Measuring perceptual centers using the phase correction response

Rudi C. Villing · Bruno H. Repp · Tomas E. Ward · Joseph M. Timoney

Abstract The perceptual center (P-center) is fundamental to the timing of heterogeneous event sequences, including music and speech. Unfortunately, there is currently no comprehensive and reliable model of P-centers in acoustic events, so P-centers must instead be measured empirically. This study reviews existing measurement methods and evaluates two methods in detail—the rhythm adjustment method and a new method based on the phase correction response (PCR) in a synchronous tapping task. The two methods yielded consistent P-center estimates and showed no evidence of P-center context dependence. The PCR method appears promising because it is accurate and efficient and does not require explicit perceptual judgments. As a secondary result, the magnitude of the PCR is shown to vary systematically with the onset complexity of speech sounds, which presumably reflects the perceived clarity of a sound's P-center.

Keywords Temporal processing · Perceptual center (P-center) · Phase correction response (PCR) · Sensorimotor synchronization

The fundamental nature of the *perceptual center* (P-center) may be recognized in its relationship to the elementary temporal perceptions of simultaneity, successiveness, temporal order, and interval duration (Pöppel, 1997) and to the higher level perception of temporal patterns, including rhythm. The P-center is the hypothetical specific moment at

R. C. Villing (✉) · T. E. Ward · J. M. Timoney
National University of Ireland Maynooth,
Maynooth, Ireland
e-mail: rudi.villing@nuim.ie

B. H. Repp
Haskins Laboratories,
New Haven, CT, USA

which a brief event is perceived to occur (Morton, Marcus, & Frankish, 1976). From this definition, it follows that when two such events are synchronized, it is their P-centers that are (approximately) synchronous, and when a sequence of events occurs, it is the pattern of P-centers that determines the perceived sequence timing: rhythmic (regular and predictable) or arrhythmic (unpredictable) and expressively or mechanically timed.

But what is an event? An event may be defined as a segment of time that an observer conceives as having a beginning and an end (Zacks & Tversky, 2001); the description and identity of the event result from some integration of the sensations and perceptions that occur during its span. Segmenting continuous experience into discrete events appears to be an important ongoing component of perception (Kurby & Zacks, 2008; Zacks, Speer, Swallow, Braver, & Reynolds, 2007). P-centers are important mainly for events and intervals that are directly sensed, rather than remembered, and take place within the timescale of the psychological present, or about 3 s (see, e.g., Fraisse, 1984; Pöppel, 1997). Furthermore, the P-center concept is applicable primarily to events that are perceived to occur at rather well-defined times—for example, musical tones, speech syllables, visual flashes, and dance movements.

Unfortunately, no single, well-defined, objectively measurable time point of an event has yet been found that reliably corresponds to the P-center. In the auditory modality, for example, rejected candidate time points include the acoustic onset (Morton et al., 1976), absolute or relative onset thresholds (Gordon, 1987), local or global intensity peaks (Gordon, 1987; Marcus, 1981), and the measured vowel onset in syllables (Marcus, 1981). There is an important consequence of the mismatch between the empirically determined P-center and any easily measurable time point of an event: The perceptual timing of a

heterogeneous sequence of events, defined by their pattern of P-centers, cannot generally be determined from any easily measurable pattern, such as the pattern of onsets. Indeed, the P-center term originated when Morton et al. discovered that they could not easily construct a perceptually regular sequence of heterogeneous recorded words for a memory experiment. P-centers can be disregarded only if the sequence is composed of homogeneous events that have well-defined boundaries and do not overlap.

Although these constraints are often not mentioned, they affect many research questions concerning timing. For example, research into sensorimotor synchronization (see Repp, 2005, for a review) is generally constrained to use homogeneous (or nearly homogeneous) event sequences in order to avoid the potential effect of P-center differences between events. Investigations of rhythmic timing and microtiming in speech or music cannot adequately measure performances in which the P-centers of events in a sequence can vary substantially relative to each other. In particular, without knowledge of P-centers, the rhythm of spoken language cannot be measured accurately, and thus questions about the perceived timing of individual languages can be answered only on the basis of flawed or indirect data at best. In general, the P-center is a necessary component of speech, music, and other temporally sensitive activities, and it may well have a part to play in achieving natural interaction and gesture timing for anthropomorphic robot and virtual human models (for a suggestive example, see Murata et al., 2008).

Yet despite their importance for event timing, it is currently rather difficult for researchers to use P-centers to time events. If there were a P-center model that accurately predicted human event timing perception, researchers could avoid measuring P-centers empirically in many cases. Indeed, developing such a model is one of the primary goals of P-center research. Nevertheless, although several P-center models for acoustic events have been proposed (Gordon, 1987; Harsin, 1997; Marcus, 1981; Pompino-Marschall, 1989; Scott, 1993), no widely accepted and generally applicable acoustic P-center model currently exists. Those that do exist make conflicting P-center predictions for the same sounds (Villing, Ward, & Timoney, 2007). In the absence of a reliable model, therefore, P-centers must be measured experimentally.

It is this need to measure P-centers that motivated the present study. A review of previously used measurement methods (described in the following section), supported by our own work with one of them, found that all suffer from various shortcomings. In response, we propose and evaluate a new method based on a synchronous-tapping task that uses a sequence with phase perturbations and adapts the *phase correction response* (PCR; Repp, 2002, 2005) to P-center measurement.

Methods for measuring P-centers

In all existing P-center measurement methods, participants must either consciously classify the temporal pattern of a set of events (as synchronous/asynchronous or isochronous/anisochronous, for example) or synchronize actions with those events. From these responses, the intervals between the P-centers of successive (or even simultaneous) events can be inferred. If, for example, a participant perceives a perfectly isochronous rhythm, the intervals between consecutive P-centers must be equal (except for some perceptual tolerance of deviations). Similarly, if a participant perceives two events as synchronous, the interval between their P-centers must be close to zero; that is, their P-centers must be synchronous. There is one limitation common to all methods, however. Without knowing the absolute location of at least one P-center in the pattern beforehand, the relative locations implied by the intervals between P-centers cannot be used to derive absolute P-center locations.

P-centers mark specific moments in time and must be defined with respect to a time origin for their values to have meaning. We distinguish several P-center variations, on the basis of the time origin used. The *absolute P-center* (or simply P-center) is defined relative to a time origin that is common to the set of events under consideration, such as the objective beginning of a continuous acoustic stimulus. This is the form required to describe a pattern of P-centers when it is either impossible or inconvenient to explicitly segment the continuous stimulus into individual events with well-defined boundaries or when there is reason to suspect that the P-centers of individual events are highly context dependent. The absolute P-center is the most general and useful form, allowing the temporal pattern of an arbitrary set of events to be measured or controlled. Unfortunately, there is no known method of directly detecting the perception of the P-center at the moment it occurs and, consequently, no way to directly measure absolute P-centers.

Morton et al. (1976) hypothesized that the P-center of a sound is independent of context, where context is taken to mean the acoustic features and timing of temporally nearby sounds. According to this *context independence hypothesis*, the temporal location of a sound's P-center is fixed, relative to the sound itself, and does not depend on the events that precede or succeed it in a sequence. (Context dependence, in contrast, would manifest itself as a P-center whose location within an event varied systematically with the temporally neighboring events.)

On the basis of the assumption of context independence, it is useful to define the *event-local P-center* (EPC): the P-center of an event relative to an event-local origin, which is normally the physical onset or start of the event. The EPC can be related to the absolute P-center by the difference $EPC = PC - EO$ (where PC is the absolute P-center and EO

is the event-local origin). Finally, the *delta P-center* (ΔPC) expresses the relationship between the P-centers of two events i and j. It is defined by the difference $\Delta PC_{ij} = EPC_i - EPC_j$ (where ΔPC_{ij} should be read as the delta P-center of event i relative to event j). Figure 1 illustrates the three P-center variants. As a concrete example, consider $EPC_i = 20$ ms and $EPC_j = 100$ ms; that is, the P-center of event j is later (further from its onset) than the P-center of event i. Correspondingly, $\Delta PC_{ij} = -80$ ms; this indicates the onset delay (or in this case, advance) that must be applied to event j to achieve perceptual synchrony with event i. Similarly, perceptual isochrony would be achieved by first making onsets isochronous and then adding the ΔPC delay, so a perceptually isochronous interval of 700 ms would require an interval of $700 + (-80) = 620$ ms between the onsets of event i and event j.

