Sensori-motor synchronisation variability decreases as the number of metrical levels in the stimulus signal increases

Guy Madison *

Department of Psychology, University of Umeå, 901 87 Umeå, Sweden

ARTICLE INFO

ABSTRACT

Timing performance becomes less precise for longer intervals, which makes it difficult to achieve simultaneity in synchronisation with a rhythm. The metrical structure of music, characterised by hierarchical levels of binary or ternary subdivisions of time, may function to increase precision by providing additional timing information when the subdivisions are explicit. This hypothesis was tested by comparing synchronisation performance across different numbers of metrical levels conveyed by loudness of sounds, such that the slowest level was loudest and the fastest was softest. Fifteen participants moved their hand with one of 9 inter-beat intervals (IBIs) ranging from 524 to 3125 ms in 4 metrical level (ML) conditions ranging from 1 (one movement for each sound) to 4 (one movement for every 8th sound). The lowest relative variability (SD/IBI < 1.5%) was obtained for the 3 longest IBIs (1600–3125 ms) and MLs 3–4, significantly less than the smallest value (4–5% at 524–1024 ms) for any ML 1 condition in which all sounds are identical. Asynchronies were also more negative with higher ML. In conclusion, metrical subdivision provides information that facilitates temporal performance, which suggests an underlying neural multi-level mechanism capable of integrating information across levels.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Timing accuracy depends on the interval timed. For sequences of intervals, as in production of an isochronous beat, the absolute error is much greater for long intervals, corresponding to slow tempi. In addition to this relation (Woodrow, 1932), timing performance is also characterised by duration–specificity, in the sense that the interval–variability relationship is non-linear. Specifically, the slope of the function relating variability to interval becomes steeper close to 1 s and steeper again above about 2 s (Madison, 2001; Mates, Radil, Müller, & Pöppel, 1994). This is particularly relevant for predictive timing, because when one attempts to synchronise with future events the asynchronies become so large that they become confusing and destroy the perception of a regular beat (Madison & Merker, 2002). These findings lead to an unresolved paradox: How can people synchronise more precisely with musical stimuli than with simple isochronous sequences at slow tempi? Here, I test the hypothesis that explicit metrical structure of music can increase precision by providing additional temporal information.

Almost all music across the world is explicitly or implicitly metrical, which constitutes a very efficient way to accommodate many different durations and rhythmical patterns within a simple and general temporal structure (e.g., Arom, 1991). This structure is often based on binary multiples or subdivisions of intervals. In music notation, a crotchet corresponds to two quavers, which in turn corresponds to two semiquavers, etc. The efficiency of this scheme lies in the fact that each level is inherently isochronous, and that rhythmic patterns are therefore constrained in phase and exhibit quantal steps of duration. By phase I mean that there is always one metrical level that is isochronous, typically that which is perceived as the beat, which puts a certain restriction on the complexity of patterns. Quantal duration emerges from the binary structure such that both the interval duration and the start time of any sound event are restricted to units that are $2^{\pm x}$ of any other interval, where x is an integer. Thus, this relatively simple organisation of time can encompass great rhythmical complexity while being strongly constrained. It has been argued that this organisation is applied because it fits the way the brain organises time (Jongsma, Desain, & Honing, 2004; London, 2002; Madison, 2009; Madison, Gouyon, Ullén, & Hörnström, 2011).

But why would the brain's representation of time be metrical? Perhaps we were at some point of evolution exposed to a selection pressure for producing complex rhythmic patterns, and development of metrical representation was the most parsimonious way to achieve it (e.g., Miller, 2000). It seems unlikely that such an elaborate mechanism would have been developed for this sole purpose, however, as one based on memory capacity alone may have served equally well as a costly signal for sexual selection. An alternative is that the $2^{\pm x}$ relationship is an emergent property of the neural organisation of the auditory system, in analogy with the processing of pitch (Janata et al., 2002).

