari. Evidence for cortical origin of the 40 Hz auditory evoked response in man. *Electroencephalogr. Clin. Neurophysiol.*, 66, 539-546 (1987).
- [12] S.C. Kadia, X. Wang, L. Liang, & D.K. Ryugo. Harmonic structure in long-range horizontal connections within the primary auditory cortex of cat. *Soc. Neurosci. Abs.*, 29 (1999).
- [13] E.M. Burns, & W.D. Ward. Categorical perception - Phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *J. Acoust. Soc. Am.*, 63, 456-468 (1978).
- [14] R. Kompass, & M.A. Elliott. Modeling as part of perception: A hypothesis on the function of neural oscillations. In E. Sommerfeld, R. Kompass, & T. Lachmann (Eds.), *Proceedings of the Seventeenth Annual Meeting of the International Society for Psychophysics*, 130-135. Berlin: Pabst (2001).
- [15] R. Näätänen. Mismatch negativity as the index of central sound representation in the human brain. *Int. J. Psychophysiol.*, 25, 78-78 (1997).
- [16] M.A. Elliott, & H.J. Müller. Evidence for 40-Hz oscillatory short-term visual memory revealed by human reaction-time measurements. *J. Exp. Psychol. Learn.*, 26, 1-16 (2000).