Available online at www.sciencedirect.com

SCIENCE DIRECT®

Cognitive Psychology 53 (2006) 59–96

Cognitive Psychology

www.elsevier.com/locate/cogpsych

Effects of auditory pattern structure on anticipatory and reactive attending ☆

Mari Riess Jones *, Heather Moynihan Johnston, Jennifer Puente

Department of Psychology, The Ohio State University, Columbus, OH 43210, USA

Accepted 26 January 2006
Available online 24 March 2006

Abstract

In three experiments, participants listened for a target's pitch change within recurrent nine-tone patterns having largely isochronous rhythms. Patterns differed in pitch structure of initial (context) and final (target distance) pattern segments. Also varied were: probe timing (Experiments 2 and 3) and instructions about probe timing (Experiments 2 and 3). In all experiments, identification of a recurrent target was poorer in patterns with wider context pitch intervals (in semitones) than in others. Effects of probe timing also occurred, with better performance for temporally expected than unexpected probes. However, when listeners were explicitly told to focus upon a target's pitch and not its timing (Experiment 3), they performed selectively better in patterns with smaller target/probe pitch distances, especially for rhythmically expected probes. Five theoretical approaches to the respective roles of pitch and/or time structure were assessed. Although no single approach accounted for all results, a modification of one theory (a Pitch/Time Entrainment model) provided a reasonable description of findings.

Keywords: Anticipatory attending; Reactive attending; Auditory patterns

---

☆ This research was sponsored, in part, by a National Science Foundation Grant, BCS-9809446, awarded to the second author. The authors are indebted to Michelle Huffman and Julia Wood for assistance with data collection as well as to Robert Ellis, Sarah Judkins, Iring Koch, and Noah MacKenzie for helpful comments on earlier drafts.

* Corresponding author.
E-mail address: jones.80@osu.edu (M.R. Jones).

0010-0285/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogpsych.2006.01.003

1. Introduction

In listening to music and related sound sequences people must attend on a moment-to-moment basis to identify aspects of the unfolding pattern. The present research considers how such attending activity is guided by auditory pattern structure. Of special interest is the role of a pattern's pitch and time structure in attentional monitoring.

In three experiments, we use a sequence monitoring task to assess selective attending to the pitch of a tone within a repeating pattern. Listeners identify the pitch of a recurrent target tone embedded within novel sequences, distinguished by different pitch/time properties. Originally, selective listening studies engaged constructs such as information channels to explain dichotic listening performance (Cherry, 1957; Moray, 1969a; Moray, Fitter, Ostry, Favreau, & Nagy, 1976), but recently other constructs, such as attentional focus (e.g., in frequency), have motivated research on selective listening. Thus, focusing attention to a certain tone frequency (or intensity) is more efficient for expected than for unexpected items where an expectancy is commonly established by relatively high prior probabilities or by an immediately preceding priming cue (Greenberg & Larkin, 1968; Mondor & Bregman, 1994; Moray et al., 1976; Scharf, 1998; Scharf, Quigley, Aoki, Peachey, & Reeves, 1987; Ward & Mori, 1996; Wearden & Bray, 2001).

In the present research, we are concerned with attention and expectancies in the context of sequence monitoring. In these situations, the manipulation of expectancies via probabilities or single cues becomes challenging; a more fruitful approach derives from the long tradition of research on serial pattern structure which seeks determinants of expectancies in relationships, including pitch and time relations, among sequential serialelements (Garner & Gottwald, 1968; Hoffman & Koch, 1998; Jones, 1974, 1981; Restle, 1970; Restle & Brown, 1970). More recently, using visual events, it has been proposed that serial structure induces implicit learning and entails the integration of sequence time relationships with serially unfolding spatial relationships (Koch & Hoffman, 2000; Restle, 1972; Shin & Ivry, 2000). To this end, serial reaction time tasks (SRT), which involve self-paced serialized motor responses (Nissen & Bullemer, 1987), have often been employed to assess the acquisition of serial relations over trials. The present research is also concerned with the role of sequence timing, but in auditory patterns that unfold in pitch space. Moreover, because we are interested in the way attending is paced by pattern structure, we eschew self-paced SRT procedures that involve correlated motor responses. Our goal is to assess the degree to which pattern structure itself, including time structure, paces attending to a succession of pattern elements (e.g., tones). Indeed, it is possible that effectively timed attending, based on event structure, paves the way for emergence of coordinated motor responses.

Recent findings in sequence monitoring tasks implicate a role for event time structure (rate, rhythm) in pacing attending and generating temporal expectancies (Jones, Moynihan, MacKenzie, & Puente, 2002; Klein & Jones, 1996; Large & Jones, 1999). With isochronous sequences, listeners were better at judging a rhythmically expected than a rhythmically unexpected ending tone, producing a symmetrical accuracy profile as a function of end tone timing, as shown in Fig. 1. This inverted U profile has been termed an *expectancy profile* (Barnes & Jones, 2000). A major goal of the present research is to identify various determinants of expectancy profiles in light of current theories about the impact of rhythm and pitch structure on expectancies and attention. In particular, we consider whether or not pitch identifications in a sequence monitoring task yield expectancy profiles that differ as a function of a pattern's pitch and/or time structure (Experiments 2

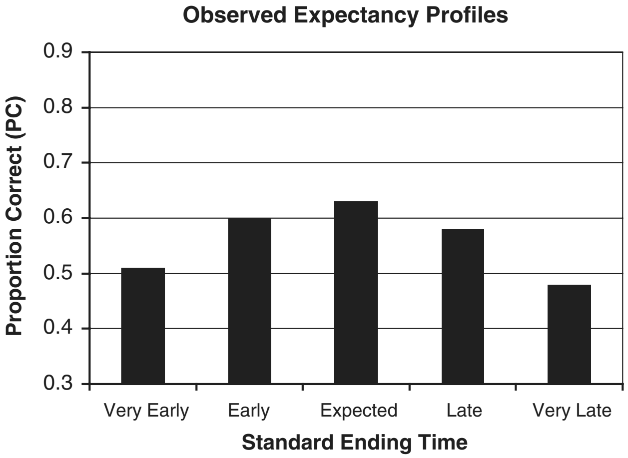

Fig. 1. A typical expectancy profile indicating maximum accuracy (proportion correct, PC) for rhythmically expected ending tones of a sequence. These data are taken from an experiment involving time judgments (Barnes & Jones, 2000).

and 3). Moreover, we also ask if such profiles are susceptible to modulation by voluntary factors instilled by instructions (Experiment 3).

