Title	Process timing and its relation to the coding of tonal harmony.
Author(s)	Elliott, Mark
Publication Date	2010
Publication Information	Aksentijevic, A., Barber, P. J., & Elliott, M. A. (2010). Process timing and its relation to the coding of tonal harmony. Journal of Experimental Psychology: Human Perception and Performance (In press).
Publisher	American Psychological Association
Item record	http://hdl.handle.net/10379/1691

Process timing and its relation to the coding of tonal harmony

[1,2]Aleksandar Aksentijevic [2]Paul J. Barber

and [,3]Mark A. Elliott

[1]Department of Psychology, Roehampton University, UK

[2]Department of Psychological Sciences, Birkbeck College, University of London, UK

[3]School of Psychology, National University of Ireland, Galway, Ireland

Journal of Experimental Psychology: Human Perception and Performance

2011 (in press)

Abstract

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 psychophysical evidence. In 6 experiments, we observed a rate- and time- specific response-time advantage for a sequence of target pips when the defining frequency of the target was a fractional multiple of a priming frequency. The effect was only observed when the prime and target tone-pip sequences were presented at 33 pips per second (pps) and when the inter-stimulus interval was approximately 100 and 250 milliseconds (ms). This evidence implicates oscillatory gamma-band activity in the representation of harmonic complex tones and suggests that synchronization with precise temporal characteristics is important for disambiguating related harmonic templates. An outline of a model is presented, which accounts for these findings in terms of fast resynchronization of relevant neuronal assemblies.

Keywords: Gamma-band oscillations, harmonic templates

Process timing and its relation to the coding of tonal harmony

One important issue 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 (Terhardt, 1989). Separate spectral pitches may be “heard out” when conscious effort is made to analyze the complex tone (von Helmholtz, 1863/1954).

According to an influential approach (e.g. Singer, 1993, 1994, 1999) neural assemblies formed via synchronized oscillatory activity in the gamma band (at around 40 Hz) are thought to be responsible for the perceptual integration of spatially distributed stimulus features. In vision, neurons in layers II and III of primary visual cortex, which are connected to cells with similar feature coding preferences, have been found to participate in dynamic synchronized assemblies (Eckhorn et al., 1988, 1990; Gray & Singer, 1989). The same type of cellular organization has been observed in the primary auditory cortex (harmonic connections between pyramidal cells; Kadia & Wang, 2003), which, when considered alongside the supra-modal nature of gamma-band synchronization (Kahana, 2006), indicates that harmonic binding in audition could involve a similar mechanism.

The human auditory 40-Hz activity was brought to broader attention by Galambos, Makeig and Talmachoff (1981) who reported the 40-Hz auditory steady-state (SSR) response – a sustained oscillatory response to sequences of clicks (Hari et al., 1989), tone-pips (Gutschalk, Oldermann & Rupp, 2009) and amplitude-modulated tones (Picton, Skinner, Campagne, Kellet, & Maiste,1987). Auditory SSRs have been shown to represent the result of summation of individual transient gamma-band responses (Bohorquez & Ozdamar, 2008). The transient auditory gamma-band activity is generally categorized into evoked and induced responses. According to Galambos (1992), the “evoked” gamma-band response (eGBR, also known as the auditory GBR or “early” 40-Hz response) represents a burst of four or five periods of a sinusoid-like oscillation that begins about 20 ms after the onset of the stimulus and is tightly time-locked to it. The response peaks at approximately 60 to 80 ms post-onset and disappears completely at between 120 and 150 ms (Makeig & Galambos, 1989; Pantev et al., 1998). The source of the eGBR is in the auditory cortex (Makela & Hari, 1987; Pantev et al., 1991). The late "induced" gamma-band response (iGBR) appears at approximately 200-400 ms post-onset and is not phase-locked to the stimulus (Galambos, 1992). Different authors have linked the eGBR with the deployment of attention (Tiitinen et al., 1993) and exchange of information between different cortical areas (**S**enkowski, Talsma, Grigutsch, Herrmann, & Woldorff, 2007).

Although the electrophysiological gamma-band is broadly located between 30 and 100 Hz, gamma-band synchronization is commonly associated with the frequency of approximately 40 Hz. However, there are numerous indications that gamma-band synchronization in the auditory system occurs at frequencies lower than 40 Hz. In a study by Stapells, Makeig and Galambos (1987), 4 out of 10 participants demonstrated SSR amplitude-maxima at presentation rates of less than 30 pips per second (pps). It should be noted that values of pps and Hz are equivalent. The former is used when referring to a specific kind of stimulus (pip train) and the later when describing brain oscillations and tonal frequencies. Evidence from the recordings of transient gamma-band activity also suggests that the optimum frequency might lie between 30 and 35 Hz. The auditory gamma-band activity has been referred to as “near-40-Hz activity” by Joliot, Ribary and Llinás (1994), who observed gamma-power peaks around 35 Hz. Similarly, auditory gamma peaks in the brain magnetic field have been reported at 30 – 33 Hz (Pantev & Elbert, 1994). An almost identical frequency profile was obtained in a biomagnetic study of the role of auditory gamma-band synchronization in the integration of pure tones (Knief, Schulte, Bertrand & Pantev, 2000). According to Makeig (1993), the onset of tones or clicks generates a burst of 30-35 Hz of gamma-band activity in the auditory cortex. Pulvermüller et al. (1994) reported a coherent 30-Hz activity related to word processing, while Pantev et al. (1991) observed an eGBR spectral peak around 33 Hz.

Neural oscillations at frequencies around 33 Hz could be associated with a more general perceptual and cognitive function. A model by Jeffreys, Traub and Whittington (1996) demonstrated that embedding an isolated interneuron in a network of cells coupled by $GABA_a$-mediated inhibitory postsynaptic potentials would cause it to oscillate at 33-Hz (see Elliott, Giersch & Seifert, 2006 for evidence of involvement of $GABA_a$ in the modulation of visual gamma-band oscillations). Palva et al. (2002) used MEG to investigate the eGBR evoked by speech and non-speech sounds. The authors observed significant differences in lateralization and timing of the speech and non-speech-triggered eGBRs. Intriguingly, the peak frequency for speech sounds in the left hemisphere was around 33 Hz. Pantev et al. (1991) reported transient auditory gamma-band activity, which clearly peaked between 32 and 33 Hz. This is in agreement with the hypothesis, corroborated by psychophysical research, according to which temporal quantization of perceptual events occurs in packets of approximately 30 ms (i.e., 33 Hz; see Pöppel, 1997). A study by Schulte, Knief, Seither-Preisler and Pantev (2000), which investigated the eGBRs associated with three types of pitch (spectral, harmonic and inharmonic), also found highest synchronization at approximately 33-35 Hz.

Auditory gamma-band activity represents an important supra-modal manifestation of the perceptual function of the cortex (Başar-Eroglu, Strüber, Schürmann, Stadler, & Başar, 1996). In light of the apparent importance of the gamma-band oscillatory activity in perceptual integration (e.g. Singer, 1999) and the documented similarities between the neurophysiological correlates of auditory and visual processing, a question arises as to the potential role of the eGBR in harmonic "feature" integration. To illustrate, Galambos and Makeig (1992) suggested that its role could consist in comparing and integrating the channel-specific parameters of the peripheral stimulus situation prior to the analysis by the neighboring neuronal aggregates. Pantev et al. (1991) proposed that the function of the eGBR could be to combine segregated auditory features into unified percepts, or more specifically, to integrate separate cortical representations of harmonics into the pitch of a complex tone. Here, it is claimed that harmonic pitch is represented as a spatially distributed pattern of active neuronal assemblies. Spatially distributed brain structures responsible for encoding harmonic relations (known as harmonic templates) have been hypothesized (Shamma & Klein, 2000) and discovered in mammalian brain (Kadia & Wang, 2003).

Based on the evidence for oscillatory gamma-band activity in the auditory cortex (Galambos et al., 1981; Makela & Hari, 1987), we hypothesized that the external gamma-band entrainment of frequency-specific loci would activate harmonically related templates, thus accentuating the perceptual difference between the integral and fractional multiples of the fundamental frequency. Since individual harmonics are related to the fundamental by a frequency ratio of the form 1: n (where n represents the position of a harmonic within the series), pure tones related by integral ratios should be processed differently relative to inharmonically related tones. Although the latter describe various relationships within the series (e.g. the relationship between the fifth and the sixth harmonic), the spectral structure of the series itself can only be specified in terms of the relationship of the harmonics with the fundamental frequency. If the grouping of harmonics reflects the early synthesis of complex tones, the entrainment of the harmonic representation within the lower gamma band should facilitate the extraction of inharmonic tones (see e.g. Schellenberg & Trehub, 1996a, 1996b for evidence of the perceptual asymmetry of harmonic and inharmonic stimuli).

We investigated the possibility that harmonic relationships are instantiated in the brain in the form of short-term oscillatory representations. We predicted that such representations should (i) be associated with a narrow range of frequencies within the gamma band; (ii) reflect activation of distinctive harmonic templates (Shamma & Klein, 2000; Kadia & Wang, 2003) and; (iii) have a precise time course that is closely related to the temporal parameters of the eGBR. The hypothesis implicating the evoked gamma-band oscillations in the integration of harmonics is based on the following assumptions. Firstly, substantial evidence suggests that gamma activity is associated with a supra-modal mechanism of feature integration (Tallon-Baudry & Bertrand, 1999; Kaiser & Lutzenberger, 2005). A cortical or thalamo-cortical oscillatory mechanism could synchronize the firing patterns within harmonic templates by binding distinct harmonic loci. It should be stressed that this hypothesis relates only to resolved harmonics and makes no claims with regard to phenomena such as periodicity, residue or inharmonic pitch.

The second assumption is that this mechanism can be studied experimentally. The motivation for the current research comes partly from the work on the effects of visual gamma entrainment on the grouping and segmentation of illusory figures in vision as carried out by Elliott and Müller (1998, 2000), who demonstrated that experimental manipulation of gamma-band synchronization yields clear and stable psychophysical effects. In the experiments reported here, we used trains of brief pure tones (pips) which are known reliably to evoke auditory SSRs (Galambos, Makeig and Talmachoff, 1981). Furthermore, Ross, Herdman & Pantev (2005a) have demonstrated that the auditory SSRs show clear right-hemisphere preference, which the authors associated with pitch processing in the primary auditory cortex.

The rationale for using pure tones was as follows. If multi-peaked cells (Kadia & Wang, 2003) can be activated by a number of harmonically related frequencies, harmonically related cells could be activated by a single frequency. A pure tone could activate and synchronize a harmonic template with the fundamental frequency corresponding to the frequency of the stimulus. In the remainder of the text, this is referred to as the "unique-template" hypothesis. Here, a more detailed description of this hypothesis is given.

Theoretically, a 500-Hz pure tone could belong to any number of harmonic representations. It could be a second harmonic of a 250-Hz representation. Here, one could expect a 500-Hz locus to be connected not only to its higher harmonics (1000, 1500… etc Hz), but also to 250 and 750 Hz as well as 1250, 1750… etc Hz). These frequencies are harmonics of a 250 Hz series but not the 500-Hz series. Equally, 500 Hz can be seen as the tenth harmonic of a 50-Hz series. Yet, a 500-Hz pure tone and harmonic series elicit pitch equivalent to 500 Hz and not to any of the lower frequencies. The function of harmonic templates is likely to be governed by a directional preference for higher harmonics of the fundamental frequency. We propose that the pitch corresponding to 500 Hz is elicited by a pure tone as well as by a 500-Hz harmonic series because the harmonic template is activated "from below", that is, from the fundamental upwards. For a 500-Hz locus to elicit a pitch equivalent to 250 Hz, another component needs to be activated at either 250 or 750 Hz. It follows that a harmonic series is uniquely determined by the position of its fundamental frequency. When a 500-Hz pure tone is presented, it activates the entire template (following Shamma & Klein, 2000). If a second harmonic is added, the pitch remains unchanged. Adding further harmonics does not change the pitch as long as the distance between any two adjacent components is the same (spectral regularity constraint). Even if the fundamental component is omitted, the pitch will not change, but will now depend on the number of co-active sites. The larger the number of active harmonics, the stronger the pitch sensation will be (von Helmholtz, 1863/1954).

Figure 1 here

Following this, a sequence of 500-Hz tones should activate both 500-Hz (fundamental) and 1000-Hz (second harmonic) templates. Subsequent presentation of a 1000-Hz tone should activate the 1000-Hz template and desynchronize the 500-Hz template (see Figure 1). Synchrony is defined in terms of temporal (phase) alignment of individual oscillations which can (but does not have to) result in the increase in the power of a corresponding electrophysiological response. Similarly, desynchronization refers to the temporal misalignment of individual oscillations which results in decreased coherence.

Such a mechanism would be necessary for disambiguating spectrally similar auditory percepts and could be involved in distinguishing harmonic complexes an octave apart. The oscillatory cycle activated by the 500-Hz tones would have to be interrupted by the onset of a deviant tone because although the components of a 1000-Hz template are found within a 500-Hz template, the former must represent a distinct structure. Otherwise, the discrimination between a 500 and 1000-Hz complex tone would not be possible.

Figure 2 here

To illustrate, three successive 500-Hz tones generate an oscillatory cycle. Individual oscillations are phase aligned so that each successive pip increases the coherence of the representation. In order for a 1000-Hz representation to become established, odd harmonics of 500 Hz (1500, 2500 etc…Hz) must be inhibited. The deviant pip is immediately followed by another triplet of 500-Hz pips, which re-activates the 500-Hz template (via the template-specific SSR) at the expense of the 1000-Hz representation. This time the disruption of the oscillatory cycle would not be as pronounced, because no inhibition needs take place. All that is needed is for the inhibited odd harmonics to be incorporated into the active 500-Hz template. The amplitude of such a representation would be asymmetrical. This mechanism, henceforward called “Harmonic Reset” (HR), is illustrated in Figure 2. It creates an oscillatory representation that waxes and wanes periodically. It is assumed that the properties of this representation could be inferred by manipulating temporal parameters of the stimuli.