The context independence hypothesis predicts two properties of the ΔPC that the various measurement methods use. First, if the roles of the two sounds in the ΔPC are swapped, the ΔPC will simply change sign—that is, $\Delta PC_{ij} = -\Delta PC_{ji}$—since EPC_i and EPC_j should be invariant under the change of role. Second, the ΔPC of any two sounds may be calculated by simple addition if the ΔPC of each of those sounds relative to a common third sound is known. Specifically, the *indirect* ΔPC of sound i relative to sound k is the sum of the ΔPCs of sounds i relative to j and j relative to k—that is, $\Delta PC_{ik} = \Delta PC_{ij} + \Delta PC_{jk}$. ($\Delta PC_{ij}$ and ΔPC_{jk} are usually *direct* ΔPCs, directly estimated from participant responses, and, assuming that there is no consistent bias, only the variance of the indirect ΔPC will be affected by the sum.) A useful consequence of this additive property is that the ΔPCs of stimulus sets used in different experiments require just one sound in common to be directly comparable if context independence holds.
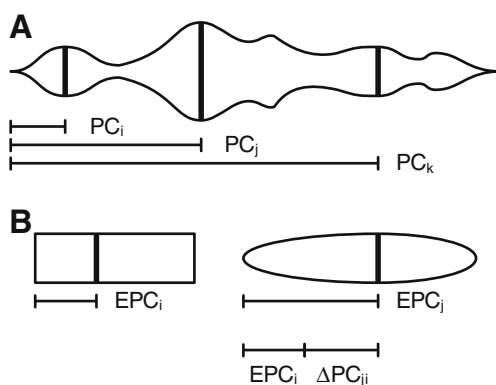
Rhythm adjustment method First described in detail by Marcus (1981), rhythm adjustment is by far the most commonly used method for measuring P-centers (see, e.g., Cooper, Whalen, & Fowler, 1986; Harsin, 1997; Pompino-Marschall, 1989; Scott, 1998). In this method, sequences are constructed by cyclic repetition of a short rhythm using just two sounds, the base sound and the test sound. Typically, a duple rhythm (base–test–base–test . . .)is used, and the base–base interval is fixed, whereas the base–test interval is adjustable.

Figure 2 shows key features of the experimental procedure. Initially, the repeating pattern is not perceptually isochronous. The participant's task is to adjust the timing of the test sound within the cycle until the *point of subjective isochrony* (where consecutive P-center to P-center intervals are equal) is reached. Each final adjustment yields one estimate of the ΔPC of the test sound with respect to the base sound. It is customary to measure the ΔPC for a pair of sounds using both possible role assignments—that is, (base, test) = (sound$_i$, sound$_j$) and (base, test) = (sound$_j$, sound$_i$). Assuming context independence, the resulting measures are averaged so that $\overline{\Delta PC}_{ij} = (\Delta PC_{ij} - \Delta PC_{ji})/2$.

For a set of N sounds, two principal approaches are possible. In the first of these, all possible $N \times (N-1)$ pairs of different sounds are measured (see, e.g., Marcus, 1981). The resulting ΔPC measurements are not all linearly independent; at least some of the measured ΔPCs can be derived from a combination of others. Therefore, multiple linear regression is used to solve for EPCs, with the exception of an unknown constant (and from these, ΔPCs may be derived). Although this approach tends to balance the errors across all sounds in the set, larger overall variability can be expected if some sounds in the set have relatively unclear P-centers (under the assumption that a participant will find it more difficult to detect anisochrony, thus making more variable adjustments, when both P-
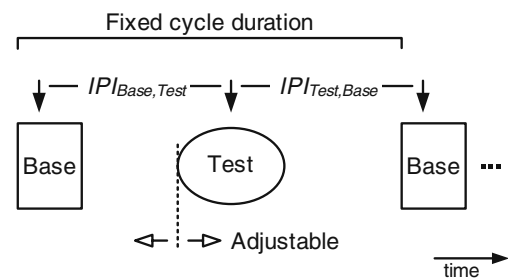

Fig. 1 A continuous stimulus with absolute P-centers measured relative to the stimulus origin (**a**), and discrete events with event-local P-centers (EPCs) measured relative to each event's onset and the delta P-center (ΔPC) difference between EPCs (**b**). P-center locations are marked by vertical lines

Fig. 2 A schematic illustration of the rhythm adjustment method. The sequence consists of cyclic repetition of two sounds, base and test. A participant adjusts the onset timing of the test sound within the cycle until the point of subjective isochrony is reached. At that point the inter-P-center interval between the base and test sounds (IPI$_{Base,Test}$) will approximate that between the test and subsequent base sound (IPI$_{Test,Base}$). Downward-pointing arrows indicate hypothetical P-center locations

centers are unclear than when just one of them is unclear.) It is also worth noting that the regression implicitly depends on P-center context independence; if P-centers are context dependent, the resulting estimates will be invalid. In the second, simpler approach, we designate one sound as a common reference, pair this reference with each of the remaining $N-1$ sounds, and directly measure ΔPCs relative to the reference. Subsequently, the indirect ΔPC for any pair not directly compared can be calculated as described previously. Although this approach is sensitive to the choice of reference sound, the advantages (such as more manageable scalability with number of stimuli and the lower overall variability expected when some sounds have less clear P-centers) make it attractive.

The main benefits of the rhythm adjustment method in general are that it is straightforward for participants to understand and can be implemented without special apparatus. For example, the method is not particularly sensitive to input delays when processing a participant's responses (in contrast to the synchronized-tapping methods to be described subsequently). It does, however, suffer from the disadvantage that participants can find the task rather difficult and fatiguing to perform reliably, since they must continuously judge whether or not the rhythm is isochronous. Indeed, judgment of isochrony seems to be even more difficult when one or both P-centers are unclear.

Other previously used methods A variant of the rhythm adjustment method involves adjusting the test sound to the *point of subjective synchrony,* rather than the point of subjective isochrony, with the base sound (Gordon, 1987; Wright, 2008). The difference between the base and test sound onset times after adjustment is an estimate of the ΔPC. Although the task superficially seems to be closely related to ensemble music performance (and is attractive for that reason), there may be other mechanisms involved in achieving synchronous musical performance (see, e.g., Goebl & Palmer, 2009). In practice, the method has yielded multimodal distributions of ΔPC observations (perhaps implying competing candidate P-centers), but these distributions may simply be artifacts of the method itself. Additional problems encountered with the method include auditory masking (the onset of one sound may mask portions of the onset of the other), stimulus fusion (the two sounds may fuse into a single composite sound), and timbre changes at short onset delays (interference patterns occur if, as a control condition, the base and test sounds are identical).

Fox and Lehiste (1987) employed a constant stimulus modification of the rhythm adjustment method in which participants were presented with a sequence of four sounds (base–base–base–test) and had to choose whether the test sound was presented *too early* or *too late*. Although easy to implement and readily adaptable to execution with multiple simultaneous participants, Fox and Lehiste noted that listeners tend to underestimate the duration of the last interval in the sequence, a behavior that may distort ΔPC measurements (see also Benguerel & D'Arcy, 1986; Repp, 1995). Additionally, the task depends on judging the temporal order of a perceived event and an internally timed moment of isochrony, which may be more difficult than anisochrony judgments in the standard rhythm adjustment paradigm.

The *tap asynchrony* method uses a pacing sequence, consisting of repeated presentations of the test sound at fixed isochronous intervals. The participant's task is simply to tap synchronously with each presentation of the test sound. The mean tap asynchrony (relative to the sound onset) is taken to be an estimate of the EPC, except for some unknown bias; that is, $\overline{A}_i = EPC_i - b$, where $\overline{A}_i$ is the mean tap asynchrony to sound i and b is the bias (relative to the sound's P-center). Subsequently, a ΔPC may be calculated as the difference in asynchronies between sounds; that is, $\Delta PC_{ij} = \overline{A}_i - \overline{A}_j = EPC_i - EPC_j$, if the bias is assumed to be invariant within a participant. Vos, Mates, and van Kruysbergen (1995) showed that the asynchrony varied systematically when parameters (durations and rise time) that affect the P-center of acoustic stimuli were varied. They concluded that participants synchronize with the P-center rather than with the (perceived) onset of sounds, but only Janker (1996) appears to have used the tap asynchrony method as described for general P-center measurement.

The tap asynchrony method is attractive because the synchronous-tapping task is performed automatically by participants and does not require them to make conscious decisions. For this reason, participants generally seem to find the task easier than rhythm adjustment. Furthermore, this method allows ΔPC estimates to be obtained quickly. Unfortunately, the method suffers from disadvantages also. When the pacing sequence consists of short, abrupt sounds (such as the clicks of a metronome), it is commonly found that a participant's taps precede the sounds by some tens of milliseconds on average, a phenomenon referred to as *negative mean asynchrony* (Aschersleben, 2002; Repp, 2005).While this asynchrony is subsumed under the bias, b, there is no guarantee that it is constant across experimental trials or test sounds. In addition, it has been shown to depend on tapping force (Aschersleben, Gehrke, & Prinz, 2004) and musical training (see, e.g., Repp & Dogget, 2007). It is also quite variable both within and between individuals.

Finally, a number of researchers have investigated P-center effects using speech production tasks, in which participants are required to produce specific speech tokens in either a rhythmic framing sentence or a simple repeating sequence (for examples, see Fowler, 1979; Fox & Lehiste,

1987; Perez, 1997; Rapp-Holmgren, 1971; Tuller & Fowler, 1980). Nevertheless, due to the complex nature of the motor task involved in speech production, the variability between repeated productions of the same token, and the limitation to speech sounds only, methods using these tasks are not suitable for general P-center measurement.

In summary, it is clear that there are problems with all existing alternatives to the rhythm adjustment method. Furthermore, even though the rhythm adjustment method is the most commonly used method, it is somewhat time-consuming to run and fatiguing for participants. For this reason, we considered a new method.