Expectancy profiles, first observed by Large and Jones (1999), have been interpreted as support for the entrainment of attending. Entrainment is a ubiquitous biological activity responsible for a variety of synchronous behaviors in organisms that connect them to various environmental regularities (light/dark cycles, meals, tidal rhythms, and so forth) (Winfree, 2000). With respect to attending, one hypothesis is that attending is associated with an internal neural activity with oscillatory tendencies and a corresponding potential for entrainment. At time scales of speech and music, such an activity reflects a primitive attunement process that may support implicit sequence learning in that it specifies how listeners tacitly come to an internal synchrony with aspects of an external sequence, i.e., the driving rhythm (Large & Jones, 1999). However, in the strict entrainment view of attending, the construct of a driving sequence rhythm is narrowly interpreted: it comprises a series of stimulus time intervals, each marked by a change in intensity conveyed by a tone's onset. Corresponding support for the development of temporal expectancies derives from tasks involving judgments of time intervals using monotone auditory patterns (McAuley & Jones, 2003).

1.1. A Strict Entrainment view: Determinants of Temporal expectancies

In the present research, we compare a strict view of a driving rhythm (given by Large & Jones, 1999) with a broader view that assumes that an auditory driving rhythm is described in terms of changes in both intensity *and* tone frequency. To develop this, we begin with the concept of a *temporal attentional focus* as a concentration of attentional energy (pulse) in time that is carried by an entraining oscillator (Large & Jones, 1999). Attentional energy is not a mentalistic construct; rather, it refers to a potentially measurable temporal concentration of neural activity associated with an internal oscillator (Large & Jones, 1999; Snyder & Large, 2005). We distinguish our description of a focus *in time* from attentional foci along space-like dimensions that have been proposed by others for visual space (Jonides & Yantis, 1988; Yantis & Jonides, 1984) or in pitch space (i.e., along the acoustic frequency continuum, e.g., Scharf, 1998). Large and Jones described this focus as an *expected region in time* (cf. Wright & Fitzgerald, 2004). Following entrainment theory, the location of a temporal focus depends on the rate and coherence of a driving rhythm. This is because listeners come to synchronize their attentional focus with the temporal onsets of successive tones, as illustrated in Fig. 2. The width, as well as location, of a recurrent attending pulse, i.e., the focus, changes as function of rhythmic coherence. Thus, in monitoring rhythmically irregular sequences a listener's attentional focus is proposed to widen (dotted line in Fig. 2, inset). As shown, a wider attentional focus corresponds to a flatter expectancy profile; in turn, this leads to predictions of lower accuracy in judging temporally expected tones within the focus. By contrast, in sequences with rhythmically regular timing, a listener's temporal focus narrows, leading to more precise attentional targeting in time and greater accuracy with rhythmically expected tones (solid line in Fig. 2, inset). Because entrainment involves inherently oscillatory attending activities, when it is efficient, it naturally incorporates an internal anticipatory process, meaning that the relative timing of an organism is in synchrony with a stimulus time pattern. In other words, a *temporal expectancy reflects a timed, anticipatory, extrapolation of an attentional focus that is driven by event rhythm.*

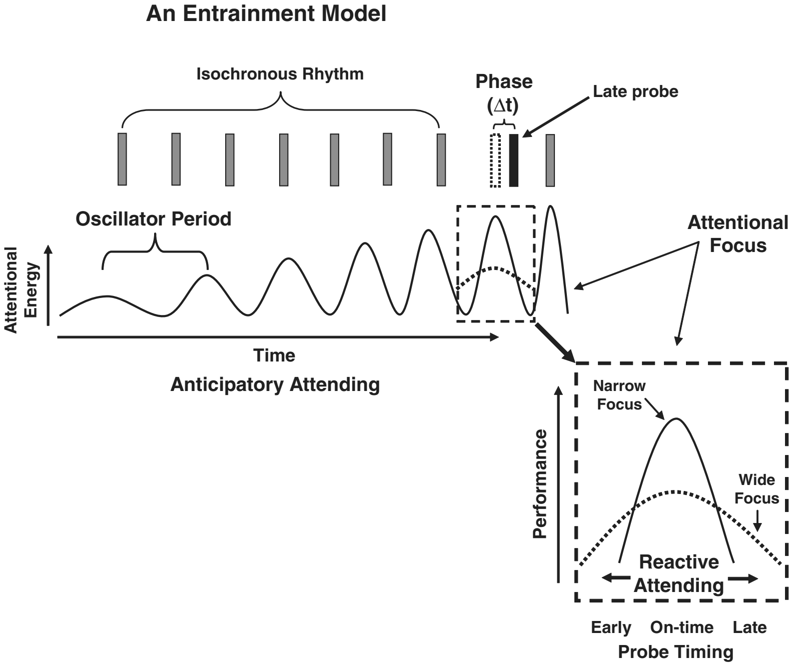

Fig. 2. Hypothetical entrainment activities of an internal oscillator with respect to a tone (gray bars) sequence containing one ill-timed (late) tone (black bar) . Both the period and the phase of an internal oscillator adjusts to synchronize with successive tone onsets. Inset shows two attentional focus (pulse) widths: narrow (solid lines) and wide (dashed lines). Anticipatory attending is periodically based on event rhythm and precedes a probe onset; attentional energy is greatest for on-time probes. Reactive attending follows an expectancy violation and is based on a translatory time shift shown for early and late probes.