In part one (Experiment 1), we describe a frequency-specific gamma-band oscillatory representation, and in part two (Experiments 2, 3 and 4), we confirm the frequency specificity of the representation and describe its time course. In part three (Experiments 5 and 6), we demonstrate that the oscillatory harmonic representation recurs, possibly as a consequence of resynchronization by a resonant loop. Following this, we propose a quantitative model of oscillatory harmonic binding and discuss our finding in the context of current theory.

General method

Since all experiments reported here were methodologically similar, we give the general description of the method here and deal with differences in the reports of individual experiments.

Stimuli

Although the auditory eGBR and SSR can be evoked by clicks (Galambos et al., 1981) or by amplitude and frequency modulated tones (Picton, et al., 1987), tone-pips (Galambos & Makeig, 1992; Gutschalk et al. 2009) appear particularly well suited to the requirements, because they allow easy manipulation of presentation rate and carrier frequency. Furthermore, the fact that tone-pips have been used in numerous neurophysiological studies should facilitate the methodological integration of the current paradigm.

Figure 3 here

Each trial consisted of two tone-pip trains. Individual pips were onset- and offset-cosine-ramped [average slope 5 decibels per millisecond], and had a plateau of 33% of the overall period. The first train (entrainer) consisted of a repeating four-pip (DBBB) sequence, with every first pip (deviant) carrying 1000 Hz, and the remaining three (baseline) pips carrying 500 Hz. The duration of the entrainer was 1000 ms. Figure 3 illustrates how experimental trials were constructed. A 3000-ms train is cut into two parts – a 1000-ms entrainer and a 2000-ms target. At the presentation rate of 33 pips-per-second (pps), the 1000-ms cut-off point coincides with the end of a deviant pip (D). The longer “target” portion is shifted in time by the value of the inter-stimulus interval (ISI). Target stimuli came in two forms. A target-absent (TA; not shown) train carried only 1000-Hz pips and was presented on 50% of all trials. The two forms of target-present (TP) stimulus were presented with equal probability on the remaining 50% of trials. Both harmonic target-present (HTP or “harmonic target”) and inharmonic target-present (ITP or “inharmonic target”) stimuli mirrored the structure of the entrainer in that triads of baseline pips carrying 1000 Hz were combined with deviant pips carrying one of two higher frequencies. The HTP deviant frequency was 2000 Hz, whereas the ITP deviant frequency was 2400 Hz. Thus, HTP stimuli were harmonically related to the entrainer (2000 Hz being the fourth harmonic of 500 Hz and the second harmonic of 1000 Hz) as well as the target baseline frequency (1000 Hz). On the other hand, the ITP stimulus was inharmonically related to the entrainer frequencies and the target baseline frequency, since the ratio between the target deviant and entrainer deviant frequencies is 5:12 (an octave plus a minor third). A rating signal detection experiment (Swets, 1996, p. 37) investigated participants’ awareness of the harmonic relationship between entrainers (500 & 1000 Hz or 600 & 1200 Hz) varied relative to targets (1000 & 2000 Hz or 1200 & 2400 Hz). Ratings from 12 participants were analyzed using A_z index (Dorfman & Alf, 1969) with a mean A_z of .602 suggesting that participants were generally unaware of the harmonicity of the entrainer-target relationship.

Design

Each participant was tested in all conditions. All experiments contained 40 trials per condition. Trials were presented in blocks of 40 and each session commenced with a 40-trial practice block. Experimental conditions were completely randomized (within and between blocks), except in Experiment 1 where presentation rate conditions were randomized between blocks. The design was asymmetrical – two levels of harmony (HTP, ITP) were presented at one level of target (TP - target-present). Specifically, the proportion of trials was: TA = 50%, TP = 50% (HTP = 25%, ITP = 25%).

Apparatus and procedure

The experiments were conducted in a sound-attenuated room with background noise level less than 30 dB SPL. Experimental trial and stimulus generation, as well as response timing was controlled by an IBM compatible computer with sound signals generated by custom-built soundcard drivers of 32-bit-per-channel resolution driving a conventional PC SoundBlaster card at 44.1 kHz sample resolution. The software controlling the experiment execution was programmed in C. The stimuli were presented diotically via Beyer Dynamic DT 150 headphones. The ambient luminance was kept at a constant, low level in order to provide a non-distracting experimental environment. Stimulus intensity was held constant throughout (average 50 dB SL; A weighting), although the peak intensity varied slightly with the change in target frequency (approximately 1 dB). This intensity level was chosen for the following reasons. According to Galambos et al. (1981), the SSRs are recorded reliably at sound intensities over 15 dB SL, but achieve stability at approximately 40dB SL (Galambos and Makeig, 1992). At the same time, transient sounds presented at intensities above 50dB tend to produce aural harmonics, which can have a significant deleterious effect on fine auditory discrimination (Ritsma, 1962). The intensity of the stimuli was examined using an artificial ear and a Dawe D-1422C Sound Level Meter, (A weighting) and the shape of the stimuli was examined using a Tektronix 7313 oscilloscope and Matlab graphing functions. Responses were recorded from a pair of custom-built response keys.

Before each session, participants were asked to complete a brief questionnaire stating their age, gender, handedness and any potential hearing problems. At the beginning of each block, participants were instructed to press any key on the computer keyboard in order to initiate the trial sequence. Each trial was preceded by an alert in the form of a luminous cross on the monitor screen lasting 500 ms. After a fixed 500-ms fore period, the experimental stimuli were presented. Participants were instructed to attend but not to respond to, the first tone-pip train and focus on the second (target) tone-pip train. The task consisted in responding rapidly and accurately to the presence (TP) or absence (TA) of two alternating tones in the target, by pressing one of two response keys. Response times (RTs) were recorded from the target onset to the key press (not release). All participants made ”target” responses with the preferred hand and were encouraged to keep their error rate below 5% (two errors per block). Erroneous responses generated feedback in the form of the word “error” appearing for 500 ms on the monitor. At the end of each block, a visual prompt informed the participants of their progress by displaying the number of completed blocks, and instructed them to initiate a new block after a 5-second break. The participants were required to keep their eyes open throughout the experiment in order to minimize the potential interference by alpha activity (Galambos, 1992).

Trials in which RTs had a z value of less than –2.5 or more than 3 were removed prior to analysis. In addition, the first two responses in each block were removed in order to eliminate atypical RTs caused by lack of preparation. Error analyses were performed on arcsine transformed scores (Howell, 1992). Data were analysed using repeated-measures ANOVA procedure in SPSS. Greenhouse-Geisser correction for sphericity violations was applied in all analyses. All participants provided informed consent. The study was approved by the Ethics Board of the School of Psychology, Birkbeck College in accordance with the guidelines provided by the British Psychological Society.

Experiment 1: Rate specificity

The aim of Experiment 1 was to investigate the oscillatory frequency specificity of the HR model. We hypothesized that targets that were inharmonically related to the entrainer would be processed more rapidly and detected more easily relative to harmonic targets at tone-pip presentation rates around 33 pps. In addition, in an earlier experiment, not reported here, we investigated rates of 26, 33, 40 and 47 pps and observed RT facilitation for inharmonic targets only at 33 pps.

Method

Participants. Sixteen participants with normal hearing (11 female, mean age 23.8 years) took part in Experiment 1. All participants were right handed.

Stimuli. Stimuli were presented at ten rates, from 30 to 39 pps. With respect to presentation rate, the duration of individual pips ranged from 33.3 ms at 30 pps, to 25.6 ms at 39 pps. At the presentation rate of 30 pps, the entrainer contains 7.5 DBBB cycles (deviant plus three baseline pips), whereas at 39 pps, there are 9.7 such cycles. The ISI between the entrainer and target was 100 ms.

Design and procedure. Altogether, there were 30 experimental conditions - two levels of Target (TA, TP), with two levels of Harmony (HTP, ITP) embedded into the latter Target level and ten levels of Presentation Rate (30-39 pps). Each combination of Rate and Harmony was allocated 40 trials, giving 1600 trials per participant (TA = 800 trials, TP = 800 trials). The rationale of the experiment was not immediately apparent to the participants, since the two Harmony conditions were embedded in the TP condition and not explicitly mentioned in the instructions. The experiment was divided into two 40-minute sessions run on different days.

Results

All participants reported that the entrainer and TP stimuli were perceived as “flutter”, with two frequencies clearly audible. They reported the sensations accompanying stimulus presentation as pleasant and even calming.

Figure 4 here

A repeated-measures ANOVA with factors Target (TA, TP) and Rate (30 – 39 pps) was performed on the correct RTs. The analysis revealed a marginally significant main effect of Target [F (1, 15) = 3.89, p = .067; MSE = 2174.85] reflecting slower responses to TA stimuli, as well as a significant Target by Rate interaction [F (9, 135) = 2.11, p = .033; MSE = 139.06]. The main effect of Rate was nonsignificant (p > .1). The significant interaction was explored by means of a second ANOVA for the factors Rate and Harmony. As illustrated in Figure 4, there was a highly significant difference between harmonic target-present (HTP) and inharmonic target-present (ITP) responses [F (1, 15) = 21.47, p < .001; MSE = 537.82], with the former being markedly slower across all presentation rates. The main effect of Rate and the Rate by Harmony interaction did not reach significance ($p > .1$ for both). In order to test the principal hypothesis, a series of Bonferroni-adjusted t-tests was carried out on the HTP-ITP differences. The tests revealed a highly significant ITP advantage at 33 pps [$t\ (15) = 6.08$, $p < .001$]. Harmonic target-present responses were marginally slower than inharmonic target-present responses at 36 pps [$t\ (15) = 3.31$, $p = .005$]. Inspection of effect sizes indicated that the effect at 33 pps ($\eta_p^2 = .71$) was almost twice as large as the effect at 36 pps ($\eta_p^2 = .38$). No significant effects were observed at other rates.

The finding of a rate-specific RT advantage for inharmonic targets at 33pps was in agreement with the evidence reviewed above which shows that auditory gamma-band activity has a characteristic frequency of around 33 Hz. There were neither substantial nor reliable differences between RTs to harmonic and inharmonic targets at neighboring frequencies demonstrating that the difference at 33 PPS is both narrowly tuned and consistent to most participants. ITP responses showed a highly prominent decrease at 33 pps, whereas the marginally significant advantage at 36 pps was due to an increase in HTP RTs. The difference between the effects was also quantitative. The effect at 33 pps was almost twice as large as that recorded at 36 pps. Unlike the ITP function, which was relatively flat with the exception of the prominent trough at 33 pps, the HTP function described a sinusoid-like contour, with maxima at 32 and 36 pps. In order to confirm this observation, the HTP response curve was subjected to an exploratory fitting procedure. The best fit was provided by a sinusoidal model of the form: $a + b*\cos(c*x – 2.5)$ ($a= 437.6$, $b = 6.2$, $c = 1.5$; $r^2 = .66$, $p = .004$). The ITP response function could not be fitted by a sinusoidal model ($p > .5$). This suggested that the marginal enhancement at 36 pps was associated with a periodicity related to the surface stimulus properties.

Figure 5 here

In order to corroborate this, the HTP response curve was compared with theoretical functions describing the relationship between the duration of four-pip cycle (DBBB) at different rates and the duration of the pip train (Figure 5). The dashed sinusoid describes the phase relationship between the DBBB cycle and the entrainer offset and the solid one the relationship between the DBBB cycle and target onset. The two functions are well aligned at 36 pps and less so at 32 pps. There is a clear inverse relationship between the alignment of the two theoretical functions and the HTP responses, especially with respect to target onset.

The overall proportion of errors was 4.6%. Arcsine transformed error data were subjected to the same analyses, as were RT scores. The ANOVA for factors Target and Rate revealed only a marginally significant effect of Rate [F (9, 135) = 1.938, p = .052; MSE = .08], which was due mainly to a peak at 38 pps in both target conditions. A second ANOVA, performed on TP scores only, yielded a highly significant main effect of Harmony, with HTP responses eliciting more errors [F (1, 15) = 20.47, p < .001; MSE = .62]. Critically, the main effect of Rate and Rate by Harmony interaction were nonsignificant.

Discussion

The principal finding of Experiment 1 was a significant RT advantage for inharmonic targets at a rate of 33 pps. No simple explanation could be found for such an effect. The evidence for generalized periodic behavior generally associated with cognitive or motor processing (Treisman, Faulkner, Naish & Brogan, 1990; Pöppel, 1997) cannot account for the fact that a rate-specific advantage is observed with targets, which are in a specific frequency relationship with the entrainer. Nor could the answer be sought in the perceptibility of specific presentation rates, since just noticeable differences for modulation frequency increase smoothly up to 320 Hz independently of carrier frequency (Lee, 1994). Rather, the answer should be sought in the relationship between the temporal properties of the stimuli and the oscillatory binding mechanisms.

With TA stimuli a certain amount of time (of the order of tens of ms) is needed to compare the periodic structures of the target and the entrainer. A TA response is made as soon as the participant is confident that the target envelope is flat. Alternatively, the entrainer confers an advantage for TP stimuli which also possess an audible envelope. This advantage is neutralized for HTP stimuli, which carry two harmonically related frequencies an octave apart. Octave has a special perceptual status in that sounds differing by an octave are easily confused (e.g. Burns & Ward, 1978). Consequently, a HTP stimulus could be confused with a flat TA stimulus since the octave relationship between its frequencies could make its envelope less salient. HTP responses describe a sinusoid with maxima at 32 and 36 Hz due to an interaction between the effect of stimulus structure and influence of an entrained representation. Change in presentation rate within a train of fixed duration means that the entrainer offset coincides with different points of a DBBB cycle. At 32 and 36 pps, the entrainer offset corresponds to the end of the cycle (zero phase angle; solid sinusoid in Figure 5), so that the participants receive an integral number of DBBB cycles (eight and nine respectively). The angular function of the DBBB cycle in relation to target onset (dashed sinusoid) also approaches zero at 36 pps. The two functions are almost perfectly aligned at 36 pps and less so at 32 pps. It appears that participants relied on temporal cues to distinguish between TA and HTP stimuli.