The phase correction response method

Research on sensorimotor synchronization—in particular, finger tapping in synchrony with an auditory sequence—has investigated the phase correction process that enables a person to stay in synchrony with a pacing sequence that may incorporate phase perturbations. The PCR, which denotes the phase shift of a tap in response to an immediately preceding phase-shifted event in an otherwise isochronous pacing sequence, is a key feature of this process (Repp, 2002, 2005). The PCR occurs involuntarily and, generally, without a participant's awareness (Repp, 2001, 2002). A phase perturbation can be either a phase shift, which affects the perturbed event and all subsequent events, or an event onset shift (EOS), which affects only the perturbed event. The PCRs elicited are equivalent because a phase shift, by definition, starts with an EOS. A schematic illustration of an EOS and the subsequent PCR is provided in Fig. 3.

As long as the perturbations are within about ±15% of the sequence baseline interonset interval (IOI), the PCR can be well described by a linear model (Repp, 2002; Schulze & Vorberg, 2002). In the linear range, each tap corrects for a certain proportion, α, of the preceding perceived tap–sound asynchrony. Therefore, an effective way to estimate α is to vary EOS magnitude within the range that elicits linear PCR and regress the PCRs onto EOS magnitude,

yielding the so called *PCR function*. The slope of the regression line is the desired estimate of α, as illustrated in Fig. 4.

To apply the PCR to P-center measurement, participants are asked to tap in synchrony with a pacing sequence, in which we denote onset-shifted events the *test events* and the other events the *base events*. A base sound is presented repeatedly at isochronous intervals, while a test sound is inserted occasionally with various EOS values and PCRs are measured in response to each test event. If the base and test events are the same (as is typically the case in PCR research), their EPCs will be identical, and hence, the point of subjective isochrony should occur when the EOS is zero (cf. Fig. 4). If, however, the onset-shifted test event differs from the preceding base events and has a different EPC, its point of subjective isochrony will occur at some EOS value other than zero. For example, if the EPC of the test event is 20 ms later than that of the preceding sounds, the expected PCR would be positive at the point of onset isochrony ($EOS = 0$), and the point of subjective isochrony would occur at an EOS of –20 ms, at which point no PCR is elicited. This point thus corresponds to the x-axis intercept of the PCR function, which now is no longer at the origin.

Since each PCR function is a line, $PCR = b_0 + b_1 x$, defined by the regression constant, b_0, and slope, b_1, the x-axis intercept ($PCR = 0$) may be calculated as $x_{Intercept} = -b_0/b_1$. This intercept value defines the onset anisochrony required to place the test event at the point of perceptual isochrony relative to the base events. We simply negate this intercept value to estimate the ΔPC of the test event relative to the base event. Like rhythm adjustment, the PCR method is used to measure the ΔPC values of mixed sound pairs, and ΔPC
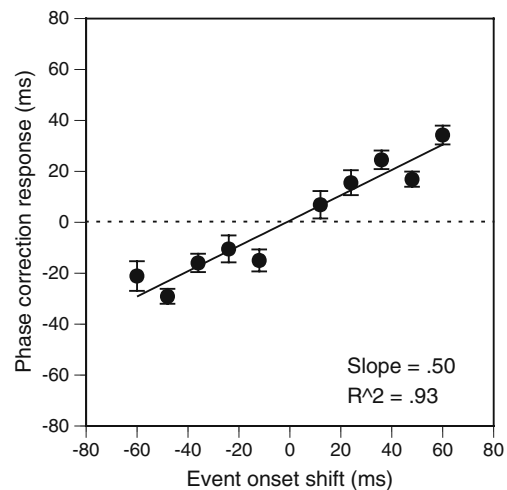
Fig. 4 Illustration of the calculation of the phase correction coefficient α as the slope of a regression line relating the phase correction response to event onset shift magnitude. Each data point is the mean of a number of observations, with standard error bars. The value of R^2 (R^2) indicates the goodness of the linear fit (very good in this example). The baseline interonset interval is 600 ms here
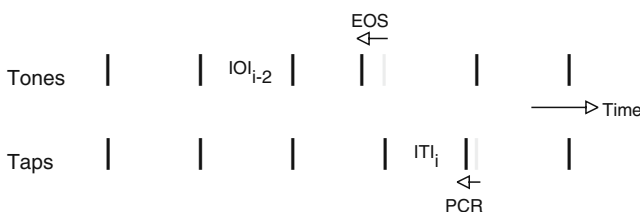

Fig. 3 Schematic illustration of an event onset shift (EOS) in a pacing sequence of tones (IOI = interonset interval) and of the phase correction response (PCR) of the subsequent tap. The PCR is measured by subtracting the baseline (preperturbation) IOI from the current intertap interval (ITI): $PCR = ITI_i - IOI_{i-2}$

estimates can be obtained for both possible role-to-sound assignments.

The PCR method seemed promising, first, because it might prove more precise than traditional methods and, second, because no explicit perceptual judgments are required: The P-center estimates are by-products of a simple synchronization task. In contrast with the tap asynchrony method, which also uses a synchronous-tapping task, the PCR method has less stringent requirements for the tap bias: The bias should be approximately constant across consecutive events but need not be constant within an entire trial sequence. Its primary drawback is that EOS values must be constrained to the range that elicits approximately linear PCRs and, thus, must be centered approximately on the point of subjective isochrony to work correctly. This means that a prior estimate of the P-center difference between two sounds, obtained with some other method, must be available to guide the relative timing of the sounds in the pacing sequence. The PCR method thus is not likely to be useful for a first assessment of sounds that are likely to show large P-center differences. Rather, it is a way of confirming and, perhaps, fine-tuning existing P-center estimates.

Aims of the present study

The primary objective of this study was to evaluate the PCR method and compare it with the most commonly used P-center measurement method, rhythm adjustment. The most fundamental question to be addressed was the following: Do both these methods measure the same percept—namely, the P-center? The agreement of ΔPC estimates between the methods was assessed by using each method to obtain estimates for the same set of stimuli: natural monosyllables that pilot experiments suggested had a wide range of P-centers and a nonspeech reference sound. The ΔPC of each syllable with respect to the reference sound provided a minimum set of estimates that allowed all syllable P-centers to be compared within and between methods.

A second important question concerned the context independence of the ΔPC estimates. All existing methods rely to a greater or lesser degree on the assumption of P-center context independence to derive ΔPC estimates. Although Marcus (1981) tested this hypothesis for rhythm adjustment, it has been examined only one other time, and then with just 1 participant (Eling, Marshall, & van Galen, 1980). We tested it in two ways. First, direct ΔPC estimates were obtained for various syllable–syllable pairs, as compared with indirect ΔPCs calculated by addition of the results for appropriate noise–syllable pairs, and were tested for significant differences. Second, ΔPC estimates were obtained for pairs of sounds in both orders (i.e., with their roles interchanged), because independence predicts a negative relationship between the ΔPCs. If there was any

context dependence due to order, its effect should be greater with the PCR method, since there is a greater difference between the presented event sequences in the two orders using this method (due to repetition of the base sound). Therefore, the PCR method afforded a more stringent test of the hypothesis of context independence.

We also applied the PCR method to homogeneous sound sequences typical of general PCR investigation. Although the PCR from such sequences could not be used to estimate P-centers, it could be used to confirm the accuracy and reliability of the PCR method, because the x-axis intercept was expected to be at zero. Pilot observations had also suggested that the slope of the PCR function might be steeper in homogeneous than in heterogeneous sequences. If confirmed, this novel finding would suggest that phase correction is less effective in the presence of sound change. We had no specific predictions regarding differences in slope among heterogeneous sequences, but we examined this issue as well.

Since P-center measurement methods are often rather time consuming to execute, we hoped to discover which method provided the better return on time invested. We therefore asked two related questions: Which method provides the most accurate between-participants ΔPC estimates (those having the smallest SD), and which method is most time efficient?

This being an international collaboration, rhythm adjustment data (Experiment 1) and PCR method data (Experiment 2) were collected in different laboratories with different equipment and different participant groups. A close agreement of results despite these differences would confirm the validity of the PCR method.

Experiment 1

In Experiment 1, ΔPCs were measured by rhythm adjustment, the most commonly used method. To test the context independence hypothesis, sound pairs used to directly estimate ΔPCs were augmented by additional pairs that could be used to derive equivalent indirect ΔPC estimates. The context independence hypothesis predicts that direct and indirect ΔPC estimates should not differ significantly. Finally, pilot experiments suggested that some sound pairs were harder to align than others. A coarse indicator of difficulty in an adjustment task is the trial duration, which is controlled by participants. It was predicted that the trial duration dependent variable would show an effect of sound pair if there were any pairs that were systematically more difficult than others.

Method

Participants The participants were 2 females and 6 males (21–45 years old) consisting of of 5 unpaid volunteers at

the National University of Ireland Maynooth and authors R. V., T.W., and J.T. These 3 authors had previously participated in a rhythm adjustment experiment, but only R.V. was practiced at the task. None of the participants had any known hearing deficiencies. All were native speakers of English and had a range of music training (0–17 years).