Efficient entrainment in rhythmically coherent patterns promotes precisely timed expectancies. Moreover, when a coherent rhythm contains a single violation, attentional synchrony is restored (within limits) via adaptive responding following this perturbation (Jones, 2004; Large & Jones, 1999). Intuitively, an ill-timed tone elicits in a listener a sort-of "double-take" that reflects an underlying corrective reaction of the entraining oscillation. Temporally unexpected elements automatically result in time shifts of an internal oscillation that have been formalized as phase corrections [Large and Jones, 1999; see (Eq. 3)]. Because a phase correction rests on a reactive shift of the attentional focus in time, it has been described as *reactive attending* (Jones, 2004). As with the underlying synchrony of anticipatory attending, the underlying process of phase correction is not mentalistic; rather it is also subject to the physical/biological constraints of entrainment. However, we distinguish *reactive attending* from *anticipatory attending* because anticipatory attending is significantly influenced by the global time structure of an event whereas reactive attending is a response to a local, often temporally deviant, aspect of pattern structure.

The terms *anticipatory* and *reactive* are simply descriptive labels for attending processes that precede and follow in time a given to-be-attended, but potentially deviant, element within an unfolding serial context.[1] Although, the distinction is evocative of that between voluntary and involuntary attending processes in visual attention, strict parallels here are risky (Jones, 2001). Converging support for this dichotomy in audition comes from ERP findings using sound sequences. Averaged ERPs to deviant items show, e.g., enhanced negativity *following* that item, whereas anticipatory ERP activity has been shown to *precede* expected items (Besson & Faita, 1995; Besson & Macar, 1987; Goschke, 1998; Snyder & Large, 2004, 2005). Fig. 2 (inset) further illustrates this distinction. Anticipatory attending is shown to realize a temporal expectancy, set in motion by the prevailing rhythm, that is directed toward a future (expected) point in time. By contrast, reactive attending involves a re-orientation toward an unexpectedly timed item that follows from a violation of this expectancy. In short, reactive attending is contingent upon temporal expectancies, i.e., upon anticipatory attending, because it is sparked by an expectancy violation.[2]

Theoretically, expectancy profiles are predicted in tasks that involve *both* anticipatory attending (to expected points in time) and reactive attending (to unexpected points in time). A sequence element that occurs at an expected time occurs *within* an anticipated focus of attending and this insures relatively good performance. However, if an element occurs *outside* this temporal focus region (i.e., is very early or very late), then reactive attending is restricted and poorer performance occurs. That is, reactive attending is limited by the temporal width of an attentional focus. This has implications for predicted expectancy profiles. Sequences that are predicted to promote a narrow focus should yield a sharp expectancy profile, characterized by good performance with on-time tones and relatively poor performance with ill-timed ones. Theoretically, a sharp expectancy profile

---

[1] See Jones (2004) for a discussion of differences in reactive attending when expectancy violations occur for early and late probes. In some cases we observe asymmetries in performance for these two different types of violations, but these asymmetries may be more likely to occur when the probe is the final sequence tone rather than the penultimate one.

[2] Caution is warranted in drawing strong conclusion from ERP findings at this point. Although this methodology is promising, often stimuli employed are very complex and rhythmic deviations are not comparable to those used here. Moreover, Bresson and colleagues conclude that although evidence for timed expectancies exist in ERP data, they attribute these findings in terms of decisional rather than attention processes per se.

follows from a narrow temporal focus and reflects: (a) Efficient anticipatory attending (to rhythmically expected elements) and (b) limited reactive attending (to rhythmically unexpected elements). Conversely, sequences that promote a wide attentional focus should yield a flat expectancy profile due to relatively poor performance with on-time tones along with moderately better performance with ill-timed ones. Theoretically, a wide, diffuse, focus in time results in relatively flat expectancy profiles, reflecting: (a) Less efficient anticipatory attending and (b) greater access to certain ill-timed tones.

The Large and Jones (1999) model raises two issues that are addressed in this research. The first issue concerns the interpretation of a driving rhythm. Attentional synchrony was originally described as depending strictly on the global coherence of a driving rhythm as this is conveyed by a series of time intervals and intensity markers. Determinants of temporal expectancies are rate and rhythm which in turn affect the width of an attentional focus, hence the shape of an expectancy profile. This emphasis upon rate (tempo) and rhythm to the exclusion of pitch structure is interesting because expectancy profiles have recently been found in certain pitch judgment tasks as well as time judgment ones. Jones et al. (2002) found that when random pitch sequences were presented in an isochronous rhythm, listeners were better in identifying the pitch of rhythmically on-time probes than in identifying ill-timed ones. This is interesting, because the strict entrainment view of Large and Jones (1999) describes *temporal*, not *pitch*, expectancies; that is, neither anticipatory nor reactive attending is responsive to pitch relationships within a driving rhythm. It is possible that attending itself is strictly temporal, hence attending at the “right time” provides general performance benefits. Nevertheless, the Large and Jones model does not address a related possibility, namely that the interpretation of a driving rhythm is too narrow to explain performance in tasks where pitch structure and pitch judgments are involved. The present research addresses this issue by manipulating properties of the driving rhythm. Using a pitch judgment task, we vary not only the timing but also the pitch structure of sequences that listeners must monitor. According to the Large and Jones model only timing variations should matter.

The second issue raised by an entrainment account concerns automaticity of pattern-driven attending. If attending is automatically guided by pattern time structure, as implied by Large and Jones (1999), then instructions should have little influence on sequence monitoring. Reactive attending, in particular, has been considered an automatic (stimulus-driven) response to unexpectedly timed tones. Along with manipulations of pitch structure we manipulate instructions about timing to test these assumptions.