* Tel.: +46 90 786 64 01 (voice); fax: +46 90786 66 95.
E-mail address: guy.madison@psy.umu.se.

Pitch also utilises integer relations between periods, as is well documented and reflected in octave generalisation, the division of the octave in perfect fourth and fifth, the missing fundamental phenomenon, and timbre as an effect of overtone partials. Resolving the question of origin may therefore require considering the rhythm processing of other animals. While it has been shown that pitch processing is similar for Marmosets (Bendor & Wang, 2005), Macaques (Wright, Rivera, Hulse, Shyan, & Neiworth, 2000), and other mammals (e.g., Amagai, Dooling, Shamma, Kidd, & Lohr, 1999), very little, if any, empirical research addresses metrical rhythm processing in non-humans. Although it seems that many species are able, if not naturally motivated, to produce isochronous sequences (Schachner, Brady, Pepperberg, & Hauser, 2009), this observation does not reveal whether the underlying mechanism is unilevel or multilevel (i.e. metrical) (for a review, see Greenfield, 1994).

Regardless of its evolutionary history, a strong implication of metrical processing in the brain is that metrical information is somehow useful. In other words, stimuli that facilitate metrical processing would lead to better performance than other stimuli (e.g., Keller, 1999; Large & Jones, 1999). I have hitherto only considered the perception and production of rhythmical patterns, but there is a simpler and more straightforward behaviour that may also have adaptive value and to which rhythm is only a means, namely temporal prediction and synchronisation. Accounts of the usefulness of prediction and synchronisation can be found elsewhere (Madison & Merker, 2005; McNeil, 1995; Merker, Madison, & Eckerdal, 2009); suffice it to say that synchronisation among humans is useful for exerting bouts of joint muscle force as well as for producing loud signals by joint vocal exclamations, which makes such signals reach farther and thus heard by a larger number of conspecifics (Merker, 1999).

Synchronisation is characterised by that (1) it is based on prediction rather than reaction, (2) it is restricted to a certain range of rates comprising interval durations from about 150 ms to several seconds, and (3) its variability is roughly proportional to the interval between perceptual or motor events, as will be reviewed in the following. Successfully synchronised movements occur simultaneously or slightly before the target events, and this predictive process must obviously be based on previous events (Madison & Merker, 2004).

Range restrictions can easily be appreciated by considering the inability to move any limb faster than about 10 Hz on the one hand (Kauranen & Vanharanta, 1996), and on the other hand the inability to perceive periodicity for intervals longer than a few seconds (Madison & Merker, 2002). Within this range there are also several points at which the interval–variability relation changes, for as yet unknown reasons, in violation of Weber's law (e.g., Madison, 2001). These findings will be briefly reviewed after the final point, since they are important for interpreting synchronisation variability.

A Weber's law type of relation holds approximately across the range of 300–1800 ms. That the variability of movement timing is larger the longer the interval between movements is as yet not accounted for by any established and verified mechanism. Additional mechanisms must therefore be inferred to mimic real performance, such as a clock with varying inter-tick intervals, an imperfect accumulator of ticks, or a memory decay of the quantity of ticks, for example (for a review, see Gibbon, Malapani, Dale, & Gallistel, 1997). Second, synchronisation studies demonstrate that the interval–variability relation can be manipulated, but they have not suggested any simple underlying principles. Manipulations include tapping to every second, third, or fourth sound (1:x synchronisation) in isochronous, uni-level sequences in which all events are physically identical (Patel, Iversen, Chen & Repp, 2005; Repp, 2003), imagining or performing events between real sounds (x:1 synchronisation) (Repp, 2010), or moving in anti-phase between sounds (Chen, Ding, & Kelso, 2001; Keller & Repp, 2005; Repp, 2005). The general findings from these and similar studies are that (1) explicit subdivision decreases variability, (2) but not as much as beating at the subdivision level (i.e., beating every 1 s while hearing sounds 500 ms apart leads to higher variability than actually tapping every 500 ms),