Stimuli The stimuli were seven naturally produced mono-syllables and a synthetic reference sound. The syllables /ba/, /la/, /pa/, /pla/, /sa/, /spa/, and /spla/ were produced by a female native speaker of English and were digitally recorded. After trimming leading and trailing silence, the recordings ranged in duration from 420 to 560 ms. Individual phoneme productions were not edited, so the recordings exhibited some natural variation in those productions. (For example, the /l/ in /la/ differed acoustically from that in /pla/.)

The reference sound was designed not only for the present study but also for anticipated use as a generally applicable reference sound that could be used in a variety of P-center experiments. In general, a good reference sound should be of short duration so that it will not overlap the previous or following event in a perceptually isochronous sequence, it should have a subjectively clear P-center (as sounds with relatively abrupt onsets tend to have), and it should not easily induce auditory streaming effects (Bregman, 1999) when alternating with other stimuli. This last point suggests that the sound should have a spectrum at least somewhat similar to that of the sound under test.

For these reasons, the reference sound was a synthetic, 200-ms, 1:1 mixture of noise and a harmonic complex. The harmonic complex had a 100-Hz fundamental frequency and phases designed to reduce the crest factor (Schroeder, 1970). Both the harmonic complex and the noise had a pink (1/f) spectrum that was intended to be relatively similar to the long-term spectral average of speech (and many natural sounds). The amplitude envelope (a cosine-shaped 20-ms onset and a 180-ms offset) was designed to elicit a relatively early P-center so that test sounds would be likely to have relatively later EPCs and, hence, ΔPCs using the noise as a reference would tend to be positive. Together, the combination of harmonic and noise components, spectral profile, and envelope were expected to mitigate the effects of streaming, and pilot experiments suggested that this was the case. Most participants described the timbre of this reference sound as noiselike, and thus we refer to it as *noise*. For convenience, the seven syllables and the reference sound are hereinafter referred to as: BA, LA, PA, PLA, SA, SPA, SPLA, and N.

Sounds were paired for measurement and formed two main groups: noise–syllable pairs and syllable–syllable pairs. Noise–syllable pairs consisted of each of the seven syllables paired with the reference sound (N). Each pair was tested in both orders (i.e., both possible assignments of sounds to the roles of base and test, as described previously in the rhythm adjustment section). There were thus 14 unique permutations from which ΔPCs could be estimated. Syllable–syllable pairs consisted of two subgroups in which all combinations of three syllables were tested. These were LA–PLA, PLA–SPLA, and LA–SPLA, and PA–SA, SA–SPA, and PA–SPA. Once again, both orders of each pair were tested, so that there were 12 permutations in all. Syllable–syllable pairs provided independent ΔPC estimates that could be compared with those measured for noise–syllable pairs to test the context independence hypothesis. Moreover, the ΔPC estimates for each triplet of syllable–syllable pairs should be internally consistent if ΔPCs are context independent.

Apparatus Custom software, running under Windows XP on a personal computer, controlled the adjustment procedure. The software allowed participants to adjust asynchrony over a ±400-ms range (permitting the sounds to overlap if so chosen), using the keyboard, mouse pointer, or mouse scroll wheel. There was no visible indication of the absolute adjusted asynchrony, and participants could make adjustments as small as 1 ms. Timing of the output audio events was sample accurate. The digital audio for each sequence was mixed in real time at a sampling rate of 48 kHz, converted to analogue by an M-Audio USB Duo 2 audio interface, and presented diotically at a comfortable level using Sennheiser HD280 Pro closed-back circumaural headphones in a quiet room.

Procedure On each trial, a pair of sounds was used to construct a cyclic sequence having a mean IOI of 650 ms and a cycle duration of 1,300 ms. The base sound was fixed to the start of each cycle, while the asynchrony of the test sound relative to the cycle midpoint was adjustable by the participant. At the start of each trial, the initial asynchrony of the test sound was randomly selected from the discontinuous range –200 to –100 ms and 100 to 200 ms. (This choice of values had three desirable properties: The initial rhythm was generally not isochronous and, thus, required adjustment; participants were exposed to trials where the test sound initially occurred both too early and too late; and finally, the asynchrony was not so large that parts of the base and test sounds would overlap.) The trial began when the participant clicked an onscreen button. The task was to adjust the asynchrony of the test sound until the rhythm of the cyclic sequence was perceptually isochronous. Participants could stop and restart the sequence with a buttonpress as necessary if, for example, they became confused about which sound was taking the base or the test role. The most recent adjustment of the asynchrony was always used when the sequence was restarted. The participant

clicked an onscreen button to end the trial. The software saved the initial asynchrony, time-stamped sequence of adjustments, and final adjusted asynchrony for each trial.

Trials were blocked, and each block consisted of trials for all 13 sound pairs in both orders (i.e., 26 trials in all). The order of trials was randomized in every block. Six blocks were presented in the course of two sessions taking approximately 45 min each. Sessions were typically a week apart.

Results

Data for repetitions of each condition were first aggregated within participants. One participant appeared unable to perform the task adequately. Other researchers have excluded participants judged unable to perform the task on the basis of screening trials (Harsin, 1997). Although we did not perform such trials, this participant's data were distinctly different from those of other participants, exhibiting much larger than average variability between replications of each condition, which was consistent with poor ability to perform the task. As a consequence, this participant's data were excluded from the analysis. The main results, averaged across the remaining participants, are shown in Table 1.

The mean trial duration was 48.2 s (SD = 14.4 s). Trial duration is an indicator of task difficulty (although subject to confounding effects, such as participant attention) and was subjected to a two-way repeated measures ANOVA[1] with the independent variables of pair (13 levels) and order (2 levels). Neither the main effects nor their interaction was significant; therefore, it seems that there were no individual conditions in which participants consistently experienced greater or lesser difficulty than average. Furthermore, although some participants reported having more difficulty with noise–syllable pairs than with syllable–syllable pairs, the noise–syllable trial durations (M = 49.8 s, SD = 17.2) were not significantly longer than the syllable–syllable trial durations (M = 46.3 s, SD = 13.8), $t(6)$ = 0.73, p = .49.

The within-participants standard deviation of the ΔPC estimate is expected to indicate both how reliably a participant can reproduce his or her own adjustments and how clear or ambiguous the ΔPC is for a particular sound pair. A two-way repeated measures ANOVA indicated that the effect of pair on the standard deviation of ΔPC was of medium size and was significant, $F(12, 72)$ = 4.11, ε = .17,

---

[1] The Greenhouse–Geisser correction was applied to all repeated measures factors with more than two levels unless two conditions were met: Mauchly's test for sphericity was not significant and ε > .8. Where used, the correction factor ε is reported so that departures of sphericity are clear. The effect size statistic generalized eta squared, $\eta_G{}^2$, is used to facilitate comparability across between-participants and within-participants designs (Bakeman, 2005; Olejnik & Algina, 2003).

p = .04, $\eta_G{}^2$ = .19. Neither the order effect nor the pair × order interaction was significant, $F(1, 6)$ = 1.04, p = .35, $\eta_G{}^2$ = .01, and $F(12, 72)$ = 1.03, ε = .18, p = .39, $\eta_G{}^2$ = .03, respectively. Closer inspection of the differences among pairs revealed that the ΔPC standard deviation was higher for noise–syllable pairs (M = 27.9, SD = 13.2) than for syllable–syllable pairs (M = 17.5, SD = 4.2), and this effect was significant, $t(6)$ = 2.61, p = .04.

From pilot experiments, it was expected that the ΔPC would differ significantly between sound pairs. Within each pair, however, ΔPCs for the two orders should not differ significantly. Table 1 shows that matching ΔPC values differed by less than 10 ms for all pairs but N–BA. A two-way repeated measures ANOVA showed the expected large and significant pair effect, $F(12, 72)$ = 282.77, ε = .29, p < .01, $\eta_G{}^2$ = .96. Both the effect of order and the pair × order interaction were small and nonsignificant, $F(1, 6)$ = 0.51, p = .50, $\eta_G{}^2$ = .01, and $F(12, 72)$ = 1.15, ε = .33, p = .36, $\eta_G{}^2$ = .05, respectively. Consequently, all subsequent analyses used ΔPC values estimated using data from both orders combined.

All direct ΔPC estimates (measured directly between the sounds in question) and indirect ΔPC estimates (measured indirectly via a third sound) of syllable–syllable pairs resulting from the data are shown in Table 2. Pairwise comparisons of direct and indirect ΔPCs for each sound pair yielded just one comparison that approached significance: PA–SA direct compared with PA–SA via N, $t(6)$ = –2.40, p = .05. With Bonferroni correction, none of the differences reached significance, so there was no evidence of P-center context dependence.

Experiment 2

The PCR method for measuring P-centers was assessed in Experiment 2. The primary objectives of this experiment were to measure ΔPC values for comparison with those of rhythm adjustment and to investigate whether there was any evidence of context dependence. In addition, homogeneous event sequences were tested, both as an additional check on the validity of PCR functions obtained and to investigate the effect of different sound sequences on the slope of the PCR function.