Fine temporal cues are not as important if a comparison is being made between harmonically distinct stimuli. In this case, the coherence of the oscillatory representation becomes relevant. When it reaches a maximum (at 33 pps), the activation of the harmonic representation is increased such that the difference between it and a harmonically distinctive stimulus affects response time. The principal finding of Experiment 1, a processing advantage for ITP stimuli at 33 pps, is not explicable in terms of surface stimulus properties. It suggests that entraining a set of harmonically related loci at 33 pps increases the salience of an unrelated frequency. These results are consistent with an oscillatory mechanism possessing a preferred period of 30 ms, which could be involved in grouping disparate spectral components into a unified harmonic percept.

Experiments 2, 3 and 4: Temporal specificity

In order to examine the observed effect further, we varied the inter-stimulus interval between the entrainer and target. Since both the entrainer and the target could evoke eGBRs (and SSRs), varying ISI would establish whether the effect observed in Experiment 1 was related to the presence of the entrainer. More importantly, it would allow a very precise investigation of the temporal parameters of the HR model.

Our central hypothesis implies that the presentation of an entraining stimulus generates an oscillatory representation whose structure mirrors the structure of the input. Temporal mismatch between input and the entrained representation signals that the two are different. This type of dynamic interaction between incoming sensory information and established mental models is important for understanding perceptual processing in general, since there is evidence that similar mechanisms operate in visual perception. Visual periodic stimuli with flicker frequencies in the gamma-range give rise to condition-specific cyclic RT patterns (Kompass & Elliott, 2001) - an effect that could reflect a bottom-up/top-down interaction at the level of synchronized neural activity. More specifically, the bottom-up processing stage involves the generation of a temporal representation or model (Kristofferson, 1977), which is tested against the incoming stimulation by means of Bayesian principles. According to this interpretation, a resonance develops between lower and higher processing stages through recurrent loops, with rate-specific grouping effects reflecting inter-stage oscillatory interactions. Once formed, the temporal model guides the encoding of the subsequent input.

Figure 6 here

Figure 6 illustrates the time-course of the evoked oscillatory representation and its temporal parameters. The strength of the representation decays linearly and no phase shift or resynchronization delay is assumed. Individual oscillations represent physiologically plausible models of real auditory eGBRs and are based on the recordings by Pantev and his colleagues (Pantev et al., 1993; Ross, Herdman & Pantev, 2005b). Each oscillation is scaled so that its temporal extent remains identical to the original recording, while its vertical extent has been increased to assist viewing. Each oscillation consists of five 30-ms periods and lasts approximately 150 ms. Each oscillation commences with a physiologically plausible 20-ms latency (Pantev et al., 1991) and peaks around 70 ms post-onset. The minimum persistence of the residual representation is 150 ms (the duration of an individual eGBR). The first coherence peak occurs at around 110 ms ISI, at the onset of the deviant-evoked oscillation. It is hypothesized that the coherence exceeds a threshold at a point at which the activation of a template increases the salience of a spectrally distinct template. One critical test of the model would be the presence of significant inharmonic RT advantage at certain specific ISI values.

Experiment 2 focused on the ISIs from 0 – 106 ms, presented in 15-ms increments (half-period of 33 Hz). A fine-grained temporal analysis was necessary for two reasons. First, if the model possessed the form outlined in Figure 6, a significant RT advantage for ITP stimuli should be observed at around 110 ms ISI. In Experiment 1 such advantage was observed at 100 ms, suggesting that the supra-threshold coherence window would need to be at least 10 ms wide. One of the assumptions behind the model is a monotonic increase in coherence caused by the superposition of successive eGBR amplitudes. This in turn constrains the time at which the threshold is exceeded.

Figure 7 here

Figure 7 illustrates the effect of rate on the point at which the coherence exceeds the threshold. Assuming linear amplitude superposition, this function can be expressed in terms of angle, with the maximum slope (0.5) at 33 pps and zero slope at 22 pps - the rate at which successive 33-Hz oscillations cancel each other out (panel A). Simple calculation shows that the point at which the threshold should be positioned in order to discriminate between those two presentation rates is around 10 ms behind the assumed 110-ms maximum point, that is, at 100 ms ISI. Consequently, no significant ITP advantage should be observed at ISIs shorter than 100 ms. In terms of Experiment 2, a significant advantage should be expected only at 106 ms ISI.

Experiment 3 was carried out to ensure that any enhancement potentially found at 106 ms ISI could not be ascribed to an end-of-range effect. In order to confirm that the findings of Experiment 2 were not due to random fluctuation or participants' sensitivity to range, Experiment 3 replicated the conditions of Experiment 2, excluding the two longest ISI values (91 and 106 ms).

Figure 8 here

If the inharmonic advantage was due to HR, it should not be observed in Experiment 3. Finally, Experiment 4 examined the width and symmetry of supra-threshold coherence window by focusing on ISIs from 106 to 182 ms. The HR model predicts a desynchronization (reset) of the representation after around 110 ms. As illustrated in Figure 8, the idealized reset slope can have any value within the grey area. The maximum slope (-0.5), which crosses the threshold at 120 ms ISI is “symmetric” in that it mirrors the positive slope, which crosses the threshold at 100 ms. The HR model predicts different absolute values for the two slopes. Specifically, the negative slope should be steeper (i.e. should cross the threshold inside the grey area) because less disruption is expected when a 1000-Hz deviant is replaced by a 500-Hz pip. If the assumption of asymmetry is correct, a significant ITP advantage should be observed at 106 but not at 121 ms ISI.

Method

Experiment 2. Sixteen volunteers with normal hearing (10 female, average age 23.5 years) participated in Experiment 2. They were naïve to the aims of the experiment. Stimuli were identical to those employed in Experiment 1. Entrainer stimuli carried 500-Hz baseline tone-pips, which were regularly interspersed with “deviant” 1000-Hz pips. TA stimuli carried 1000-Hz pips only, whereas TP stimuli appeared in two forms. As in Experiments 1 and 2, HTP trains carried 1000-Hz pips, with every fourth pip carrying 2000 Hz and ITP trains carried 1000 and 2400-Hz pips arranged in the same way. The ISI between the entrainer and target was varied from 0 to 106 ms in 15-ms increments (0, 15, 30, 45, 61, 76, 91 and 106 ms). Each TP condition was presented 40 times, resulting in 1280 trials per participant. Conditions were fully randomized across trials, which were presented in 32 40-trial blocks. Each participant took part in two 640-trial sessions on separate days.

Experiment 3. Fourteen right-handed volunteers (10 female, mean age 23.1 years) with normal hearing took part. Two participants were left-handed. Factor ISI had six levels (0, 15, 30, 45, 61 and 76 ms). Together with two levels of Target (TA, TP) and Harmony

(HTP, ITP), there were 18 conditions, giving 960 trials per participant. There were two experimental sessions that took place on different days. Each session lasted approximately 30 minutes.

Experiment 4. Thirteen volunteers with normal hearing (eight female, mean age 27.3 years) took part in Experiment 4. One participant was left-handed. There were 18 conditions giving 960 trials per participant. The levels of ISI were 106, 121, 136, 151, 167 and 182 ms. Experiment consisted of two 480-trial sessions carried out on different days. Each session lasted approximately 30 minutes.

Results

Experiment 2. The ANOVA for factors Target and ISI yielded a significant main effect of Target [F (1, 15) = 13.82, p = .002; MSE = 3064.82], which was in agreement with the previous experiments. However, a highly significant main effect of ISI as well as a significant Target by ISI interaction [F (7, 105) = 10.92, p < .001; MSE = 144.65 and F (7, 105) = 14.63, p < .001; MSE = 125.40, respectively] confirmed the observation that the TA – TP difference was due to a downward trend in TA latencies. At the ISI of 0 ms, TA responses were nearly 50 ms slower relative to TP responses. The monotonic decrease (~ 40 ms/100 ms ISI) ended at 91 ms ISI, so that no TA – TP difference was present at the last two ISI values. A second ANOVA was performed on TP stimuli only and the factors of interest were ISI and Harmony. As in Experiment 1, a significant main effect of Harmony [F (1, 15) = 12.90, p = .003; MSE = 652.32] was due to somewhat slower HTP responses. A marginally significant ISI by Harmony interaction [F (7, 105) = 1.96, p = .068; MSE = 165.03] reflected a clear divergence of HTP and ITP responses at 106 ms ISI. Post-hoc Bonferroni-adjusted t tests confirmed that HTP-ITP difference was significant only at 106 ms ISI [t (15) = 5.55, p < .001].

Figure 9 here

The principal effect observed in Experiment 2 was a significant difference between HTP and ITP responses at 106 ms ISI (see Figure 9). The effect (28.5 ms) was larger than the effect observed in Experiment 1 (17.7 ms). Interestingly, there was a divergence between HTP and ITP functions at 106 ms ISI, demonstrating that the improvement in ITP performance was accompanied by an increase in HTP latencies. This finding provided further support to the findings of Experiment 1, namely, that the effect did not consist solely in facilitating the processing of an inharmonic target, but also in reducing the salience of a stimulus, which was harmonically related to the entrainer. The divergence between the curves could not be observed in Experiment 1, because the manipulation of presentation rate exposed the dependence of HTP responses on temporal regularity of the oscillatory representation. With presentation rate fixed, the divergence between the two response curves emerged. The pattern could be ascribed to the *strength* of the representation as predicted by HR model and was consistent with an increase in the signal-to-noise (S/N) ratio between harmonic and inharmonic targets.

Errors accounted for 2.8% of all responses. TA and TP responses produced, on average, equal number of errors, since an ANOVA for factors Target and ISI revealed no significant effects. In agreement with Experiment 1, an ANOVA for factors Harmony and ISI yielded a highly significant main effect of Harmony [F (1, 15) = 16.81, p = .001; MSE = .301; HTP > ITP].

Experiment 3. The general trend previously observed in Experiment 2 was replicated in Experiment 3. The downward slope for TA RTs mirrored the pattern in the previous experiment, which was confirmed by an ANOVA with factors Target and ISI. The main effect of Target was significant [F (1, 13) = 8.83, p = .011; MSE = 3050.83], while the main effect of ISI was highly significant [F (5, 65) = 4.59, p = .001; MSE = 190.55]. The difference in trend for TA and TP responses was reflected in a highly significant ISI by Target interaction [F (5, 65) = 7.01, p < .001; MSE = 119.64]. A second ANOVA for factors ISI and Harmony revealed only a significant main effect of Harmony [F (1, 13) = 5.41; p = .037; MSE = 1590.98], with HTP responses being somewhat slower than ITP responses. A series of Bonferroni-adjusted t tests detected no significant differences. Errors accounted for 3.8% of all responses. No significant effects were observed.

Experiment 4. The main effects of Target and ISI were nonsignificant (p > .1), though the interaction between ISI and Target was significant [F (5, 60) = 2.98, p = .018; MSE = 133.99]. This was due to opposing trends in TA and TP functions. TA responses became slightly faster with the increase in the interval, while TP responses became slower. The downward trend in TA responses was less pronounced compared to Experiments 2 and 3 and was confined to the ISIs of 121 and 136 ms. This trend was compatible with the previously offered masking explanation. Specifically, it is possible that the masking by the terminal pip extended to about 136 ms ISI. At the same time, the slowing down of TP responses suggested that the priming by the entrainer had a limited temporal range.

Figure 10 here

As in previous experiments, the main effect of Harmony was significant [F (1, 12) = 7.08, p = .021; MSE = 1601.90]. The interaction between Harmony and ISI was marginally nonsignificant [F (5, 60) = 2.00, p = .091], while the main effect of ISI failed to reach significance (p > .1). A series of Bonferroni-adjusted t tests revealed a significant inharmonic advantage at 106 ms ISI [t (12) = 3.28, p = .007]. No significant effects were observed in other ISI conditions (see Figure 10).

Overall, 4.4 % of all responses were errors. A significant main effect of ISI [F (5, 60) = 2.75, p = .027; MSE = .08] was due to improved accuracy in both TA and TP conditions between 121 and 136 ms ISI. No further effects were revealed.

Discussion

Experiments 2, 3 and 4 examined the time course of the hypothetical oscillatory harmonic representation by varying the inter-stimulus intervals in 15-ms increments from 0 to 182 ms. The most important finding was the significant HTP-ITP difference at 106 ms ISI. As indicated by the absence of a similar difference at shorter ISIs (Experiment 3), the effect could not be ascribed to participants' sensitivity to ISI range. The result indicated that the effect observed in Experiment 1 depended critically on the presence of the entrainer. Furthermore, the effect was compatible with the notion that harmonic templates are synchronized by means of oscillatory activity whose frequency lies around 33 Hz. According to our hypothesis, oscillations evoked by 500-Hz entrainer pips create an oscillatory representation of the 500-Hz template. The result of Experiment 2 suggested that the coherence of this representation exceeds a behavioral threshold at around 110 ms ISI, that is, 20 ms after the projected onset of the deviant. Since 106 ms represents the highest ISI value in Experiment 2, establishing the temporal extent of the coherence window required further investigation. While Experiment 3 excluded the possibility that the effect was an artifact of presentation conditions, Experiment 4 confirmed the presence of an asymmetric coherence window. Statistically significant advantage for ITP responses was again observed at 106 ms ISI and the absence of a significant difference at 121 ms ISI suggests that the coherence window was asymmetric and that it lasted less than 20 ms. Of particular interest was the absence of a significant ITP advantage in the region between 121 and 182 ms ISI. This finding supports the HR hypothesis according to which the oscillatory representation of the 500-Hz template is desynchronized by the related 1000-Hz representation.