and (3) variability actually increases when the subdivision intervals are shorter than 200–250 ms (Repp, 2003). It should be noted that none of the studies employing synchronisation (in-phase rather than anti-phase) used stimuli that were explicitly metrical in the sense that slower events representing higher metrical levels were more salient (e.g., louder, different in pitch, or more frequent) than those on lower metrical levels. Rather, events on and between beats were identical, unless they were subjective as in anti-phase and x:1 synchronisation. While it is possible to intentionally attribute the beat to different levels in sequences (Repp, 2003, 2012; Zendel, Ross, & Fujioka, 2011), and while this occurs spontaneously for certain IOIs (Woodrow, 1911 and references therein), it is likely to require more effort and therefore to be less effective than when stimuli are perceptually distinct with respect to a metrical structure.

I will now review breakpoints in the relationship between variability and interval duration in some detail before stating specific hypotheses. Measures of timing performance are a discontinuous function of the duration to be timed, with breakpoints in the range from a few hundred milliseconds to 3 s. Such breakpoints might reflect underlying timing mechanism in the sense of indicating distinguishable contributions by mechanisms that operate in different ranges of durations (e.g., Clarke, Ivry, Grinband, Roberts, & Shimizu, 1996; Lewis, 2002; Mangels, Ivry, & Shimizu, 1998), or as signatures of a specific mechanism (e.g., Miall, 1996). At any rate, breakpoints are hard to reconcile with influential timing models, as reviewed by Buonomano and Laje (2010), Gibbon et al. (1997), or Ivry (1996), for example.

The clearest breakpoint is found at 1800–2000 ms inter-onset interval (IOI) for various temporal tasks. For the discrimination of two intervals, Getty (1975) found that Weber's law holds for 400–2000 ms, but not for 2000–3200 ms. Reproduction of an interval was more accurate when it was subdivided by counting, but only when intervals were shorter than 1.8 s (Grondin, 1992) or ~1.2 s (Grondin, Meilleur-Wells, & Lachance, 1999). Woodrow (1932) noted a breakdown in synchronisation performance to isochronous sound sequences above 2 s, associated with a sudden leap in relative variability (standard deviation divided by duration). Mates et al. (1994) showed that below 1.8 s interval duration, responses occurred on average slightly before the stimuli with a narrow distribution around the time of the stimulus. For 1.8, 2.4, 3.6, and 4.8 s IOI, however, an increasing proportion of the responses occurred some 125–150 ms after the stimuli, indicating simple reaction rather than anticipation, and resulting in a bimodal distribution. Other studies did not find this strong inability to synchronise with long IOIs, probably due to instructions to do so and better trained participants (Repp, 2010; Repp & Doggett, 2007). In conclusion, the breakpoint close to 2 s can be described as a limit for predictive timing, above which the human mean timing error is too large to provide useful anticipation of future events even if their time of occurrence is in principle perfectly predictable (cf. Lejeune & Wearden, 2009).

A more narrow optimal range of temporal performance is frequently reported, for example 400–600 ms (Drake & Botte, 1993), 700–1000 (Lunney, 1974; Woodrow, 1932), 700–1150 (Halpern & Darwin, 1982), and 750–950 ms (Madison, 2001). The differences between these estimates may simply be a consequence of the range of durations employed. The range from 400 to 800 ms IOI is generally considered optimal, and is characterised by slower increase in absolute error with duration, which translates to a constant or slightly decreasing relative error (standard deviation or discrimination error divided by duration) in the range of 2.5–6% (e.g., Fraisse, 1984). Optimal ranges not only might be a consequence of a division between mechanisms or modes of operation, but might also, for example, reflect adaptation to durations that frequently occur in everyday life.

The preceding review of range limits concerns the results for 1:1 synchronisation, but it will be useful in interpreting the results of 1:2, 1:4, and 1:8 synchronisation, corresponding to metrical levels 2–4. The critical difference from previous 1:x tapping studies is that the stimulus sequence will feature metrical levels that are perceptually