In sum, a strict entrainment model offers two null predictions that provide a common theoretical backdrop for the three experiments we report. These involve the respective roles of pitch structure and instructions in sequence monitoring. To preview, we juxtapose these null hypotheses with alternative proposals provided by several contrasting views of expectancy and/or attention.

1.2. Pattern Structure and Task

To motivate a forthcoming discussion of alternative theories, we first outline important aspects of our patterns that relate to different theoretical predictions about pattern timing and pitch structure.

With regard to timing, all patterns in this research are based on a coherent, i.e., isochronous, rhythmic frame. They comprise nine tones with the eighth tone designated the target tone, as shown in Fig. 3A. The time intervals are marked by uniform intensity changes associated with onsets of tones (that vary in frequency); these time intervals are identical with one exception: one tone is sometimes temporally displaced in order to assess effects of an expectancy violation. Each nine-tone pattern cycles three times. If rhythmic isochrony is violated, the time change occurs in the third cycle with the target tone (now termed probe) arriving unexpectedly in time (i.e., early or late). In the first experiment the rhythm is isochronous in all cycles, whereas in later experiments the relative timing of the *probe tone*

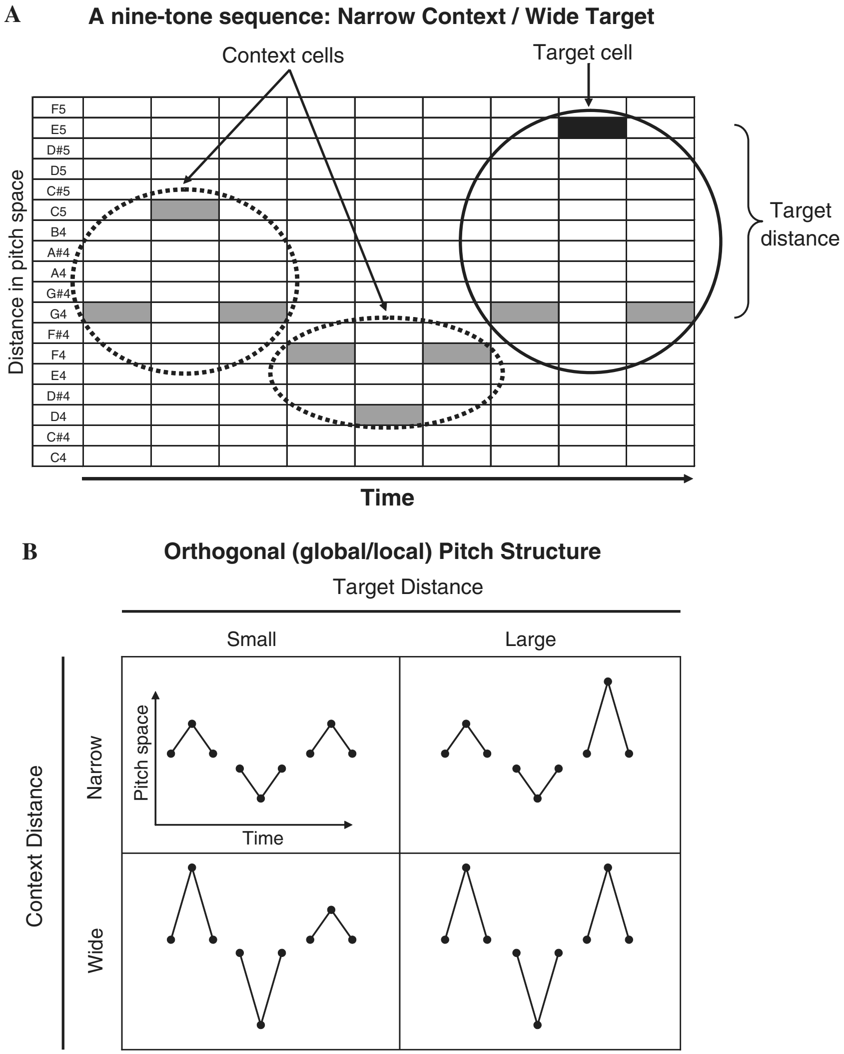

Fig. 3. Stimulus sequences: (A) an exemplar chromatic tone sequence in which the first two cells (global context) contain Narrow pitch intervals and the third cell (local target distance) contains Large intervals. (B) Outline of four different pitch structure conditions (Experiments 1–3) created by orthogonally crossing levels of context with levels of target distance; ordinate represents distance in pitch space and the abscissa represents change in time.

varies, arriving (equally often) early, on-time, or late. Unlike earlier research, in these experiments the probe is the penultimate tone, not the final sequence tone.

The probe timing variable is critical to testing certain hypotheses about sequence monitoring and expectancy profiles. An observed expectancy profile depends on *both* the contextual timing within a sequence and on the temporal violations of this driving rhythm. Thus, if a probe occurs early or late, then attending should be 'mistimed,' leading to an expectancy profile of the same form reported in related paradigms (i.e., where timing of a ending tone varied, cf. Fig. 1). A strict version of the entrainment model (i.e., Large & Jones, 1999) implies that people "use" the intensity/time pattern of a rhythmic sequence to pace attending; thus, they should perform best when a probe tone is rhythmically expected and worst when it is not, regardless of sequence pitch structure.