Method

Participants There were 9 participants: 8 paid volunteers and author B.H.R. The volunteers (3 men, 5 women) had agreed to serve in a series of sensorimotor and perceptual experiments at Haskins Laboratories and were all highly trained musicians (graduate students at the Yale School of Music, 22–28 years old, who had studied their respective instruments for 13–21 years and played at a professional

Table 1 Direct ΔPC estimates obtained using the rhythm adjustment method

Pair	SD ΔPC		ΔPC		ΔPC (both orders)	
	Fwd	Rev	Fwd	Rev	M	SE
N–BA	19.30	31.45	11.90	-0.19	5.86	4.92
N–LA	24.50	34.41	39.26	38.64	38.95	6.35
N–PA	21.04	26.21	51.02	43.60	47.31	3.74
N–PLA	27.26	31.27	54.17	57.21	55.69	5.86
N–SA	22.03	24.46	109.21	110.71	109.96	3.68
N–SPA	31.40	36.29	181.98	189.79	185.88	8.51
N–SPLA	29.65	30.98	176.45	173.83	175.14	5.50
LA–PLA	13.97	13.43	16.33	9.95	13.14	0.99
LA–SPLA	17.39	19.02	138.57	133.81	136.19	3.02
PLA–SPLA	22.58	19.51	117.86	119.50	118.68	2.62
PA–SA	14.67	18.08	56.14	48.64	52.39	2.06
PA–SPA	22.24	16.88	128.05	128.38	128.21	2.48
SA–SPA	14.81	17.90	67.79	67.76	67.77	2.11

Note. Each pair (sound$_j$–sound$_i$) acted in the roles base–test in the forward order (Fwd) and test–base in the reverse order (Rev). All ΔPC values shown are for sound$_i$ relative to sound$_j$; thus, reverse order ΔPCs, measured for sound$_j$ relative to sound$_i$, were negated. The *SD* ΔPC measure is the average within-participants standard deviation of the ΔPC. The between-participants mean (M) and standard error (SE) of the ΔPC estimates when data from both orders is pooled are shown in the final two columns. All values are in milliseconds

level). Although music training was not required for the task, we took advantage of the ready availability of this rhythmically skilled and highly motivated group of participants. Author B.H.R. was 63 years old at the time, has been an active amateur pianist all his life, and is highly experienced in synchronization tasks.

Stimuli The same eight sounds as in Experiment 1 were used. Again there were seven noise–syllable pairs, consisting of each syllable paired with the reference sound, and six syllable–syllable pairs (LA–PLA, PLA–SPLA, LA–SPLA, PA–SA, SA–SPA, and PA–SPA). These two groups were used to form mixed sequences (in which the base sound and test sound differed), and each pair was tested in both orders

(with each sound serving once as the base sound and once as the test sound). In addition all eight sounds were tested singly in homogeneous sequences (in which the same sound served as base and test sound). Taken together, there were 34 distinct sequences to be tested. These were divided into three sets: Sets 1 and 2 both contained various mixed pair sequences and shared the N–BA sequences in common (for consistency checking); set 3 also contained some mixed pair sequences but consisted primarily of homogeneous sequences.

The PCR method requires initial ΔPC estimates for all sounds to be used in mixed sequences so that EOS perturbations of the test sound can be approximately centered about the point of subjective isochrony. Since the

Table 2 Direct and indirect ΔPC estimates for syllable–syllable pairs obtained with the rhythm adjustment method

Pair	ΔPC		ΔPC via N		ΔPC via Syllable		
	M	SE	M	SE	Syl.	M	SE
LA–PLA	13.14	2.40	16.74	3.78	SPLA	17.51	2.88
LA–SPLA	136.19	2.48	136.19	3.96	PLA	131.82	3.17
PLA–SPLA	118.68	3.22	119.45	6.12	LA	123.05	3.28
PA–SA	52.39	1.17	62.65	4.17	SPA	60.44	3.51
PA–SPA	128.21	3.52	138.57	7.41	SA	120.17	3.19
SA–SPA	67.77	3.35	75.92	4.70	PA	75.82	4.06

Note. For each pair (sound$_j$–sound$_i$), the direct ΔPC of sound$_i$ relative to sound$_j$, ΔPC_{ij}, is reproduced from Table 1 for comparison. Indirect ΔPCs were calculated via a third sound, k, such that $\Delta PC_{ij} = \Delta PC_{ik} + \Delta PC_{kj}$. The identity of sound k was either the reference noise (ΔPC via N) or a syllable (Syl.). All values are in milliseconds

results of Experiment 1 were not yet available, author R.V. ran himself in a pilot adjustment experiment (with a mean IOI of 600 ms and an adjustment range of just ±250 ms), testing all noise–syllable pairs 4 times in both orders. His estimated ΔPCs relative to N (analyzed as in Experiment 1) were 7, 42, 53, 55, 106, 184, and 183 ms for BA, LA, PA, PLA, SA, SPA, and SPLA, respectively. These values were used to calculate the (estimated) delay necessary to achieve subjective isochrony for any pair of sounds; any EOS was added to this delay. Thus, for example, when LA was the base sound, SPLA was the test sound, and $EOS = 0$, SPLA was delayed by $42 – 183 = –141$ ms (i.e., advanced by 141 ms).

Each trial consisted of a nearly isochronous sequence of varying length in which a base sound occurred repeatedly and a test sound was inserted from time to time. Each sequence contained 11 test sounds, with the number of intervening base sounds varying randomly from 4 to 6. The first test sound occurred in the eighth sequence position at the earliest. The IOI between base sounds was 700 ms, which prevented any overlap of base and test sounds. The 11 test sounds occurred at temporal offsets (EOS values) ranging from –50 to 50 ms, in increments of 10 ms, relative to the estimated point of subjective isochrony. (Thus, e.g., in the LA–SPLA condition considered above, delays of SPLA would range from –191 to –91 ms.) The order of EOS values within a sequence was random.

Apparatus The experimental procedure was controlled by customized MAX/MSP 4.6.3 software (designed for MIDI applications) running on an Intel iMac computer (OS 10.4.10). The timing accuracy of the sequential audio output, which was controlled by the MSP (signal processing) component of the software, was verified by acoustic measurements to be within 1 ms. Taps were registered by a Roland SPD-6 electronic percussion pad connected to the computer via a MOTU Fastlane MIDI interface. Sound sequences were presented diotically over Sennheiser HD540 Reference II headphones.

Procedure Each stimulus set, repeated 5 times in different random orders (blocks), required a separate session of about 1 h. The order of sets 1 and 2 was varied between participants; the two sessions were typically 1 week apart. Set 3 was presented at a later time.

Participants sat in front of the computer and tapped manually on the percussion pad, which they held on their lap. Participants were free to tap in any style they preferred. They started each sequence by pressing the space bar on the computer keyboard and started tapping with the third sound they heard. They were instructed to stay in synchrony throughout and to ignore any small deviations from temporal regularity in the sequence. After each presentation of the block of trials, they saved their data in a file.

Results

The PCR to each test sound EOS was calculated by subtracting the baseline IOI (700 ms) from the interval between the two taps coinciding, respectively, with the test sound and the following base sound (cf. Fig. 3). Occasionally, a PCR could not be calculated, because one or both of the critical taps had failed to be registered or were anomalous (double taps or unusually large asynchronies, having z scores >3.29). A total of 0.3% of the PCR data was excluded due to these causes. Simple linear regression of the PCRs on EOS magnitude was used to estimate the parameters of the PCR function separately for each participant, sound pair, and order. The test sound delay (added for stimulus presentation on the basis of the preliminary ΔPC estimates) was subtracted from the PCR function's x-axis intercept prior to calculating the corresponding ΔPC estimate.

The main results, averaged across participants, are shown in Table 3 (mixed sequences) and Table 4 (homogeneous sequences). Sounds within each pair are ordered so that the less complex sound, which is also the sound with the earlier EPC, comes first. Within Table 3, noise–syllable sequences are followed by syllable–syllable sequences. All results for the pair N–BA were averaged; this pair had been presented in two separate sessions as a consistency check (with highly consistent results).

ΔPC estimates The ΔPC estimates for homogeneous sequences should all be zero since the (identical) test events should be subjectively isochronous when there is no onset shift. It is clear from Table 4 that these ΔPC values deviate very little from zero (less than 5 ms in all cases). Each deviation was subjected to a *t* test, and although the LA and PA deviations were individually significant, with Bonferroni correction none of the deviations reached significance.

As with rhythm adjustment, ΔPC estimates for mixed sequences should differ significantly between sound pairs but should not differ significantly between orders for any single pair. A two way repeated measures ANOVA conducted on the ΔPC estimates for mixed sequences showed the expected large and significant effect of pair, $F(12, 96) = 325.88$, $\varepsilon = .21$, $p < .01$, $\eta_G^2 = .93$. The main effect of order was small and far from significance, $F(1, 8) = 0.29$, $p = .61$, $\eta_G^2 = .01$, but the pair × order interaction approached significance, although its effect was small, $F(12, 96) = 2.41$, $\varepsilon = .39$, $p = .058$, $\eta_G^2 = .05$. Bonferroni post hoc tests revealed an individually significant difference between orders only for PLA–SPLA, $t(8) = –2.33$, $p < .05$; no other comparisons were significant.