Figure 11 here

In order to interpret the observed effects, a temporally specified model of a harmonic oscillatory representation is proposed based on known properties of the auditory gamma-band response (see Figure 11). The cascade of oscillations represents the unfolding of the representation over 200 ms ISI, while the attenuation of successive oscillations over time reflects its gradual desynchronization. The oscillatory response evoked by the deviant pip is represented by a thick curve and placed at a phase angle of 180 degrees with respect to the preceding 500-Hz oscillations. The phase shift reflects the temporary disruption of the 500-Hz cycle by the deviant. Inhibiting superfluous odd harmonics creates a delay in the formation of a 1000-Hz representation, which, given the temporal constraints, manifests itself as a 180-degree phase shift. The re-establishment of the 500-Hz representation requires no inhibition and that is why no phase shift occurs when a 500-Hz pip replaces the deviant. The resulting complex oscillatory trace is shown in the upper part of the figure. Since the longest ISI presented in the above experiments was 182 ms, it was not possible to ascertain whether the behavioral effects of the oscillatory trace re-emerged at longer intervals. This was the aim of the next two experiments.

Oscillatory resonance: Experiments 5 and 6

Experiments conducted thus far had revealed that the effect of entrainment of a harmonic representation possessed well-defined temporal properties in that it occurred at a specific frequency and at a specific time. The residual oscillatory representation reached a supra-threshold level at around 110 ms ISI and no significant effects were observed at longer intervals. This could have been due to the natural decay of the gamma-band oscillations. It was noted earlier that the summation of the successive 500-Hz amplitudes was responsible for bringing the residual representation above the threshold. Exogenous driving could have led to a similar artificial increase in the persistence of the trace. A second period of supra-threshold coherence could emerge at longer ISIs. The HR model predicts a second coherence window at around 220 ms ISI (2 x 110 ms). However**,** a linearly decaying trace could not produce another such window without some form of resynchronization. As shown earlier, the duration of the supra-threshold window was estimated at 15 to 20 ms. Such temporal specificity means that the threshold is high relative to the coherence baseline. The height of the window is approximately 1/10 of the height of the entire curve. Assuming linear decay that reaches zero point at 400 ms ISI, the coherence at 220 ms ISI should be approximately 50 % lower than at 110 ms – well below the threshold. The presence of a second window would imply re-synchronization of the oscillatory trace via some form of oscillatory resonance. Experiment 5 investigated this hypothesis while the aim of Experiment 6 was to confirm the findings of Experiment 5 using a different sampling rate.

Method

Experiment 5. 12 right-handed volunteers (seven female, mean age 22.4 years) took part in Experiment 5. Six levels of ISI (197, 212, 227, 243, 258 and 273 ms) were combined with two levels of Target (TA and TP, the latter containing two levels of Harmony) to give 960 trials per participant. In other respects, Experiment 5 was identical to Experiment 4.

Experiment 6. 12 paid volunteers (6 female, mean age 24.6 years) participated in Experiment 6. The design and procedural details were identical to those in Experiment 5, with the exception of ISI values, which were 0, 50, 100, 150, 200 and 250 ms.

Figure 12 here

Results

Experiment 5. An ANOVA for factors Target and ISI was performed on mean RT scores. Following the already observed pattern, there was a significant main effect of Target [F (1, 11) = 9.95, p = .009; MSE = 940.97], with TA responses being slower. There was also a significant main effect of ISI [F (2.65, 29.19) = 4.22, p = .02; MSE = 288.54] caused by a general decrease in RTs at 258 and 273 ms but no interaction (p > .1). A second ANOVA with factors ISI and Harmony revealed that both HTP and ITP responses were faster at these points (258 and 273 ms; see Figure 12), as confirmed by a significant main effect of ISI [F (5, 55) = 3.10, p = .002]. As in previous experiments, there was a significant main effect of Harmony [F (1, 12) = 27.79, p < .001]. The two factors did not interact (p > .1). A series of Bonferroni-adjusted t tests revealed a significant inharmonic advantage at 258 ms ISI [t (11) = 3.515, p = .005]. No significant effects were observed at other intervals.

Errors accounted for 2.9% of all responses. There were no significant differences in accuracy with respect to Target. An ANOVA for factors Harmony and ISI revealed a significant main effect of ISI [F (5, 55) = 2.62, p = .034; MSE = .14]. Finally, there was a highly significant main effect of Harmony [F (1, 11) = 26.67, p < .001; MSE = .05], which was caused by the already observed decrease in accuracy for HTP responses.

Figure 13 here

Experiment 6. An ANOVA for factors Target and ISI yielded a significant main effect of Target [F (1, 11) = 12.20, p = .005; MSE = 1491.08] as well as a highly significant main effect of ISI [F (5, 55) = 37.63, p < .001; MSE = 113.96]. Both these effects, as well as a highly significant interaction between ISI and Target [F (5, 55) = 19.78, p < .001; MSE = 134.80] reflected the familiar pattern observed in Experiments 2, 3 and 4 (see Figure 13). As revealed by an ANOVA for factors ISI and Harmony, the main effect of Harmony was highly significant [F (1, 11) = 19.67, p = .001; MSE = 568.63], with ITP response times being faster, as in all previous experiments. Six Bonferroni-adjusted t tests were performed on HTP-ITP differences for each ISI condition. The only significant points were 100 ms ISI [t (11) = 4.77, p = .001] and 250 ms ISI [t (11) = 4.19, p = .001].

3.2 % of all responses were incorrect. The only significant effect was that of Target [F (1, 11) = 5.55, p = .038; MSE = .13]. Visual inspection of the data indicated that TP responses were more accurate. In addition, a significant interaction between ISI and Harmony [F (1, 11) = 2.32, p = .022; MSE = .16] was caused by HTP maxima at 0 and 100 ms ISI. The presence of a pronounced peak in the ITP error function at 100 ms ISI was due to high error scores in the ITP/100 ms-ISI condition for five participants. Although no clear explanation could be offered for this effect, it is possible that the broader ISI range and the concomitant increase in uncertainty could have affected participants' estimate of target onset. The peaks in the ITP error function occurred at 0 and 100 ms, while no such peak was found at 250 ms ISI, implying no clear relationship between latencies and accuracy.

Discussion

The main finding of Experiment 5 was the presence of a second coherence window at 250 ms ISI. The presence of two windows was confirmed by Experiment 6, providing strong support for the unique-template hypothesis (see Introduction). Each 500-Hz entrainer pip evoked a single 33-Hz oscillation, which synchronized the components of a harmonic template comprising resolved harmonics of 500 Hz. Each successive oscillation strengthened the oscillatory representation, so that at some point around 100 ms from the onset of the first 500-Hz pip the activity exceeded a behavioral threshold. Deviant pips carrying 1000 Hz activate a template, which comprises some of the already active harmonic sites. In order for the new representation to be established, some of the components of the old template had to be inhibited. This was achieved through a phase reset of the 500-Hz representation. The exact duration of the inhibitory period could not be estimated based on the sampling rate employed in Experiments 2, 3 and 4, but a broad estimate would be of the order of 5 to 10 ms. The principal issue addressed in Experiment 5 was whether this form of temporal dependence could be observed at longer ISIs. Extrapolation from the proposed model suggested that a second coherence window should occur at around 220 ms ISI.

Figure 14 here

Figure 14 illustrates the shape of the oscillatory trace as inferred from the results of Experiments 5 and 6. The persistence of the oscillatory trace of at least 250 ms exceeded the expected duration of a naturally decaying eGBR. Importantly, the presence of a second coherence window suggested a resonant mechanism within the theta range with a period of approximately 150 ms (6.7 Hz).

If the above inference is correct, the second oscillatory cycle could originate in the higher processing centers (e.g. secondary auditory cortex – A2 in Figure 14) from where it would be mediated by a theta rhythm, possibly involved in the encoding of the representation into working memory (Sederberg, Kahana, Howard, Donner & Madsen, 2003; Klimesch, Schack & Sauseng, 2005). The effect would emerge at 250 ms ISI only when the relevant primary circuits were resynchronized. The 30-ms shift (corresponding to the driving frequency of 33 Hz) between the predicted and observed locations of the second window could reflect the re-synchronization delay. These findings are supported by the results of Experiment 6 which replicated the reported effects using a different ISI sampling rate.

General Discussion

The most important effect observed in the current experiments was the rate- and time-specific RT advantage for inharmonic targets. Experiment 1 revealed RT advantage for inharmonic targets when the stimuli (tone-pip trains) were presented at 33 pps. The effect was ascribed to an oscillatory representation consisting of a cascade of evoked 33-Hz oscillations. In Experiments 2, 3 and 4, the inter-stimulus interval between the entrainer and target was varied in order to test the predictions of the HR model. The experiments confirmed that the effect was caused by the specific temporal profile of the oscillatory representation evoked by the entraining stimulus. ISIs were sampled in 15-ms increments from 0 to 182 ms. The only significant RT difference was observed in a narrow time window around 100 ms ISI. This information was used to infer the duration and form of the supra-threshold activation. Finally, Experiments 5 and 6 confirmed the presence of a second window of activation at approximately 250 ms post-stimulus, suggesting resynchronization via a resonant feedback loop, possibly between distinct processing areas.

The 500-Hz entrainer pips were assumed to activate a harmonic template consisting of a number of harmonics (at least one) of the fundamental frequency of 500 Hz. Since each pip is assumed to evoke a single eGBR, the maximum synchronization is achieved at approximately 90 ms post-stimulus. To this should be added a constant latency shift of approximately 20 ms (Pantev et al., 1991), placing the point of maximum enhancement at around 110 ms ISI. In the absence of a deviant pip, the coherence would keep increasing up to about 300 ms post entrainer onset (see Gutschalk et al., 2009). This situation is not conducive to behavioral investigation simply because no variation can be observed in the time course of the oscillatory representation. The deviant disrupts the representation of the 500-Hz template presumably because a separate yet related representation has to be disambiguated from the active one. Besides disrupting the preceding cycle, the deviant-evoked oscillation would partly strengthen the succeeding one.

One question concerns the origin of the HR mechanism. Was the reset due to a top-down desynchronization of the active template or an inherent temporal mismatch between the two representations? The structural difference between the two templates could have been encoded in the higher processing centers (possibly in the form of a meta-template), which, in turn, could have modulated the gamma-band activity in the primary cortex via efferent pathways. However, given the rapidity of the reset, it is possible that primary cortical circuits retain some ability to disambiguate competing templates. Specifically, harmonic distance between competing templates could be encoded in gamma-band oscillatory phase.

Figure 15 here

One way in which a deviant-evoked oscillation could disrupt the preceding cycle would be via a constant 15-ms delay (half-period of 33 Hz), which would result in a 180-degree phase mismatch. Thus, deviant-evoked oscillations are identical to those evoked by baseline pips, except for additional 15-ms latency (Panel A of Figure 15). The position of successive coherence windows is shifted by some 30 ms relative to the HR model which posits that each template possesses a unique phase profile in accordance with the “unique template” hypothesis described earlier (Panel B).

Figure 16

Our results are compatible with the notion of harmonic templates possessing inherent oscillatory profiles. Figure 16 illustrates possible relationship between two harmonic templates and their oscillatory cycles. Two identical templates are not distinguishable in terms of gamma phase. A 500-Hz and a 750-Hz template are separated by a 90-degree phase shift. They do not require a complete phase cancellation because they share relatively few harmonics. Finally, two related yet non-identical templates (500 and 1000 Hz) are distinguished by a 180-degree phase shift.

Complete phase mismatch is needed to disambiguate two templates sharing a large number of harmonics. If disambiguation is associated with phase differences on millisecond scale, oscillatory frequency of 33 Hz provides sufficient temporal resolution to allow the encoding of about 15 distinct templates – a value close to 12 semitones of an octave. This form of encoding is harmony-specific because phase difference does not encode frequency distance between templates (in which case maximum phase mismatch might occur for templates close in frequency). Rather, gamma-band phase disambiguates harmonically related but distinct structures.

Apart from its importance for the current research, the HR mechanism might be relevant for perceptual processing in general. Some form of phase reset is necessary for object disambiguation because binding-related synchronization of feature-coding loci is short-lived and must not impede the processing of novel stimuli. As demonstrated by Joliot, et al. (1994), perceptual segregation of successive sounds is intimately linked with the gamma-band oscillatory cycle. A novel stimulus evokes its own gamma-band response. If it is related to the previous stimulus (e.g. harmonically or spatially), it must disrupt or abolish the previous cycle in order to assert its presence (e.g. Rohrbaugh, Varner, Paige, Eckardt and Ellingson, 1990; Snyder & Large, 2002).

Given the complexity of the oscillatory activity in the brain, considerations of neuronal synchronization should be informed by general models of temporal processing. Such models suggest that the fine temporal structure, arguably detectable in psychophysical measurements (e.g. Elliott & Müller, 2000; Kompass & Elliott, 2001), could represent the "missing link" between the psychophysical and neural levels of explanation. This view receives support from models such as the "shifting correlation" model of auditory streaming by Wang (1996), based on Terman and Wang's (1995) global-competition-local-cooperation network of oscillators.

One way of describing the organization of temporal periodicities in perception is exemplified by Geissler's Time Quantum Model (TQM; Geissler, 1987). TQM posits a hierarchical temporal structure, based on a psycho-neural time quantum, which provides a dynamical framework for perceptual organization. Specifically, the dynamics of binding can be described in terms of shifting patterns of synchronization, constrained by the value of the time quantum (Geissler & Kompass, 2001). According to this scheme a 33-Hz rhythm would represent the seventh subharmonic of a fundamental quantal period of approximately 4.5 ms. From this generalized perspective, the apparent preference for low gamma-band rhythms might be viewed as arising from certain neuroanatomical and neurochemical constraints. In other words, gamma-band oscillations represent an accessible manifestation of the function of a complex dynamical system. For example, Pöppel (1997) considers 30 ms a "common temporal unit" of perceptual and cognitive processing. Furthermore, Burle and Bonnet (2000) have suggested that temporal oscillations in the gamma range could underlie the operation of a general time-keeping device regulating temporal relationships in perceptual and cognitive processing.