We assess the null prediction of Large and Jones by varying pitch interval widths in these sequences. One aim is to examine the respective effects of global and local pitch structure on expectancy profiles. Examples of the pitch manipulations appear in Fig. 3. In practice, global and local pitch structure have often been confounded (e.g., sequences contain either all small pitch intervals or all large ones); here we orthogonally cross global with local pitch structure. Our manipulations focus upon pitch interval size (as opposed to tonal/harmonic properties); constituent tones come from the chromatic musical scale. The nine-tone patterns comprised three cells with each cell involving three tones (as in the exemplar pattern of Fig. 3A). Pitch relationships in the first two cells constitute global pitch structure; these form a serial *context variable*. In the third cell, pitch intervals convey local pitch structure because they surround the target pitch; these form a *target distance variable*. In all experiments, we orthogonally manipulate the size of pitch intervals comprising context (Narrow, Wide) and target distance (Small, Large) cells, as shown in the four panels of Fig. 3B. This resulted in four different pitch structure, i.e., context/target distance, conditions: Narrow/Small, Narrow/Large, Wide/Small, and Wide/Large. These pitch structure variables distinguish pattern types in all experiments.

Our main experimental design requires selective attending to certain tones within patterns from these four pitch conditions. In each pattern, listeners must track the pitch of a recurring target to determine how this pitch changes when it becomes the probe in the third cycle. In some experiments (Experiment 1), the local time of the probe does not vary whereas in others it does (Experiments 2 and 3). Together, these pitch and time manipulations permit tests of the Large and Jones (1999) entrainment hypothesis. Specifically, we ask: "what aspects of pattern structure determine whether or not a target/probe tone falls outside a listener's temporal focus of attention?" According to a strict version of entrainment, these aspects relate only to a pattern's intensity/time changes. Therefore, in all four pitch pattern conditions performance should be poorer for ill-timed than for on-time probes and identical expectancy profiles should appear in all four pitch structure conditions (in Experiments 2 and 3). Alternatively, other theories incorporate a role for pattern pitch structure. We consider these next.

1.3. Four Pitch Structure Theories: Determinants of Pitch Expectancies

We contrast four alternative approaches with the strict entrainment view of Large and Jones (1999). Unlike the Large and Jones theory, all of these theories postulate a role for *pitch* expectancies in the present task. However, depending on the theory, pitch expectancies reflect an orientation to the pitch dimension or they emerge from attunement to pitch relationships within a pattern. These approaches also differ in respective emphasis on time and rhythm. They fall into two categories, namely theories that feature explicit roles for *both* pitch and time structure (Pitch/Time Entrainment, Pitch Space accounts) and those which primarily emphasize the role of pitch structure on pitch expectancies (Implication–Realization, Scene Analysis theories). The former offer predictions about manipulations of both pitch structure and probe timing, whereas the latter offer predictions mainly about the pitch structure variables.

1.3.1. Theories of pitch structure and timing

The two models in this category share a common concern with an expectancy-generated attentional focus that involves pitch. However, their interpretation of an attentional focus differs.

A *Pitch/Time Entrainment view* shares with the Large and Jones model an emphasis upon pattern rhythm as a determinant of entrainment. But it broadens the definition of an effective driving rhythm by including frequency change as well as changes in intensity and/or time. In this view, a pitch/time pattern context induces a *moment-to-moment attending trajectory that is extrapolated from pattern structure, as a dynamic expectancy in both pitch space and time*. Such trajectories function as joint expectancies about 'where' in pitch space and 'when' in time future tones may occur. Thus, expectancies necessarily reflect an attunement to pattern structure, but they are dynamic extensions of this structure as well. Presumably, future tones that occur within an anticipated *pitch–time focus* (neighborhood) will be identified more efficiently than those far from this pitch–time focus (Jones, 1990).

Relevant research suggests that both global and local pitch relationships contribute to rhythmic coherence and to establishing pitch–time expectancies. Listeners were best in identifying an unexpected tone's pitch when it was relatively close in pitch space to its expected location; related constraints were associated with time manipulations (Boltz, 1989; Dowling, Lung, & Herrbold, 1987; Jones, Kidd, & Wetzel, 1981). This research invites enlarging the construct of a driving rhythm to include pitch structure. Accordingly, the Pitch/Time Entrainment theory interprets an attentional focus as a *region in pitch space and time*. This theory leads to two hypotheses that concern, respectively, anticipatory and reactive attending, but which differ from those of Large and Jones (1999).

The first hypothesis assumes that anticipatory attending is influenced by a pattern's invariant relationships, including its initial serial structure. Narrow pitch context intervals are assumed to elicit anticipatory attending with a narrow attentional focus *in time*, whereas patterns with Wide pitch context intervals will elicit less precise anticipations, due to wider and more diffuse attentional focus in pitch–time. In other words, pitch structure is assumed to affect the width of an attentional focus, *including its temporal width.* This leads to the prediction that Narrow context conditions will promote more efficient anticipatory attending (i.e., better performance) to on-time target/probes than Wide context conditions.

The second hypothesis of the Pitch/Time Entrainment theory addresses reactive attending to ill-timed probes. It assumes that local pitch structure, associated with target distance, will contribute to anticipatory attending in that it also affects focus width. That is, this hypothesis is consistent with the first hypothesis in that both hypotheses assume that small pitch intervals invite narrow attentional foci in time. Therefore, with regard to local pitch structure, a Small target distance is more likely to instill a narrow temporal focus than is a Large target distance. In turn, this places constraints on reactive attending; it implies that with Large target distances attentional energy for on-time probes will be dissipated in favor of remote ill-probes that are more likely to fall within a diffuse attentional focus (relative to Small target distances). This is because the attentional focus (in pitch–time) should widen with large pitch intervals that surround a target/probe tone. By contrast, with Small target distances, a narrow attentional focus will restrict reactive attending to early and late probes. Thus, this second hypothesis predicts that performance with ill-timed probes should be *poorer* with Small target distances than with Large ones.

In sum, a Pitch/Time Entrainment view assumes that both anticipatory and reactive attending are influenced by pitch structure with the result that distinctly different expectancy profiles should emerge in different pitch structure conditions. These predictions are summarized in Fig. 4 (solid lines). Sharpest profiles should emerge with monitoring of patterns that instill the narrowest attentional focus, namely those in the Narrow/Small condition, whereas the flattest profiles (due to wide foci) should occur with patterns in the Wide/Large condition.