Since there was no reliable effect of order, the direct ΔPC estimates for both orders were averaged to form a single estimate for each pair, as is customary for rhythm

Table 3 PCR slope and ΔPC estimates from mixed event onset shift sequences

Pair	Order	Slope		ΔPC	
		M	SE	M	SE
N–BA	Fwd	0.56	0.05	3.82	4.86
	Rev	0.65	0.06	9.54	2.78
N–LA	Fwd	0.55	0.06	39.46	5.44
	Rev	0.68	0.06	46.45	3.57
N–PA	Fwd	0.54	0.08	62.77	8.53
	Rev	0.70	0.06	59.68	2.22
N–PLA	Fwd	0.50	0.06	56.36	7.17
	Rev	0.69	0.04	58.29	2.99
N–SA	Fwd	0.54	0.08	125.05	7.89
	Rev	0.66	0.06	113.38	2.16
N–SPA	Fwd	0.52	0.08	181.07	4.68
	Rev	0.63	0.05	180.99	4.62
N–SPLA	Fwd	0.54	0.06	183.73	6.65
	Rev	0.68	0.08	182.76	4.35
LA–PLA	Fwd	0.54	0.06	17.08	6.53
	Rev	0.61	0.06	10.92	3.50
LA–SPLA	Fwd	0.50	0.05	137.64	8.50
	Rev	0.61	0.04	137.02	7.81
PLA–SPLA	Fwd	0.56	0.04	133.78	5.55
	Rev	0.53	0.06	113.37	7.35
PA–SA	Fwd	0.66	0.06	55.00	3.41
	Rev	0.62	0.06	47.51	4.04
PA–SPA	Fwd	0.54	0.05	117.89	6.17
	Rev	0.58	0.07	115.38	3.88
SA–SPA	Fwd	0.53	0.08	70.72	5.66
	Rev	0.60	0.05	66.69	4.16

Note—. Each pair (sound$_j$–sound$_i$) acted in the roles base–test in the forward order (Fwd) and test–base in the reverse order (Rev). All ΔPC values shown are for sound$_i$ relative to sound$_j$; thus, reverse order ΔPC values, measured for sound$_j$ relative to sound$_i$, were negated. All values are in milliseconds

Table 4 PCR slope and ΔPC estimates from homogenous event onset shift sequences

Sound	Slope		ΔPC	
	M	SE	M	SE
N	0.84	0.07	0.66	1.11
BA	0.74	0.05	0.49	1.45
LA	0.70	0.05	4.43	1.66
PA	0.70	0.06	4.12	1.61
PLA	0.62	0.05	3.35	2.84
SA	0.61	0.06	-0.14	2.88
SPA	0.53	0.05	2.42	2.40
SPLA	0.45	0.04	2.13	3.42

Note. ΔPC estimates shown are for each sound relative to itself. All values are in milliseconds

adjustment. We also calculated indirect ΔPC estimates where the data permitted. All these ΔPC estimates are shown in Table 5. As usual, context independence predicts that indirect and direct ΔPC estimates should not differ significantly, and this was tested by pairwise comparisons of direct and indirect estimates for each syllable–syllable pair (10 comparisons). Only the comparison of the direct estimate with the indirect estimate (via N) for the pair PA–SPA reached individual significance, $t(8) = -2.97$, $p < .05$. With Bonferroni correction, none of the comparisons were significant, and so there was no evidence of context dependence.

PCR function standard error of the estimate and slope Within-participants PCR variability was summarized by the PCR function standard error of the estimate (*SEE*). This statistic did not exhibit any consistent pattern (grand $M = 24.6$, $SD = 7.1$) and is not shown for that reason. A one-way repeated measures ANOVA showed no significant effect of sound on the *SEE* for homogeneous sequences, $F(7, 56) = 1.35$, $\varepsilon = .52$, $p = .28$, $\eta_G^2 = .03$. For mixed sequences, a

Table 5 Mean direct and indirect ΔPC estimates from the phase correction response method

Pair	ΔPC	
	M	SE
N–BA	6.68	3.34
N–LA	42.96	2.89
N–PA	61.22	4.88
N–PLA	57.33	3.78
N–SA	119.22	4.41
N–SPA	181.03	3.55
N–SPLA	183.24	3.64
LA–PLA	14.00	3.02
via N	14.37	2.92
LA–SPLA	137.33	7.12
via N	140.28	3.03
PLA–SPLA	123.57	4.81
via N	125.91	1.52
via LA	123.33	6.00
PA–SA	51.26	2.07
via N	57.99	2.11
PA–SPA	116.64	3.71
via N	119.81	4.80
SA–SPA	68.71	2.21
via N	61.81	4.30
via PA	65.38	3.35

Note. For each pair (sound$_j$–sound$_i$), the ΔPC estimate shown is for sound$_i$ relative to sound$_j$; that is, ΔPC_{ij}. Indirect ΔPC estimates via a third sound, k, were calculated as usual, such that $\Delta PC_{ij} = \Delta PC_{ik} + \Delta PC_{kj}$. All values are in milliseconds

two-way repeated measures ANOVA revealed that the order effect was nearly significant but small, $F(1, 8) = 5.25$, $p = .051$, $\eta_G^2 = .01$. Neither the pair effect nor the pair × order interaction was significant, $F(12, 96) = 1.09$, $\varepsilon = .33$, $p = .38$, $\eta_G^2 = .02$, and $F(12, 96) = 1.11$, $\varepsilon = .29$, $p = .37$, $\eta_G^2 = .02$, respectively.

The slope of the PCR function affects the confidence interval of within-participants ΔPC estimates, with shallower slopes resulting in larger confidence intervals and less certain estimates. In general, slopes were not excessively shallow, although they were rather variable (grand $M = 0.60$, $SD = 0.19$). Slope also showed a clear participant effect: Some participants exhibited consistently larger or smaller slopes than did others. Inspection of the data also revealed some systematic variation. There was a wide range of mean slopes obtained from homogeneous sequences, with the steepest slope for N and the shallowest slopes for the syllables starting with consonant clusters. A one-way repeated measures ANOVA on these data showed that the differences were substantial and highly significant, $F(7, 56) = 12.59$, $\varepsilon = .55$, $p < .001$, $\eta_G^2 = .37$.

The experimental design included two subsets of sounds incorporating all possible combinations of base and test sound: N, LA, PLA, and SPLA, and N, PA, SA, and SPA. Figure 5 shows the PCR slope for each combination of base sound and test sound measured. Several effects are apparent. First, the range of slopes for mixed sequences tends to be smaller than the range for homogeneous sequences. Second, slopes show systematic variation by test sound for each base sound. This variation seems to follow the same trend as the corresponding homogeneous sequence slopes, except when N is the base sound. Finally, slopes for each test sound were generally (but not always) larger when the sequence was homogeneous rather than mixed. A two-way repeated measures ANOVA on the subset of sounds N, LA, PLA, and SPLA showed no significant effect of the base sound, $F(3, 24) = 1.00$, $\varepsilon =$

.75, $p = .40$, $\eta_G^2 = .02$, a highly significant, moderate-sized test sound effect, $F(3, 24) = 25.52$, $p < .01$, $\eta_G^2 = .21$, and a small to medium interaction effect that approached significance, $F(9, 72) = 2.62$, $\varepsilon = .40$, $p = .06$, $\eta_G^2 = .09$. Planned contrasts indicated that homogeneous and mixed sequence slopes were not significantly different, $F(1, 8) = 3.17$, $\varepsilon = .28$, $p = .11$. A similar two-way repeated measures ANOVA on the subset defined by N, PA, SA, and SPA once again showed no significant effect of the base sound, $F(3, 24) = 1.26$, $\varepsilon = .65$, $p = .31$, $\eta_G^2 = .02$, a moderate, highly significant effect of the test sound, $F(3, 24) = 14.34$, $\varepsilon = .80$, $p < .01$, $\eta_G^2 = .11$, and a small nonsignificant interaction effect, $F(9, 72) = 1.72$, $\varepsilon = .46$, $p = .17$, $\eta_G^2 = .06$. Again, planned contrasts indicated that the differences between homogeneous and mixed sequence slopes were not significant, $F(1, 8) = 2.73$, $\varepsilon = .25$, $p = .14$.

Method comparison

ΔPC estimate consistency

An important motivation for this work was to investigate whether the new PCR method would measure the same percept as the rhythm adjustment method. The syllable–noise ΔPC estimates for both methods are shown in Fig. 6, and it is apparent that there are no obvious systematic differences between them. To confirm this, a two-way mixed ANOVA was performed. This revealed no significant effect on the ΔPC estimates by either method or the method × pair interaction, $F(1, 14) = 0.50$, $p = .49$, $\eta_G^2 = .01$, and $F(12, 168) = 1.78$, $\varepsilon = .37$, $p = .14$, $\eta_G^2 = .08$, respectively.