The results of the current experiments link 33-Hz activity with the integration of harmonic representations. At the same time, this association cannot be exclusive. The role of 33-Hz activity could be described as facilitatory in the sense of improving the S/N ratio between competing representations resulting in the within-object congruence and between-object disambiguation. Gamma-band oscillatory activity could play just such a role by providing a temporal structure for interactions between cells that carry specific information (Singer, 1993, 1994, 1999). The observed effects suggest a cyclical short-term harmonic representation, compatible with a sensory memory trace (Cowan, 1984). The behavioral effects of this representation are not facilitatory in the way that commonly reported memory effects are. Rather, the facilitation of inharmonic responses is a consequence of increased coherence in a co-active harmonic representation.

The proposed mechanism is also consistent with a sensory model in the sense of Kompass and Elliott (2001), against which the incoming stimulus is evaluated (see also Herrmann, Munk & Engel, 2004). This model involves a resonant loop between the primary and secondary processing areas that maintains the oscillatory representation for some 300 to 400 ms post-stimulus. The fact that the effect of entrainment is opposite to that observed in a similar task in the visual modality (i.e. processing advantage for the primed structure; Elliott & Müller, 1998, 2000) is explained by the special anatomical and functional status of harmony, namely, the early emergence of harmonic structures in the brain and the prevalence of harmonic complexes in auditory perception (Schellenberg & Trehub, 1996a, 1996b). While strengthening the external validity of the current research, this observation suggests that investigating the status of the observed oscillatory persistence as a memory phenomenon might require a modification of the current task that would allow the facilitatory aspect of harmonic entrainment to emerge in behavioral responses.

References

Başar-Eroglu, C., Strüber, D., Schürmann, M., Stadler, M., & Başar, E. (1996). Gamma-band responses in the brain: A short review of psychophysiological correlates and functional significance. *International Journal of Psychophysiology, 24,* 101-112.

Bohorquez, J., & Ozdamar, O. (2008). Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution. *Clinical Neurophysiology, 119,* 2598-2607.

Brosch, M., & Schreiner, C. E. (2000). Sequence sensitivity in cat primary auditory cortex. *Cerebral Cortex, 10,* 1155-1167.

Burle, B., & Bonnet, M. (2000). High-speed memory scanning: a behavioral argument for a serial oscillatory model. *Cognitive Brain Research, 9,* 327-337.

Burns, E. M., & Ward, W. D. (1978). Categorical perception - Phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America, 63,* 456-468.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96,* 341-390.

Dorfman, D. D., & Alf, E. (1969). Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals — rating method data. *Journal of Mathematical Psychology, 6,* 487-496.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H. J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics, 60,* 121-130.

Eckhorn, R., Reitboeck, H. J., Dicke, P. A., Arndt, M., & Kruse, W. (1990). Feature-linking across cortical maps via synchronization. In R. Eckmiller (Ed.), *Proceedings of the International Conference on Parallel Processing in Neural Systems and Computers* (pp. 1-4). Amsterdam: Elsevier.

Elliott, M. A., & Müller, H. J. (1998). Synchronous information presented in 40-Hz flicker enhances visual feature binding. *Psychological Science, 9,* 277-283.

Elliott, M. A., & Müller, H. J. (2000). Evidence for 40-Hz oscillatory short-term visual memory revealed by human reaction-time measurements. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26,* 1-16.

Elliott, M. A., Giersch, A., & Seifert, D. (2006). Some facilitatory effects of Lorazepam on visual binding. *Psychopharmacology, 184,* 229-238.

Galambos, R., Makeig, S., & Talmachoff, P. J. (1981). A 40-Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences, USA*, 78, 2643-2647.

Galambos, R. (1992). A comparison of certain gamma-band (40-Hz) brain rhythms in cat and man. In E. Başar & T. M. Bullock (Eds.), *Induced Rhythms in the Brain* (pp. 201-216). Boston, MA: Birkhäuser.

Galambos, R., & Makeig, S. (1992). Physiological studies of central masking in man. II: Tone-pip SSRs and the masking level difference. *Journal of the Acoustical Society of America, 92,* 2691-2697.

Geissler, H-G. (1987). The temporal architecture of central information processing: Evidence for a tentative time-quantum model. *Psychological Research*, *49,* 99-106.

Geissler, H-G., & Kompass, R. (2001). Temporal constraints on binding? Evidence from quantal state transitions in perception. *Visual Cognition*, *8, (3/4/5),* 679-696.

Gray, C. M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences, USA, 86,* 1698-1702.

Gutschalk, A., Oldermann, K., & Rupp, A. (2009). Rate perception and the auditory 40-Hz steady-state fields evoked by two-tone sequences. *Hearing Research, 257,* 83-92.

Helmholtz, H. L. F. von (1954). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. (A. J. Ellis, Trans.). New York, NY: Dover Publishing. (Original work published 1863).

Herrmann, C. S., Munk, M. H. J., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Science, 8,* 347-355.

Howell, D. C. (1992). *Statistical Methods for Psychology*. Belmont, CA: Duxbury Press.

Jeffreys, J. G. R., Traub, R. D., & Whittington, M. A. (1996). Neuronal networks for induced '40 Hz' rhythms. *Trends in Neuroscience, 19,* 202-208.

Joliot, M., Ribary, U., & Llinás, R. (1994). Human oscillatory brain activity near 40 Hz coexists with cognitive temporal binding. *Proceedings of the National Academy of Sciences, USA, 91,* 11748-11751.

Kadia, S. C., & Wang, X. (2003). Spectral integration in the A1 of awake primates: Neurons with single- and multipeaked tuning characteristics. *Journal of Neurophysiology, 89,* 1603-1622.

Kaiser, J., & Lutzenberger, W. (2005). Human gamma-band activity: A window to cognitive processing. *Neuroreport, 16,* 207-211.

Kahana, M. (2006). The cognitive correlates of human brain oscillations. *Journal of Neuroscience, 26,* 1669-1672.

Klimesch, W., Schack, B., & Sauseng, P. (2005). The functional significance of theta and upper alpha oscillations for working memory: a review. *Experimental Psychology, 52,* 99-108.

Knief, A., Schulte M., Bertrand, O., & Pantev, C. (2000). The perception of coherent and non-coherent auditory objects: A signature in gamma frequency band. *Hearing Research, 145,* 161-168.

Kompass, R., & Elliott, M. A. (2001). 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* (pp. 130-135). Berlin: Pabst Verlag.

Kristofferson, A. B. (1977). A real-time criterion theory of duration discrimination. *Perception & Psychophysics, 21,* 105-117.

Lee, J. (1994). Amplitude modulation rate discrimination with sinusoidal carriers. *Journal of the Acoustical Society of America, 96,* 2140-2147.

Makeig, S., & Galambos, R. (1989, October). *The 40-Hz band evoked response lasts 150 ms and increases in size at slow rates*. Poster presented at the annual meeting of the Society for Neuroscience, Phoenix, AZ.

Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography and Clinical Neurophysiology, 86,* 283-293.

Makela, J. P., & Hari, R. (1987). Evidence for cortical origin of the 40 Hz auditory evoked response in man. *Electroencephalography and Clinical Neurophysiology, 66,* 539-546.

Malsburg, C. Von der (1981). *The Correlation Theory of Brain Function*. (Int. Rep. No. 81-2). Gõttingen, Germany: Department of Neurobiology, Max Planck Institute for Biophysical Chemistry.

Malsburg, C. Von der, & Schneider, W. (1986). A neural cocktail-party processor. *Biological Cybernetics, 54,* 29-40.

Palva, S., Palva, J. M., Shtyrov, Y., Kujala, T., Ilmoniemi, R. J., Kaila, K., & Näätänen, R. (2002). Distinct gamma-band evoked responses to speech and non-speech sounds in humans. *Journal of Neuroscience, 22,* 1-5.

Pantev, C., Makeig, S., Hoke, M., Galambos, R., Hampson, S., & Gallen, S. (1991). Human auditory evoked gamma-band magnetic fields. *Proceedings of the National Academy of Sciences, USA, 88,* 8996-9000.

Pantev, C., Elbert, T., Makeig, S., Hampson, S., Eulitz, C., & Hoke, M. (1993). Relationship of transient and steady-state auditory evoked fields. *Electroencephalography and Clinical Neurophysiology, 88,* 389-396.

Pantev, C., & Elbert, T. (1994). The transient auditory evoked gamma-band field. In C. Pantev, T. Elbert, & B. Lütkenhöner (Eds*.), Oscillatory Event-Related Brain Dynamics* (pp. 219-229). New York: Plenum Press.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased cortical representations in musicians. *Nature, 392,* 811-814.

Picton, T. W., Skinner, C. R., Campagne, S. C., Kellet, A. J. C., & Maiste, A. C. (1987). Potentials evoked by the sinusoidal modulation of the amplitude and frequency of a tone. *Journal of the Acoustical Society of America, 82,* 175-178.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences, 1,* 56-61.

Pulvermüller, F., Preissl, H., Eulitz, C., Pantev, C., Lutzenberger, W., Elbert, T., & Birnbaumer, N. (1994). Brain rhythms, cell assemblies and cognition: Evidence from the processing of words and pseudowords. *Psycholoquy, 5,* 1-30.

Ritsma, R. J. (1962). Existence region of the tonal residue. I. *Journal of the Acoustical Society of America, 34,* 1224-1229.

Rohrbaugh, J. W., Varner, J. L., Stephen, R., Paige, S. R., Eckardt, M. J., & Ellingson, R. J. (1990). Auditory and visual event-related perturbations in the 40 Hz auditory steady-state responses. *Electroencephalography and Clinical Neurophysiology, 76,* 148-164.

Ross, B., Herdman,A. T., & Pantev, C. (2005a). Right hemispheric laterality of human auditory 40 Hz auditory steady state responses. *Cerebral Cortex, 15,* 2029-2059.

Ross, B., Herdman, A. T., & Pantev, C. (2005b). Stimulus-induced desynchronization of human auditory 40-Hz steady-state responses. *Journal of Neurophysiology, 94,* 4082-4093.

Schellenberg, E. G, & Trehub, S. E. (1996a). Children's discrimination of melodic intervals. *Developmental Psychology, 32,* 1039-1050.

Schellenberg, E. G., & Trehub, S. E. (1996b). Natural musical intervals: Evidence from infant listeners. *Psychological Science, 7,* 272-277.

Schulte, M., Knief, A., Seither-Preisler, A., & Pantev C. (2000). Gestalt recognition in a virtual melody experiment. In: T. Katila, R. Hari, R. Ilmoniemi, & M. Mäkijärvi, (Eds.), *Proceedings of the 12th International Conference on Biomagnetism* (pp. 107-110). Espoo, Finland: Helsinki University of Technology.

Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J., & Madsen, J. R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. *Journal of Neuroscience, 23,* 10809-10814.

Senkowski, D., Talsma, D., Grigutsch, M., & Herrmann, & Woldorff, M. G. (2007). Good times for multisensory integration: Effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia, 45,* 561-571.

Shamma, S., & Klein, D. (2000). The case of the missing pitch templates: How harmonic templates emerge in the early auditory system. *Journal of the Acoustical Society of America, 107,* 2631-2644.

Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annual Review of Psychology, 55,* 349-374.

Singer, W. (1994). Time as coding space in neocortical processing: A hypothesis. In E. Buzsáki, R. Llinás, W. Singer, A. Berthoz, & A. Christen (Eds.), *Temporal coding in the Brain* (pp. 51-79). Berlin: Springer-Verlag.

Singer, W. (1999). Time as coding space? Current Opinion in Neurology, 9, 189 – 194.

Snyder, J. S., & Large, E. W. (2002). Neurophysiological correlates of meter perception: Evoked and induced gamma-band (20-60 Hz) activity. *Proceedings of the Seventh International Conference on Music Perception and Cognition*. Adelaide, Australia: Causal Productions.

Stapells, D. R., Makeig, S., & Galambos, R. (1987). Auditory steady-state responses: threshold prediction using phase coherence. *Electroencephalography and Clinical Neurophysiology, 67,* 260-270.

Swets, J. A. (1996). *Signal Detection Theory and ROC analysis in Psychology and Diagnostics: Collected Papers*. Mahwah, NJ: Lawrence Erlbaum Associates.

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Science, 3,* 151-162.

Terhardt, E. (1989). Warum hören wir Sinustöne? (Why do we hear pure tones?) *Naturwissenschaften*, *76,* 496-504.

Terman, D., & Wang, D. L. (1995). Global competition and local cooperation in a network of neural oscillators. *Physica D, 81,* 148-176.

Tiitinen, H., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., & Näätänen, R. (1993). Selective attention enhances the auditory 40-Hz transient response in humans. *Nature, 364,* 59-60.

Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1990). The internal clock: evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception, 19,* 705-743.

Wang. D. (1996). Primitive auditory segmentation based on oscillatory correlation. *Cognitive Science, 20,* 409-456.

Author note

Aleksandar Aksentijevic, The Department of Psychological Sciences, Birkbeck, University of London, now at Department of Psychology, Roehampton University; Paul Barber, The Department of Psychological Sciences, Birkbeck, University of London; Mark A. Elliott, School of Psychology, National University of Ireland, Galway.

This article is based on the research reported in the corresponding author's doctoral dissertation. We wish to thank Denis O'Hora for his useful comments.