A *Pitch-Space* model offers a different view of attending. It draws on attentional research in vision to provide an interpretation of an attentional focus with a fixed width that is centered at a specified location on the pitch dimension (i.e., in pitch space, but not in time). The rationale is that pitch is a task-relevant dimension whereas time is not (Egeth & Yantis, 1997; Yantis & Egeth, 1999). In this account, an *expectancy is a top-down voluntary search process, conferred by instructions and task demands, which orients attending to features along the pitch dimension (versus the time dimension)*. Attending is not oriented by unfolding serial pitch relationships within a sequence but by instructions and task demands. Consequently, the Pitch Space model does not incorporate a role for global pitch structure and offers no related predictions about anticipatory attending.

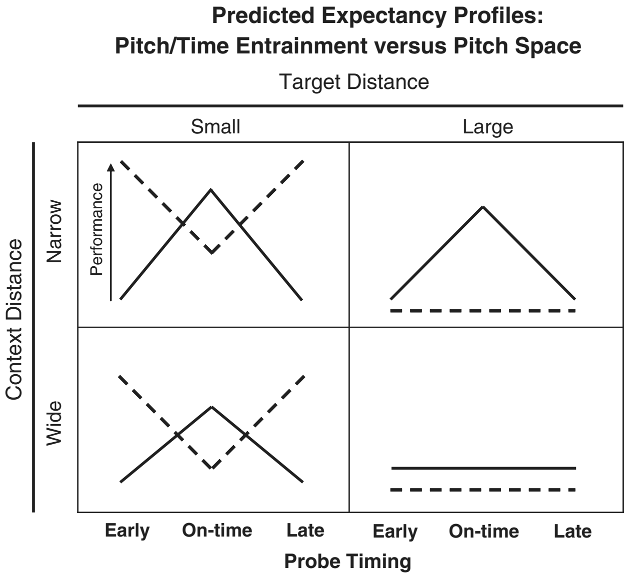

Fig. 4. Predicted expectancy profiles (over probe timing levels) based on a Pitch/Time Entrainment interpretation of attentional focus (solid lines) and a Pitch Space interpretation (dashed lines) for each of the four pitch structure conditions in all experiments. Higher vertical positions within each cell correspond to better performance. Predictions for on-time probes hold for all three experiments; those for early and late probes apply to Experiments 2 and 3.

Nevertheless, the Pitch Space model leads to clear predictions about local pitch and time structure. If a listener orients to pitch as dimension, then a listener's attentional focus will be centered on pitch features of successive tones; consequently, tones proximal in pitch space are likely to be included within a common attentional focus (Mondor & Bregman, 1994). Thus, performance with all probes, even ill-timed ones, should benefit from greater local pitch proximity (Small Target Distance). Moreover, following visual attention interpretations, variations in probe timing should result in stimulus-driven attention shifts. Due to their bottom-up salience, early and late probes may be attention-getting much as abrupt onsets are in visual attention. An ill-timed probe can summon attention, particularly when it is also close in pitch space to the preceding tone. In visual attention, identification of a target is facilitated when it occurs with an abrupt onset, due to attentional capture, namely a stimulus-driven shift of a viewer's focus of attention to the target's spatial location (e.g., Jonides & Yantis, 1988). In auditory arrays, a Pitch Space account predicts that a parallel facilitation will accompany a shift of attention in pitch space to an early or late probe tone with pitch-proximal tones. Thus, a Pitch Space view predicts expectancy profiles in various pitch conditions that differ as a function of target distance and probe timing, as shown in Fig. 4 (dashed lines). Although there is some evidence that performance may be facilitated by the occurrence of certain structurally unexpected targets, this comes from research using a different task involving fast sequences in which probe timing was not varied (Mondor & Terrio, 1998). Thus, specific predictions involving pitch proximity and timing variations within slow sequences have not been previously evaluated.[3] These predictions of the Pitch Space model differ in critical ways from those of the Pitch/Time Entrainment model. In particular, in this view local pitch proximity should not only result relatively poor performance for on-time probes, but it should yield *better* (not worse) performance with ill-timed probes in patterns with Small target distances than with ill-timed probes in patterns with Large target distances.

Two important differences that distinguish the Pitch Space from the Pitch/Time Entrainment theory are shown in Fig. 4. First, with respect to global pitch structure, a Pitch/Time Entrainment theory predicts a benefit from Narrow contexts with on-time probes, due to more efficient anticipatory attending whereas a Pitch Space model predicts no such benefit. Second, with respect to local pitch structure, the two theories predict contrasting interactions of Target Distance with Probe Timing. This is most striking in sequences with Small Target Distance. In particular, the Narrow/Small condition the Pitch/Time Entrainment view (solid lines) predicts a very sharp (inverted U) expectancy profile that contrasts dramatically with the Pitch Space model (dashed lines) which predicts a shallow U profile due to attentional capture by early or late probes.

---

[3] The methodology of Mondor and Terrio (1998) differs from that of present research in many respects. In addition to using much faster pitch sequences (IOIs of 200 ms versus 600 ms in the present research), that were either ascending or descending in pitch, these investigators manipulated the magnitude and direction of a deviant pitch change (i.e., not probe time relationships) in probes that varied in their serial placement within a sequence over trials. Most importantly, although Mondor and Terrio found better performance with structurally unexpected than expected targets, this finding derives from manipulations that confounded the probability of a target's occurrence with structural expectancy (i.e., in a session, structurally unexpected tones were three times more likely to occur than structurally expected ones). In the studies we report here, the three types of pitch targets (same, higher, and lower) and the three types of probe times (early, on-time, and late) are all equally likely to occur.

Finally, predictions about both global and local pitch structure can also be contrasted with null predictions from the Large and Jones (1999) model. This strict entrainment theory predicts identical, sharp, and expectancy profiles for all four pitch structure conditions.