Accuracy and efficiency

Method accuracy was evaluated by comparing the between-participants standard deviations of ΔPC estimates averaged

Fig. 5 Mean slope of the phase correction response (PCR) function for all combinations of the sounds N, LA, PLA, SPLA (**a**) and N, PA, SA, SPA (**b**). The slopes for all test sounds are clustered for each base sound. Mixed sequence slopes are shown with empty symbols, whereas homogeneous sequences (with identical base and test sounds) are shown with filled symbols. Error bars show 95% confidence intervals for the mean slope
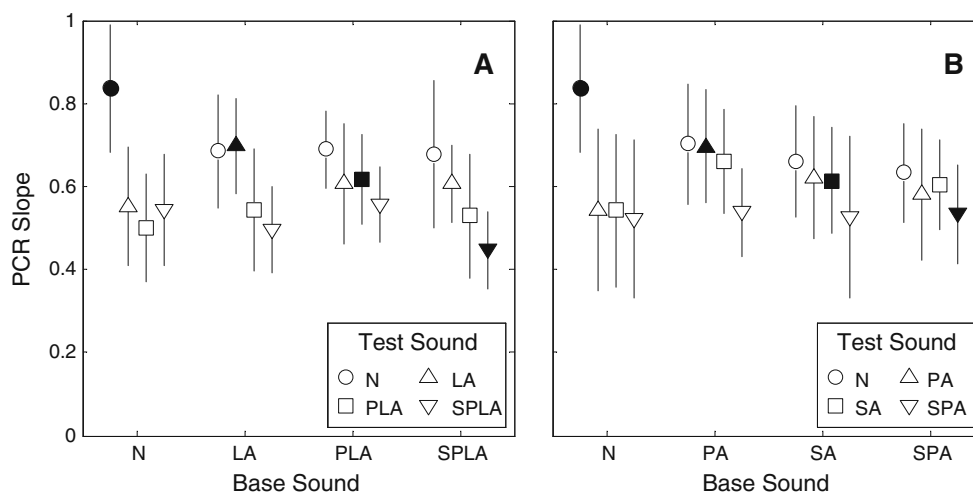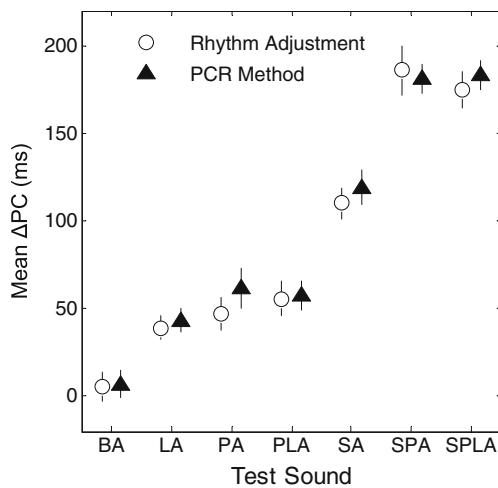
Fig. 6 Between-participants ΔPC estimates from both methods compared. Symbols indicate the mean ΔPC relative to the reference noise, N, in milliseconds. Error bars indicate the 95% confidence interval of the mean

across all syllable–noise ΔPCs separately for each method.[2] The results indicated that this averaged standard deviation for rhythm adjustment (10.4 ms) was slightly less than that for the PCR method (11.4 ms), but the difference was not significant, $t(12) = -0.77$, $p = .46$.

Since the accuracy of estimates obtained by both methods was rather similar, efficiency depended primarily on the time requirements for each participant. For both methods, the time requirement for N trials of M test sounds (all paired with the same references sound) can be estimated simply as $N \times M \times T$, where T is the trial duration (with some allowance for breaks between trials). In this study, rhythm adjustment used 12 trials per pair and took 48.2 s per trial on average, whereas for the PCR method, the values were 10 trials per pair and 43.0 s per trial. Thus, rhythm adjustment was somewhat less efficient as executed.

Furthermore, we suspected that the nature of the regression used to estimate ΔPC values could permit shorter PCR method trials (by using fewer EOS levels). The PCR method data were reanalyzed using just 6 of the original 11 EOS levels—namely, –50, –30, –10, 10, 30, and 50 ms. The between-participants ΔPC estimates differed from the originals by less than 3 ms in all cases, although the averaged between-participants standard deviations increased from 11.4 to 13.2 ms. Using these EOS levels, each PCR trial could be completed in just 23.5 s, providing

a useful 45% reduction in participant time required for each sound pair.

General discussion

P-center measurement

In this research, a new method for P-center measurement, the PCR method, was introduced and evaluated in comparison with the commonly employed rhythm adjustment method. To that end, the study had several objectives: to determine whether or not both methods produce consistent ΔPC estimates, to confirm or disconfirm the P-center context independence hypothesis, and to evaluate the accuracy and efficiency of each method.

Consistency of estimates Our results showed that the PCR method and rhythm adjustment method are consistent. This finding is important for a number of reasons. First, to our knowledge, no previous study has explicitly compared P-center measurement methods. Instead, a variety of measurement methods have been used, with no evidence that they all measure the same percept. In fact, specific problems reported with other methods, such as the existence of multimodal P-center distributions (Gordon, 1987; Wright, 2008) and underestimated interval durations (Fox & Lehiste, 1987), would suggest that it is dangerous to simply assume that all measurement methods are equally valid and comparable.

Second, although Vos et al. (1995) concluded that tap asynchrony varies with the P-center and Janker (1996) applied this to P-center measurement, the validity of this approach has apparently not been tested by comparison with independently measured P-centers. Our results provide indirect support for the notion that the average difference in tap asynchrony between consecutive events varies reliably with their P-centers, since this is an alternative approach to calculating the PCR. Furthermore, because a tap asynchrony is simply a biased EPC, the consistency of PCR results with those of rhythm adjustment suggests that the average bias is stable across consecutive events, even when those events differ. Whether this bias is stable across an entire sequence (trial) or between trials is an empirical question with important consequences yet to be answered: Without a stable bias, the tap asynchrony method can never yield reliable P-center estimates.

Finally, we observe that the consistency of ΔPC estimates obtained using different measurement methods, in different laboratories, and using different participants supports the nature of the P-center as a reliable and universal percept and corroborates previous research (e.g., Marcus, 1981), which indicates that P-centers do not depend on individuals or groups.

---

[2] The data provided no evidence that participants reliably perceived different P-centers; on the contrary, there was good agreement in mean ΔPC estimates between participants. Therefore, we assume that between-participants variability in ΔPC estimates is due to task-independent factors common to both methods (such as human limits for anisochrony detection) and task-dependent factors that may make one method's estimates more accurate than the other's. It is these relative differences in accuracy, if any, that are of concern here.

Context independence No evidence of P-center context dependence was found for either the rhythm adjustment or the PCR method. Therefore, these results support previous P-center context independence findings for rhythm adjustment (Eling et al., 1980; Marcus, 1981) and extend those findings to the PCR method. The most important implications of this context independence for P-center measurement are that indirect ΔPC estimates may be calculated from averaged direct measures and that, as a consequence, ΔPC estimates from different experiments or studies using these methods may be compared, provided the stimulus sets share at least one sound in common.

Accuracy and efficiency The accuracy of measurement methods is always of concern, and our results show that there is no significant difference between the methods in this regard. The PCR method in this study used highly skilled musicians, however, and while we would expect no significant difference in ΔPC estimates, the accuracy of the method may be reduced if musically less skilled participants are employed; higher within-participants variability could be expected, and more participants may be required to achieve an equivalent between-participants standard error.

Our consideration of method efficiency results from previous experience measuring P-centers using rhythm adjustment. Participants found the method somewhat long and fatiguing, and this, in turn, limited the number of P-centers that could be measured in a session or study. All else being equal, both participants and researchers benefit if the most time-efficient method is selected. Our study showed that the PCR method was more time efficient than rhythm adjustment. Although the PCR method does require some additional time not accounted for in the participant time estimates (because initial ΔPC estimates are required to construct the presentation sequences), this requirement is not onerous; it is quite feasible for initial ΔPC estimates to be obtained from a pilot experiment using just 1 individual (as was the case in this study) or a small number of participants.

PCR function slope

The phase correction response is characterized by the parameter α, estimated as the slope of the PCR function, which reflects the weight given to the timing of external pacing events, relative to internally planned tap events; it is an index of the strength of sensorimotor coupling. It is known, for example, that α is smaller in synchronization with visual stimuli (Repp & Penel, 2002) and in synchronization of continuous drawing movements with a metronome (Repp & Steinman, 2010), presumably due to weaker sensorimotor coupling in each case. In this study, α can be interpreted as indicating how confidently a participant perceives the P-centers of pacing events. If the P-center is difficult to locate accurately, a participant cannot assign it much confidence and should, instead, rely more on continuation of their established internal timing. On the other hand, if the pacing P-center can be located accurately, responding quickly to any perturbations in the pacing sequence is a better strategy for staying synchronized.

The results of Experiment 2 show a clear effect of sound on α for homogeneous sequences. It is largest for the N sound; participants adjust their taps most confidently and rapidly to phase perturbations of this sound. In contrast, α is smallest for SPLA, the most complex syllable with one of the latest EPC estimates. Participants appear to adjust more tentatively and more slowly when this sound is perturbed from perceptual isochrony. To explain these results, we consider the precision of the P-center percept in some more detail.