Correspondence concerning this article should be addressed to Aleksandar Aksentijevic, Department of Psychology, Roehampton University, Whitelands College, Holybourne Avenue, London, SW154JD, United Kingdom. Electronic mail may be sent to a.aksentijevic@roehampton.ac.uk.

Figure captions

Figure 1. The “unique-template” hypothesis. 500-Hz pips activate a 500-Hz template, The 1000-Hz (deviant) pip activates a 1000-Hz template. This requires partial inhibition of the components belonging to the latter but not the former template (500 Hz, 1500 Hz and 2500 Hz).

Figure 2. Harmonic Reset (HR) as a consequence of the unique-template hypothesis. Bottom: Acoustic waveform of the entrainer. Three 500-Hz pips (B) synchronise the 500-Hz template. The representation is reset by a 1000-Hz deviant pip (D). Top: The result is a periodic oscillatory representation with asymmetric period.

Figure 3. The construction of a target-present (TP) trial. A 3000-ms tone-pip train is split into two portions separated by an arbitrary ISI (D = deviant, B = baseline). The configuration shows a train presented at 33 pps. Target frequencies are integral or fractional multiples of the entrainer frequencies (not shown; see text). TA stimuli (also not shown) carry a single frequency.

Figure 4. Mean response times (± 1 SEM) for HTP and ITP conditions plotted against presentation rate in Experiment 1. Asterisk marks a significant difference at 33pps.

Figure 5. HTP response curve (thick) and its relationship with a function describing the phase angle between the DBBB cycle and entrainer offset (dashed) and a function describing the phase angle between the DBBB cycle and target onset (solid). The dashed curve describes the variation in the alignment between the DBBB cycle and entrainer offset caused by the change in presentation rate. The mismatch between the two phase-angle functions is due to the 100-ms ISI.

Figure 6. Oscillatory representation following the entrainer offset. Stimuli are presented at 33 pps. Deviant-evoked oscillations are highlighted. The trace is assumed to last no more than 400 ms and to decay linearly in the absence of re-synchronization (slope 0.5). Two coherence peaks are predicted at around 110 and 220 ms ISI.

Figure 7. Predicted location of the threshold with respect to ISI. Panel A: Under the assumption of amplitude summation, the difference in slope between two adjacent rates is found by dividing the angle of the maximum slope by the number of rates between 33 pps (maximum coherence) and 22 pps (complete phase cancellation). Panel B: The threshold is placed at the lowest point at which the 33-Hz oscillation can be discriminated from the 32-Hz oscillation. For a representation whose coherence increases monotonically and reaches a maximum at 110 ms ISI, this point lies at 100 ms ISI.

Figure 8. Examining the symmetry of the oscillatory representation. Panel A: Coherence of the trace exceeds the threshold at approximately 100 ms ISI (after four eGBR cycles). If there is asymmetry in the relationship between the two templates, the negative slope should cross the threshold within the grey area. Panel B: A family of curves describing the envelope of the representation.

Figure 9. Mean response times (± 1 SEM) for HTP and ITP conditions plotted against plotted against ISI in Experiment 2. Asterisk marks the significant difference at 106 ms.

Figure 10. Mean response times (± 1 SEM) for HTP and ITP conditions plotted against plotted against ISI in Experiment 4. Asterisk marks the significant difference at 106 ms.

Figure 11. Oscillatory harmonic representation, inferred from Experiments 2, 3 and 4. Panel A: Temporal evolution of the oscillatory trace within 200 ms ISI. Thick curves represent deviant-evoked oscillations. The shaded band represents a 15-ms period during which the trace exceeds the behavioural threshold. The trace is assumed to decay linearly. Panel B: Detail of the coherence window.

Figure 12. Mean response times (± 1 SEM) for HTP and ITP conditions plotted against plotted against ISI in Experiment 5. Asterisk marks the significant difference at 258 ms.

Figure 13. Mean response times (± 1 SEM) for HTP and ITP conditions plotted against plotted against ISI in Experiment 6. Asterisks mark the significant differences at 100 and 250 ms.

Figure 14. Oscillatory trace as inferred from Experiments 1 to 6 and its temporal evolution within 400 ms ISI. As shown in the upper part of the figure, the trace is quasi-periodic and exceeds the behavioural threshold within two windows (shaded bands) created by the phase reset of the entrained 500-Hz oscillatory representation. The lower part of the figure shows the second (resonant) oscillatory cycle, which re-synchronises the primary circuits. For simplicity, it is assumed that the second representation is identical to the first. The onset of the second cycle is at around 120 ms ISI. A possible relationship between the observed effect and the underlying anatomical structures is shown in the centre.

Figure 15. Comparison between predictions of a delay-reset mechanism and an “inherent phase” mechanism. Panel A: The reset is achieved through a 15-ms delay of the deviant-

evoked cycle. Maximum coherence windows occur at approximately 140 and 280 ms ISI (shaded bands). Panel B: Reset mechanism as inferred from the current experiments. All oscillations have the same onset latency (20 ms). The coherence windows are found at around 110 and 250 ms ISI (shaded band).

Figure 16. Inferred relationship between harmonic templates and their oscillatory cycles. (a) Two identical templates with fundamental frequency (f_0) of 500 Hz possess identical phase profiles. (b) A 750-Hz template is distinct and requires a small shift (in this case 90 degrees) to be disambiguated from a 500-Hz template. (c) Two harmonically related templates (f_0 = 500 and 1000 Hz) with opposing phase profiles. Complete phase cancellation is necessary because the two representations are almost identical (a 1000-Hz template is completely subsumed under a 500-Hz template).

Figure 1

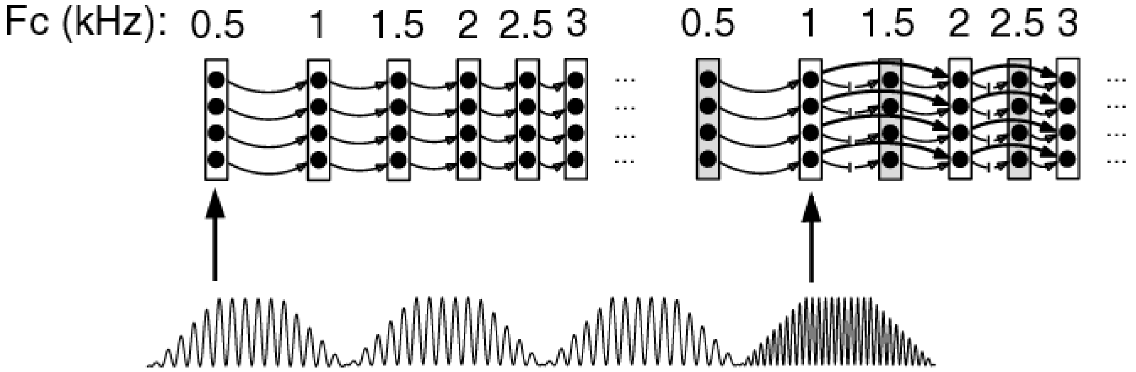

Figure 2

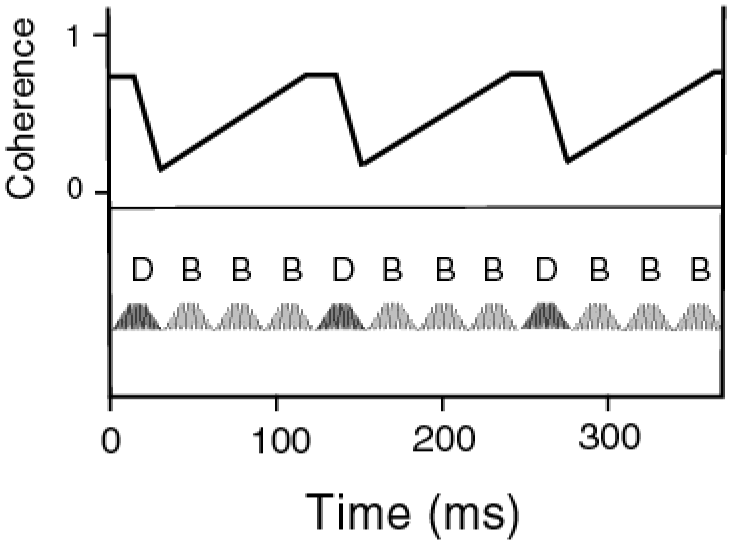

Figure 3

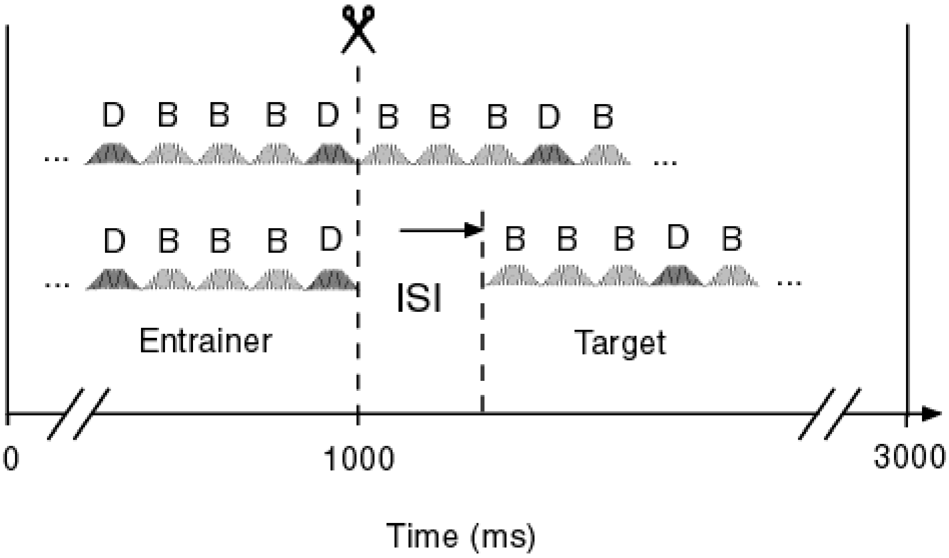
D B B B D B B B D B
...
...
D B B B D
...
ISI
B B B D B
...
Entrainer
Target
0
1000
3000
Time (ms)

Figure 4

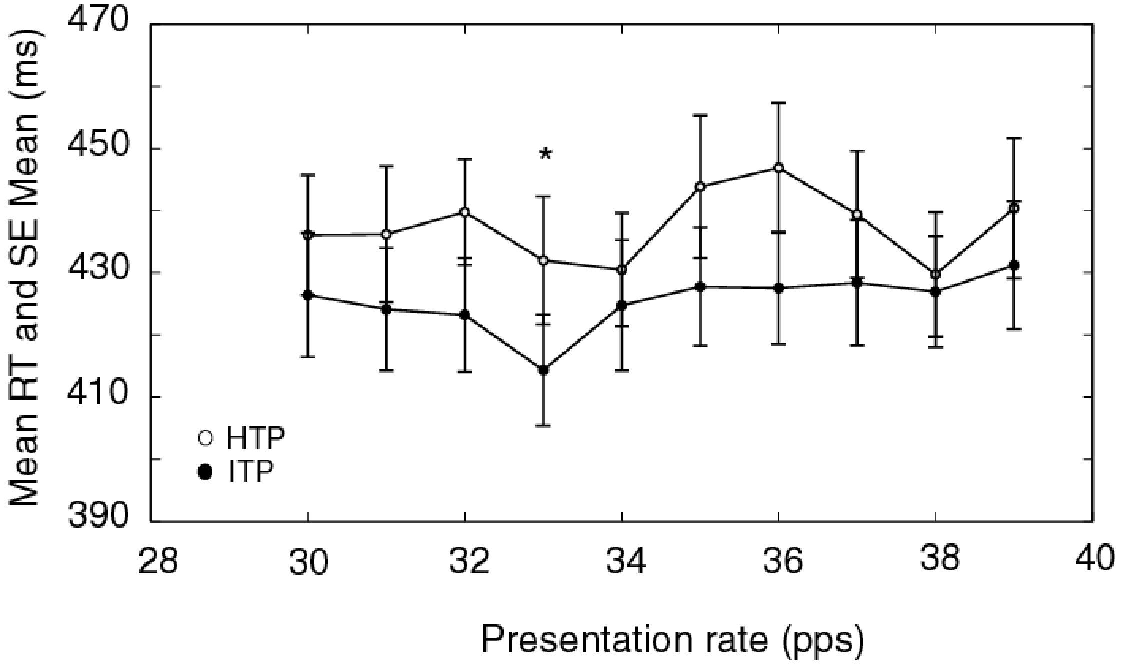
Mean RT and SE Mean (ms)
470
450
430
410
390
28
30
32
34
36
38
40
Presentation rate (pps)
*
○ HTP
● ITP

Figure 5

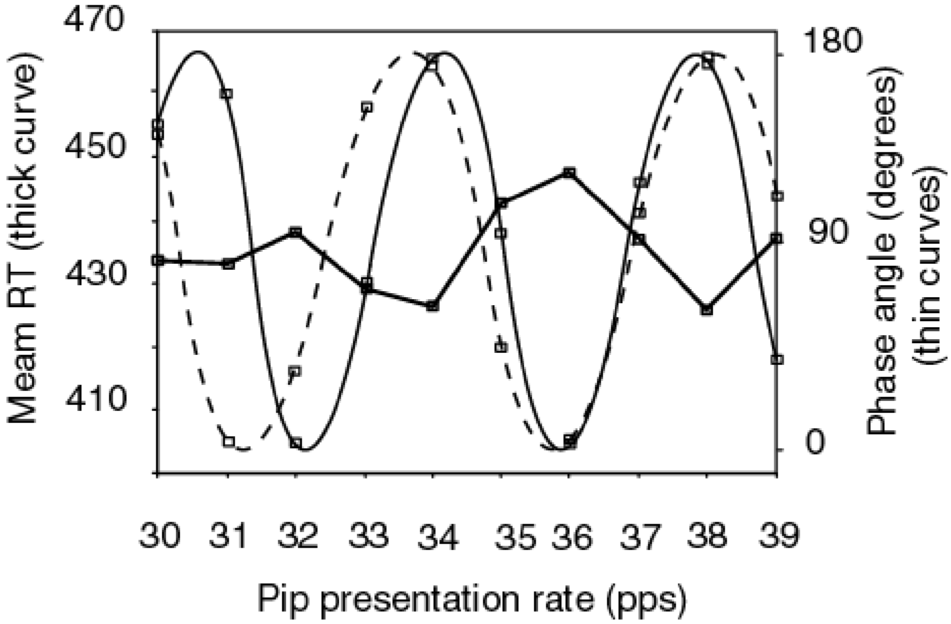
470
450
430
410
Meam RT (thick curve)
180
90
0
Phase angle (degrees)
(thin curves)
30
31
32
33
34
35
36
37
38
39
Pip presentation rate (pps)