1.3.2. Pitch structure theories

The two remaining theories feature a major role for Gestalt rules. However, they differ in their approach to pitch expectancies.

The *Implication–Realization (I-R) theory* of Narmour proposes that *pitch (melodic) expectancies that depend upon innate (automatic) Gestalt principles applied to pitch intervals* (Narmour, 1990, 1991). This theory identifies pitch structure as a source of "bottom-up," melodic expectancies based on Gestalt rules (Narmour, 1991, 1992; Schellenberg, 1996; Schellenberg, Adachi, Purdy, & McKinnon, 2002). Narmour's original theory also provides guidelines for top-down (learned) as well as bottom-up (innate) expectancies to accommodate a range of sequential effects. Simpler versions of the I-R approach, offered by Schellenberg, engage fewer Gestalt rules, but all I-R approaches feature pitch proximity importantly (Schellenberg, 1996; Schellenberg et al., 2002). Smaller pitch intervals are proposed to imply a strong local expectancy that the immediately following pitch interval (the realized interval) should also be small.

In practice, assessments of I-R theories have addressed local effects of pitch structure in that these often entail pitch manipulations of a final sequence tone. Thus, a melody may end with an implicative (open) and realized (closed) pair of pitch intervals. Here, I-R models predict that listeners rely upon innate Gestalt laws to generate pitch expectancies about the final tone, as these are reflected by high goodness ratings. Typically, regression analyses confirm that local pitch proximity is a strong determinant of these local expectancies (Cuddy & Lunney, 1995; Schellenberg, 1996; Schmuckler, 1989; Thompson, Cuddy, & Plaus, 1997; Unyk & Carlsen, 1987).

In sum, in various I-R models pitch expectancies are largely determined by Gestalt principles, especially the proximity principle. In the present research, we adapt the I-R rationale to address selective listening. We assume that in monitoring chromatic pitch sequences, expectancies based on local and/or global pitch proximity will lead to better overall performance with expected than with unexpected pitches. These models provide clear predictions about local proximity effects (e.g., in Narrow/Small, Wide/Small conditions); generally, I-R theories imply better overall performance for patterns with small rather than large pitch intervals. Finally, however, this approach provides less specificity about the role of temporal expectancies on rhythmically expected versus unexpected probe tones in pitch identifications.

Scene Analysis (Bregman, 1990) shares with I-R theory a featured role for Gestalt principles, especially pitch proximity. In this two-stage model, Gestalt principles represent "hard-wired" relationships functional mainly in an initial perceptual stage that addresses perception of fast auditory sequences; here pitch proximity is an important determinant of automatic groupings among tones with neighboring frequencies. In a second, schema-driven stage, pitch expectancies operate that are not necessarily based on Gestalt principles. Unlike Narmour, for Bregman a *pitch expectancy is defined in terms of selective attending that is determined by a learned, domain-specific, schema.* These schemas effortfully guide attending to selected tones within slow sequences (as in the present research). In this respect, if pitch-proximal tones automatically form a group (pre-attentively) in the first stage, then this grouping can interfere with selective attending to a single (within group) tone in the second, schema-driven stage (Bregman & Rudnicky, 1975; Mondor & Terrio, 1998), Scene Analysis predicts that selective attending is best in pitch patterns where the target tone "stands out" from an established perceptual group. In the present design, this is most likely with patterns of the Narrow/Large condition. As with I-R theories, Scene Analysis does not offer specific predictions about temporal expectancies or the impact of pitch structure on rhythmically generated expectancy profiles.

1.3.2.1. Theoretical summary. Four major theoretical approaches address the role of pitch structure in determining pitch expectancies during selective listening. Although they differ in assumptions and predictions, they share the conclusion that pitch distance, often cast as pitch proximity (local and/or global), plays an important role in the monitoring of relatively slow auditory patterns. A fifth approach to sequence monitoring, the Large and Jones (1999) model, concentrates on temporal, but not pitch, expectancies. Because it provides a spare description of pattern structure, it offers an omnibus null hypothesis regarding the respective roles of global and local pitch structure in a pitch identification task.

1.4. Plan of experiments and hypotheses

Our theoretical overview invites predictions about the role of pattern structure on selective attending. Our plan is to initially compare some of these predictions in Experiment 1, where all five theories provide predictions. This is a baseline experiment in which none of the sequences contain rhythmic irregularities. Thus, in Experiment 1, hypotheses from all five theories concern the effects of pitch proximity in global (pattern context) and local (target distance) pitch structure on pitch identifications of probe tones that always occurred as on-time probes. In Experiments 2 and 3, where we vary probe timing, we further evaluate the pitch structure hypotheses assessed in Experiment 1 for the three models that specifically address probe timing, i.e., Large and Jones (1999), Pitch/Time Entrainment, and Pitch Space approaches. Experiment 3 differs from Experiment 2 with respect to a manipulation of instructions about probe timing. Whereas in Experiment 2 listeners were not told about probe timing variations (implicit instructions), in Experiment 3 they were informed of this variable (explicit instructions). Experiment 3 continues to focus on the same three theories, but it assesses the Pitch Space and entrainment views about the role of instructions versus pattern structure on attending. This experiment considers whether instructions about probe timing can over-ride influences of pattern structure on performance.

2. Experiment 1: A baseline experiment with no rhythmic violations

Experiment 1 examines the role of local versus global pitch structure in monitoring sequences where all probes occur at temporally expected times. An omnibus null hypothesis predicts effects of neither pitch variable, whereas various alternative hypotheses emphasize different roles for global and local pitch proximity in selective attending to pitch.