Some sounds have subjectively well-defined and clear P-centers. Short sounds, percussive sounds, and the N sound in this study fall into this category. The P-centers of sounds with longer and more gradual or more complex onsets seem to have P-centers that are somewhat more ambiguous or, at least, more difficult to detect accurately. This phenomenon is generally not reported in the literature, with the possible exception of Rasch (1979), who suggested that "shorter and sharper rises of notes make better synchronization both necessary and possible" (p. 128). In particular, the phenomenon does not appear to have been formally identified to date, nor have there been any detailed studies examining it. As a consequence, we introduce the term *P-center clarity* to describe the subjective precision of a P-center.

Although we did not formally investigate P-center clarity, it seems that α may be directly related to the perceived clarity of the P-center for homogeneous sequences. For mixed sequences, however, the situation is more complex. The perturbed test sound had a significant effect on the PCR function slope, whereas the base sound did not appear to have an effect. The direction of the effect was generally the same as that for homogeneous sequences, suggesting that the PCR slope of mixed sequences was related to the perceived clarity of the test sound's P-center. Mean slopes for mixed sequences appeared to be smaller than those for homogeneous sequences for each test sound, but this effect did not reach significance. Nevertheless, a reduction in slope, which can be interpreted as reduced confidence in localizing the test sound P-center, suggests that a change of sounds results in a perceptual penalty. A possible explanation for the penalty is the increased cognitive load when perceptual expectations, spectral and temporal, created by the repeated base sound are suddenly violated by the inserted test sound. Although the insertion of oddball stimuli into a sequence of predictable stimuli has been shown to affect perceived stimulus duration (Tse,

Intriligator, Rivest, & Cavanagh, 2004) and might affect perceived interval duration, the data provided no evidence of this. (A change in perceived interval duration should cause a change in PCR function intercept, rather than PCR function slope.)

These results raise an interesting question: Is α constant throughout a sequence, or does it adapt to changes? Before the first EOS is encountered, there is no difference between a homogeneous and a mixed sequence, so it would be natural to expect that the initial value of α in a sequence would be identical for both sequence types. After the first EOS, it is possible that there is a step change in α for mixed sequences that remains approximately constant thereafter. An alternative hypothesis is that α adapts gradually but continuously throughout the sequence. Yet another alternative is that α depends only on the identity of the most recent pacing sound and, therefore, may change after each sound. The experiments in this study cannot easily distinguish between these hypotheses, but the possibility that the strength of sensorimotor coupling is continuously variable warrants further investigation.

Conclusions

We have shown the PCR method to be a useful new method for measuring P-centers (specifically, ΔPCs). It is essentially interchangeable with the more commonly used rhythm adjustment method in terms of both the mean and variability of ΔPC estimates that result. The PCR method's compelling advantage is that it does not require conscious decision making by participants, an advantage when some of the P-centers to be measured are relatively unclear. The PCR method is also more time efficient than rhythm adjustment, which is a definite advantage when trying to measure many P-centers. In the context of ΔPC measurement, the main advantage of the rhythm adjustment method is its simplicity, in terms of both apparatus and subsequent data analysis.

Our data do not provide any evidence of P-center context dependence for either the rhythm adjustment or the PCR method. This finding is important because the assumption of P-center context independence is the foundation on which ΔPC comparison within and between experiments, using any of the methods in this study, relies.

We have introduced the term P-center clarity to describe the subjective precision with which an event's P-center is perceived. Although not specifically manipulated in this study, we note that clarity seems closely related to both the abruptness of the event onset and the lateness of the P-center, relative to the event's onset. When sounds with relatively unclear P-centers are approximately isochronously timed, the dispersion of acceptable points of subjective

isochrony might be expected to be wider than for sounds with clear P-centers. However, our data appear to exhibit just one potential effect of P-center clarity: The slope of the PCR function gets shallower for sounds with more complex onsets and less clear P-centers.

A final intriguing question raised by this study is whether the strength of sensorimotor coupling (measured by α) depends on the nature of the sequence and, furthermore, whether it may change throughout the sequence.

References

Aschersleben, G. (2002). Temporal control of movements in sensorimotor synchronization. *Brain and Cognition, 48*, 66–79.

Aschersleben, G., Gehrke, J., & Prinz, W. (2004). A psychophysical approach to action timing. In C. Kaernbach, E. Schröger, & H. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition* (pp. 117–136). Hillsdale: Erlbaum.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*, 379–384.

Benguerel, A.-P., & D'Arcy, J. (1986). Time-warping and the perception of rhythm in speech. *Journal of Phonetics, 14*, 231–246.

Bregman, A. S. (1999). *Auditory scene analysis: The perceptual organization of sound* (2nd ed.). Cambridge: MIT Press.

Cooper, A. M., Whalen, D. H., & Fowler, C. A. (1986). P-centers are unaffected by phonetic categorization. *Perception & Psychophysics, 39*, 187–196.

Eling, P. A., Marshall, J. C., & van Galen, G. P. (1980). Perceptual centres for Dutch digits. *Acta Psychologica, 46*, 95–102.

Fowler, C. A. (1979). 'Perceptual centers' in speech production and perception. *Perception & Psychophysics, 25*, 375–388.

Fox, R. A., & Lehiste, I. (1987). The effect of vowel quality variations on the stress-beat location. *Journal of Phonetics, 15*, 1–13.

Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology, 35*, 1–36.

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception, 26*, 427–438.

Gordon, J. W. (1987). The perceptual attack time of musical tones. *Journal of the Acoustical Society of America, 82*, 88–104.

Harsin, C. A. (1997). Perceptual-center modeling is affected by including acoustic rate-of-change modulations. *Perception & Psychophysics, 59*, 243–251.

Janker, P. M. (1996). Evidence for the P-center syllable-nucleus-onset correspondence hypothesis. *ZAS Papers in Linguistics (ZASPIL), 7*, 94–124.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences, 12*, 72–79.

Marcus, S. M. (1981). Acoustic determinants of perceptual-center (P-center) location. *Perception & Psychophysics, 30*, 247–256.

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review, 83*, 405–408.

Murata, K., Nakadai, K., Yoshii, K., Takeda, R., Torii, T., Okuno, H. G., … Tsujino, H. (2008, September). *A robot uses its own microphone to synchronize its steps to musical beats while scatting and singing.* Paper presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008), Nice, France.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8*, 434–447.

Perez, P. E. (1997). *Consonant duration and stress effects on the p-centers of English disyllables* (Unpublished doctoral dissertation). University of Arizona, Tucson

Pompino-Marschall, B. (1989). On the psychoacoustic nature of the P-center phenomenon. *Journal of Phonetics, 17*, 175–192.

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

Rapp-Holmgren, K. (1971). A study of syllable timing. *STL-QPSR, 12*, 14–19.

Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acustica, 43*, 121–131.

Repp, B. H. (1995). Detectability of duration and intensity increments in melody tones: A partial connection between music perception and performance. *Perception & Psychophysics, 57*, 1217–1232.

Repp, B. H. (2001). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 600–621.

Repp, B. H. (2002). Automaticity and voluntary control of phase correction following event onset shifts in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 410–430.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review, 12*, 969–992.

Repp, B. H., & Dogget, R. (2007). Tapping to a very slow beat: A comparison of musicians and non-musicians. *Music Perception, 24*, 367–376.

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simulta-neous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 1085–1099.

Repp, B. H., & Steinman, S. R. (2010). Simultaneous event-based and emergent timing: Synchronization, continuation, and phase correction. *Journal of Motor Behavior, 42*, 111–126.

Schroeder, M. (1970). Synthesis of low-peak factor signals and binary sequences with low autocorrelation. *IEEE Transactions on Information Theory, 16*, 85–89.

Schulze, H.-H., & Vorberg, D. (2002). Linear phase correction models for synchronization: Parameter identification and estimation of parameters. *Brain and Cognition, 48*, 80–97.

Scott, S. K. (1993). *P-centres in speech: An acoustic analysis* (Unpublished doctoral thesis). University College London, London.

Scott, S. K. (1998). The point of P-centres. *Psychological Research, 61*, 4–11.

Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics, 66*, 1171–1189.

Tuller, B., & Fowler, C. A. (1980). Some articulatory correlates of perceptual isochrony. *Perception & Psychophysics, 27*, 277–283.

Villing, R., Ward, T., & Timoney, J. (2007, July). *A review of P-centre models.* Paper presented at the Rhythm Perception and and Production Workshop, Kippure Estate, County Wicklow, Ireland.

Vos, P. G., Mates, J., & van Kruysbergen, N. W. (1995). The perceptual centre of a stimulus as the cue for synchronization to a metronome: Evidence from asynchronies. *Quarterly Journal of Experimental Psychology, 48A*, 1024–1040.

Wright, M. (2008). *The shape of an instant: Measuring and modeling perceptual attack time with probability density functions* (Unpublished doctoral dissertation). Stanford University, Palo Alto, CA.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127*, 3–21.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin, 133*, 273–293.