Figure 6

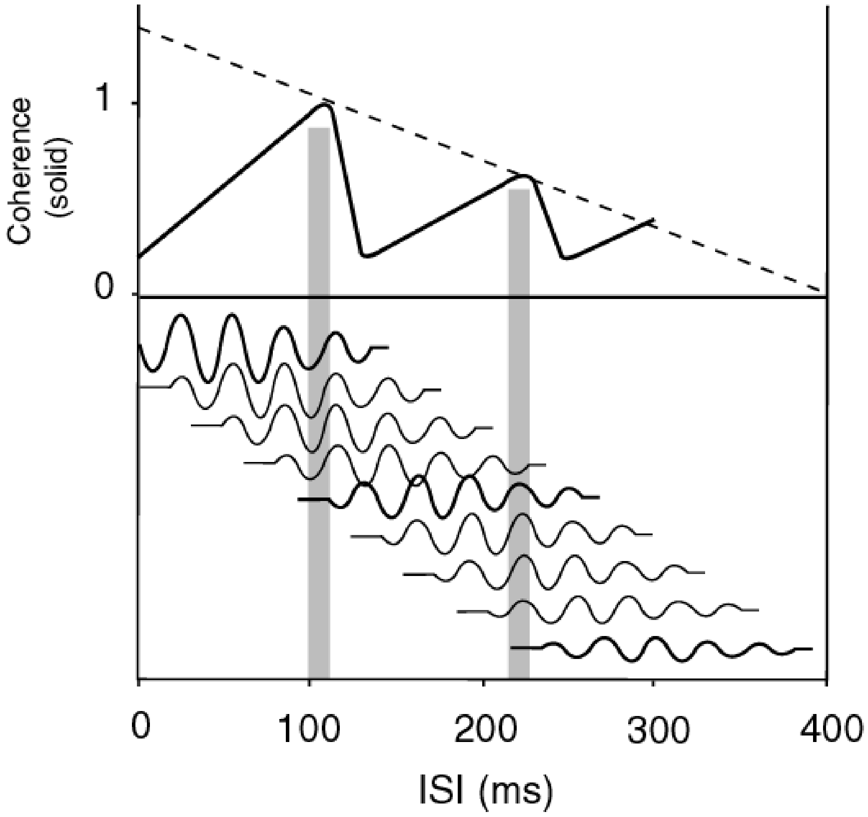

Figure 7

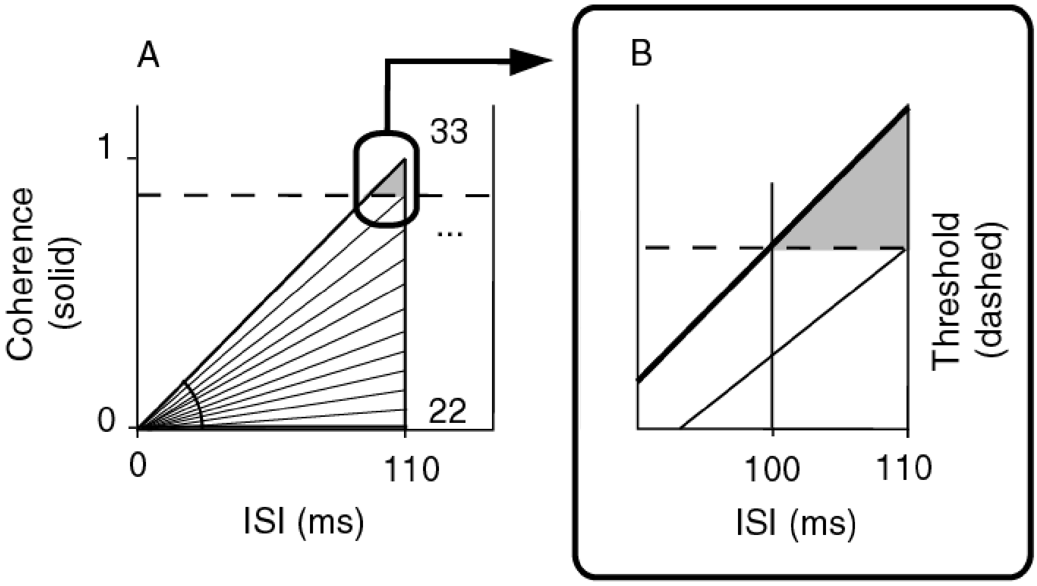
A
Coherence (solid)
1
0
0
110
ISI (ms)
33
...
22
B
Threshold (dashed)
100
110
ISI (ms)

Figure 8

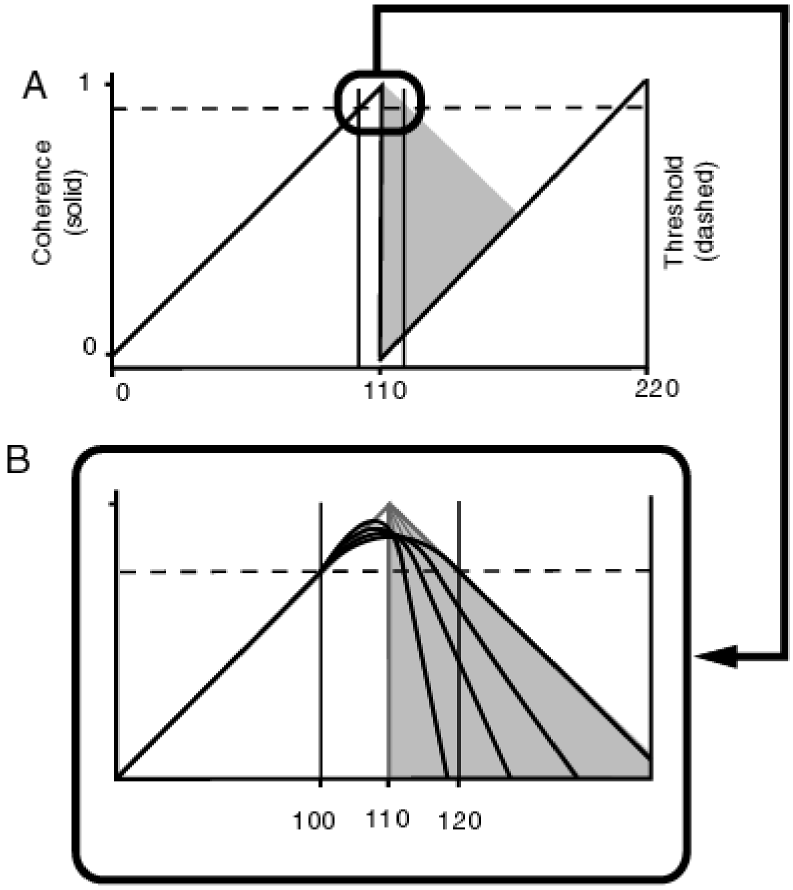

Figure 9

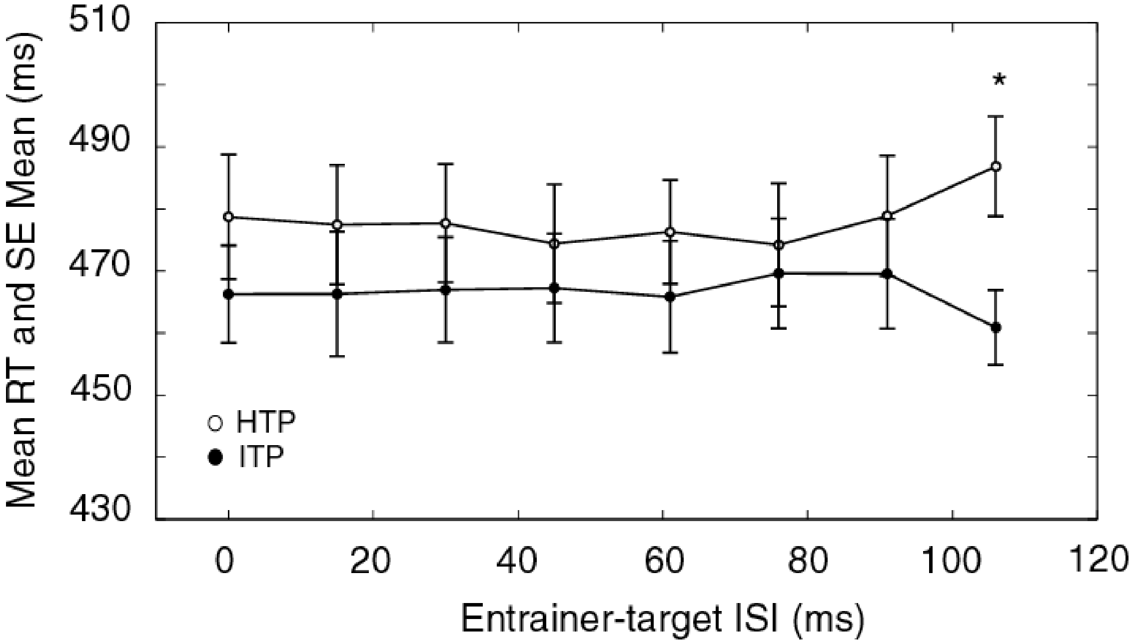
Mean RT and SE Mean (ms)
510
490
470
450
430
0
20
40
60
80
100
120
Entrainer-target ISI (ms)
○ HTP
● ITP
*

Figure 10

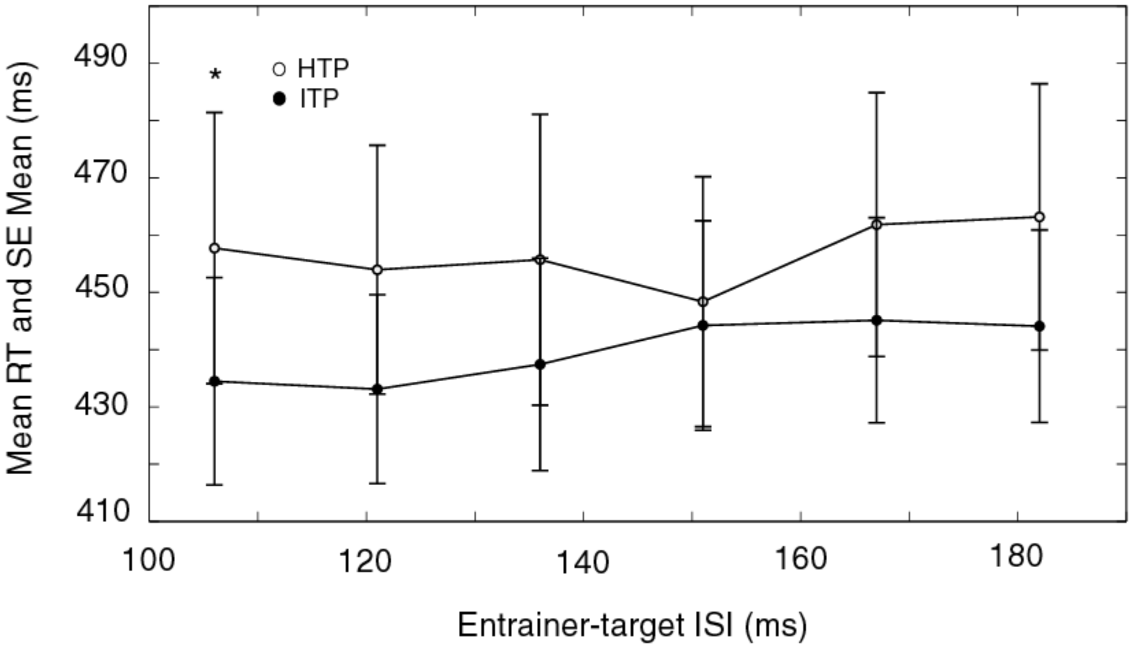
*
○ HTP
● ITP
Mean RT and SE Mean (ms)
410
430
450
470
490
100
120
140
160
180
Entrainer-target ISI (ms)

Figure 11

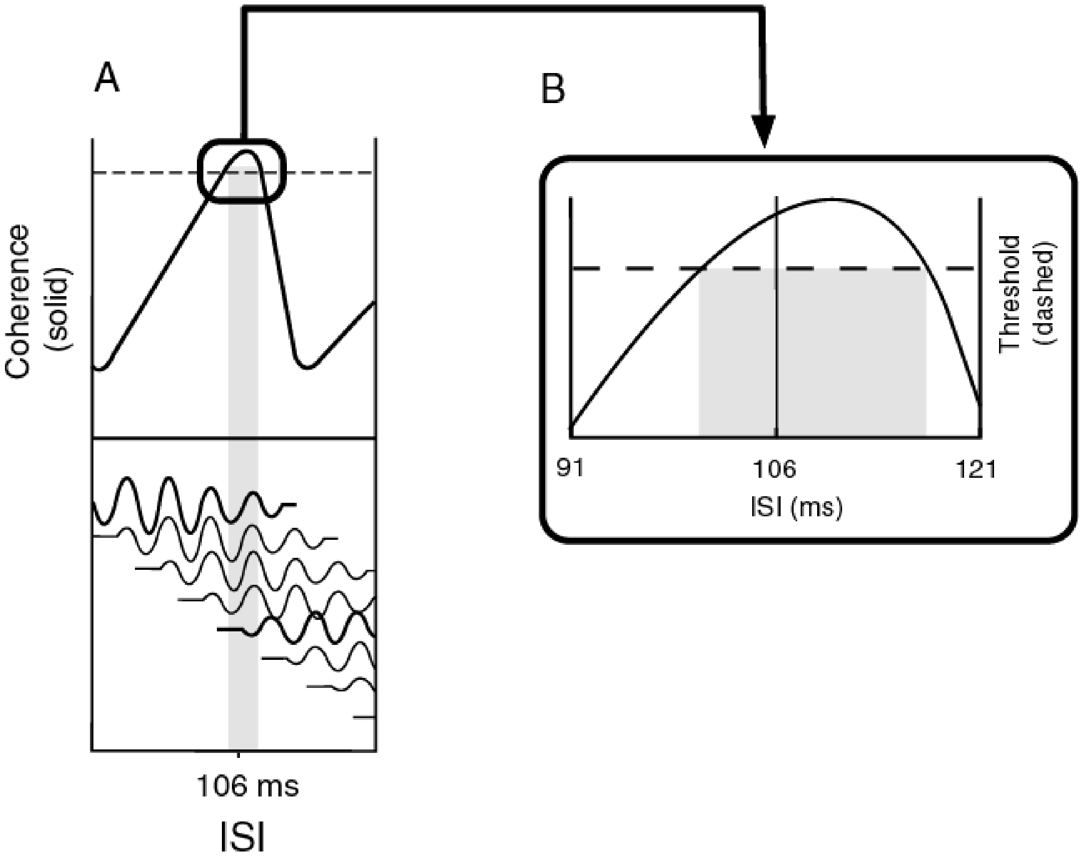
A
B
Coherence (solid)
106 ms
ISI
Threshold (dashed)
91
106
121
ISI (ms)

Figure 12

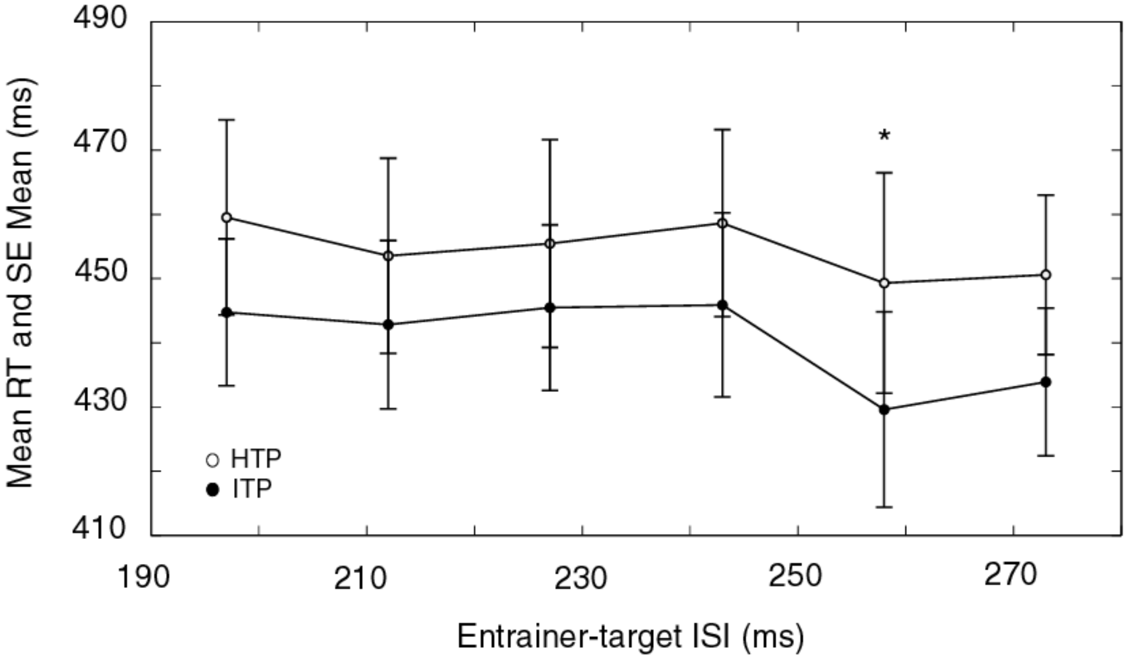
490
470
450
430
410
190
210
230
250
270
Mean RT and SE Mean (ms)
Entrainer-target ISI (ms)
○ HTP
● ITP
*

Figure 13

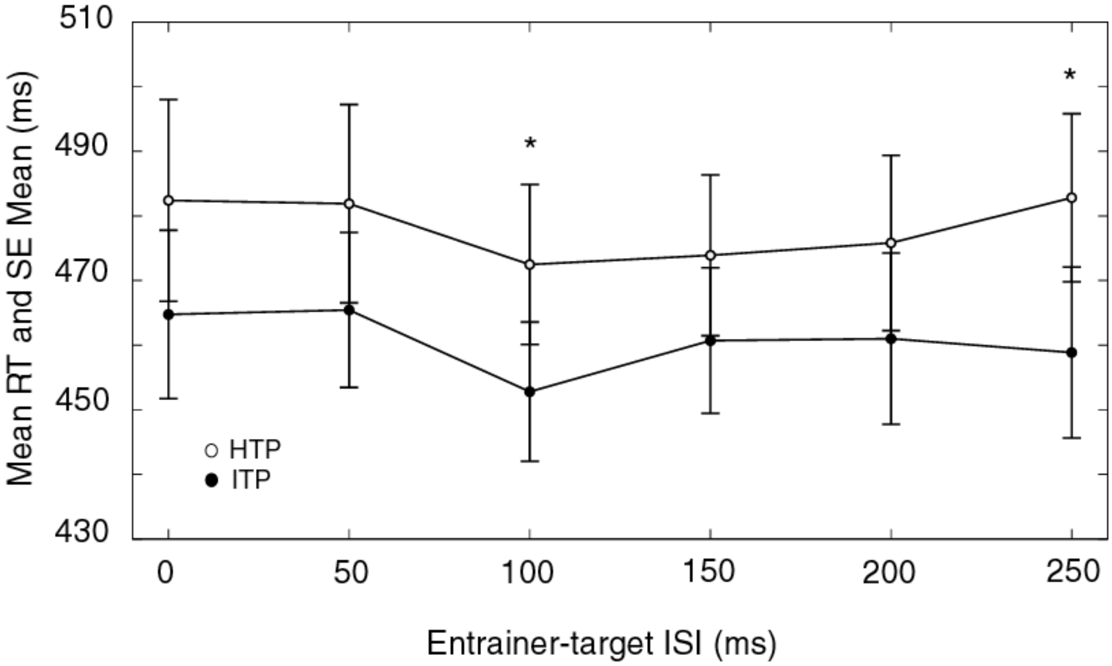
510
490
470
450
430
Mean RT and SE Mean (ms)
0
50
100
150
200
250
Entrainer-target ISI (ms)
○ HTP
● ITP
*
*

Figure 14

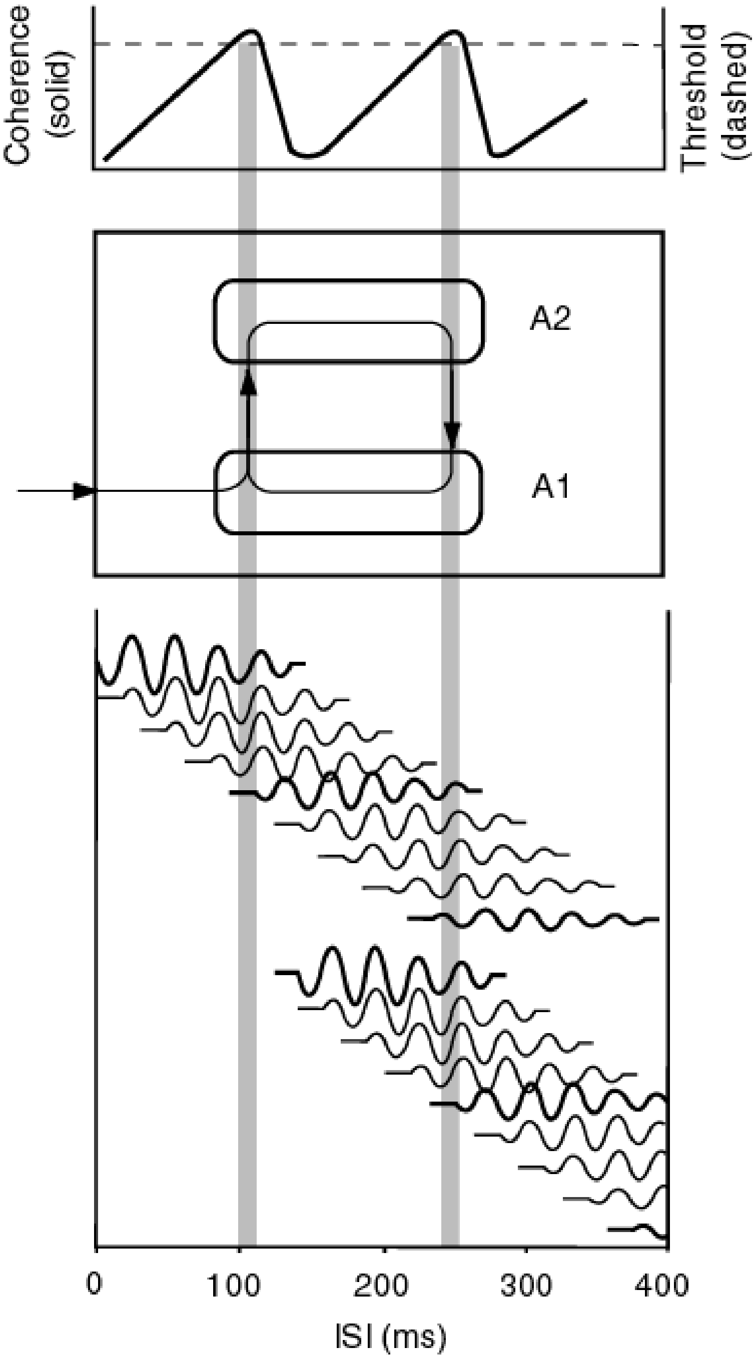
Coherence
(solid)
Threshold
(dashed)
A2
A1
0
100
200
300
400
ISI (ms)

Figure 15

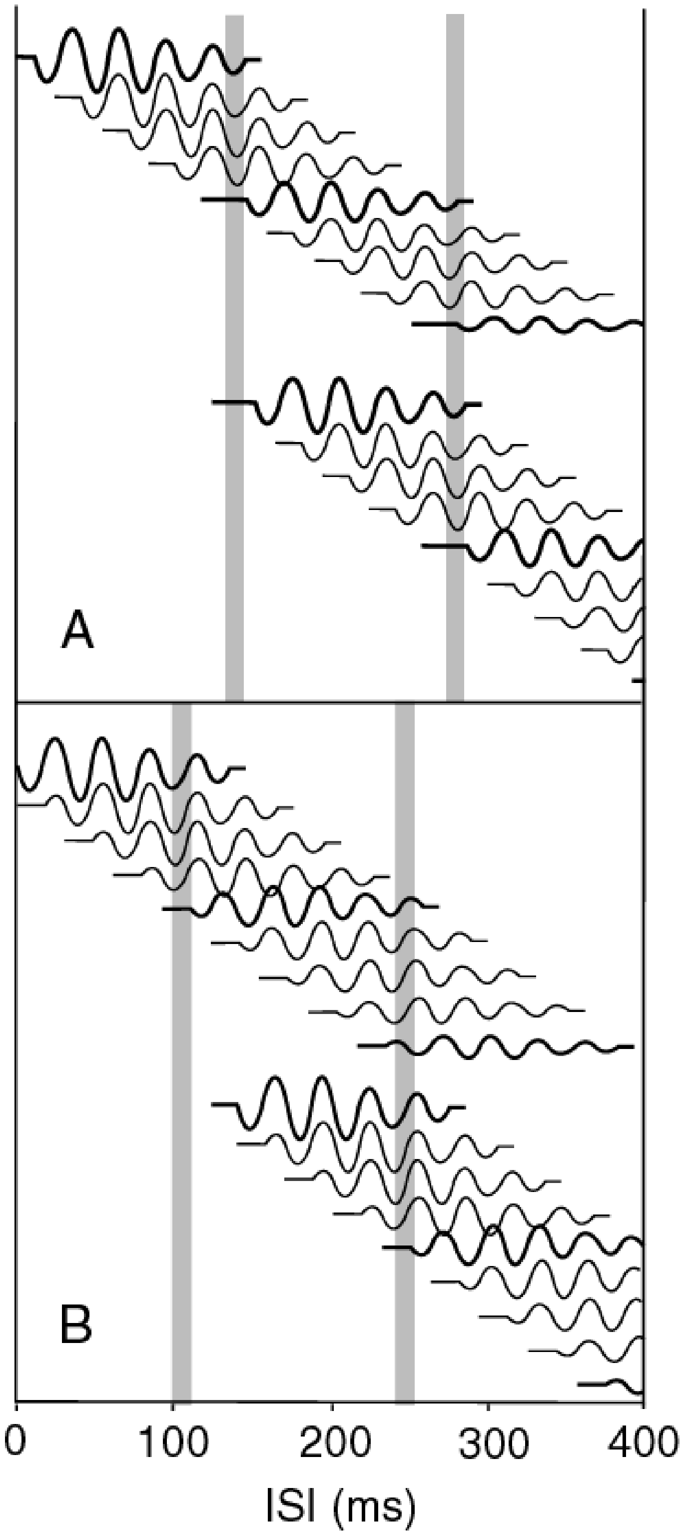

Figure 16