If global pitch structure guides selective attending then its effects may be evident in performance with probes that occur later in a sequence (cf. Fig. 3). Global pitch proximity occurs when the size of pitch intervals in context cells of a pattern are Narrow (versus Wide). According to Pitch/Time Entrainment reasoning, a Narrow context should promote efficient anticipatory attending to later sequence tones because it establishes a narrow focus in pitch/time; it predicts superior performance for on-time probes in the Narrow conditions. The I-R approach, broadly interpreted, leads to a similar prediction. Scene Analysis, on the other hand, implies an interaction of context with target distance where best performance should occur in the Narrow/Large condition (i.e., a probe tone "stands out" better following the grouping elicited by a Narrow context). Finally, the Pitch Space model shares with the strict entrainment model a null prediction about manipulations of global pitch structure.

Next consider local pitch structure and the target distance variable. If local pitch proximity between a probe and neighboring tones influences selective attending, then performance with on-time probes should be facilitated in Small Target Distance conditions (versus Large Target Distance). Both of the attentional focus theories (Pitch/Time Entrainment and Pitch Space) as well as the I-R models predict such facilitation. But, Scene Analysis offers a contrasting prediction, namely that local proximity will interfere with performance. It predicts that best performance should appear in Narrow/Large and that relatively poor performance should appear in Wide/Small and Narrow/Small.

Finally, it is possible that best performance will occur in patterns in which pitch intervals of the context variable (first two cells of a pattern) are congruent with (i.e., similar to) those of the target distance variable (third cell of a pattern). This prediction does not grant special status to proximity; rather it enlists a memory code for pitch patterns which is strengthened through repetition of similar pitch intervals within global and local contexts. Thus, to the extent prior context and local context match, certain memory approaches imply that positive priming will benefit congruent pitch structures (Narrow/Small and Wide/Large) over incongruent ones (Narrow/Large and Wide/Small).

2.1. Methods

2.1.1. Participants

Thirty-six non-musicians with normal hearing and fewer than 6 years of formal musical training participated in this experiment for credit in a psychology course at The Ohio State University. Two subjects' data were eliminated as they responded to fewer than 90% of trials.[4]

2.1.2. Equipment

All aspects of stimulus generation and response collection were controlled by MIDILAB Version 6.0 software (Todd, Boltz, & Jones, 1989). Sounds were generated on an IBM PC compatible computer interfaced with a Roland MPU-401 MIDI Processing Unit that controlled a Yamaha TX81Z tone generator set to a sine wave voice. The stimuli were transmitted to a separate sound attenuated room and amplified using the Rane HC-6 Headphone console. Each participant listened to the stimuli over Beyerdynamic DT 770 headphones at a comfortable listening level (ca. 70 dB).

---

[4] In this and following experiments, these criteria eliminated subjects who fell into one of two categories: 1. The individual fell asleep during an early morning experimental session; 2. The individual failed to return for the second day session.

2.1.3. Stimuli and conditions

Experimental patterns were composed of nine tones and comprised 3 three-tone-cells as shown in Fig. 3A (e.g., for a Narrow/Large pattern). The pitch interval between the central tone and flanker tones in the first 2 three-tone-cells was varied to create a Narrow or Wide pitch interval context. For the Narrow context, the pitch interval was 5 semitones (STs) in the first cell and 3 STs in the second cell. For the Wide context, the pitch interval was 11 STs in the first cell and 9 STs in the second cell. The pitch interval between the central tone and flanker tones in the third three-tone-cell was varied to create a Small or Large target pitch interval. For Small targets the pitch interval was 3 STs and for Large targets the pitch interval was 9 STs. These variables were crossed to create sequences that had Narrow context/Small targets, Narrow context/Large targets, Wide context/Small targets, and Wide context/Large targets. In addition, the pitch interval between the last tone of a three-tone-cell and the first tone of the next three-tone-cell was always 2 STs. For example, given a Narrow context/Large target the change in ST between tones would be: +5, −5, −2, −3, +3, +2, +9, −9. Different instances of the four types of patterns were created with two different initial pitches of C#4 and G4. Instances of each pattern type were constructed such that over a session, the conditional probability of a target pitch, given the pitch of the immediately preceding tone, was .50 when considered over the set of eight different patterns within a session. Thus, neither the immediately preceding pitch nor the preceding pitch interval could probabilistically cue a given target pitch.

On each trial a specific nine-tone sequence was presented three times (three cycles); on the third cycle, the pitch of the eighth tone (the probe) could change. Equally often the probe remained the same or was higher or lower by 1 ST. All nine-tone melodies were isochronous with a 600 ms rate; all tone durations were 137 ms. In addition, 1800 ms elapsed between the onset of the last tone of one cycle and the onset (tone) of the next cycle.

A control set of trials was also presented to listeners to assess discrimination of the target/probe pairs in the absence of surrounding pattern tones. In control trials, listeners heard only two 137 ms tones, separated by an IOI of 6378 ms. The pitch of the first tone was E4, A#4, or E5 as the sequence target always assumed one of these values; probe tones were equally often 1 ST lower, the same pitch or 1 ST higher.

2.1.4. Pattern validation

To assess pattern manipulations, two supplementary experiments were conducted using expert listeners (at least 10 years of formal musical training). Our aim was to determine the extent to which patterns were judged as chromatic (versus diatonic) and whether this covaried with structural manipulations. In addition, we assessed the degree to which different tones in a sequence functioned as good completions when they ended truncated versions of the pattern.

Expert listeners first heard two presentations of each of the eight sequences (randomized order). They were instructed to rate each pattern on a 7 point scale along a dimension labeled diatonic/chromatic (1 = highly diatonic and 7 = highly chromatic). Mean ratings were 2.81, 3.19, 4.1, and 4.25 for Narrow/Small, Narrow/Large, Wide/Small, and Wide/Large conditions, respectively. Although this indicates that patterns with wide context intervals tend to be heard as somewhat more chromatic, a two-way ANOVA (two levels of context crossed with two levels of target distance) revealed no statistically significant main effects or interactions. Although there was a non-significant tendency to judge patterns in the Narrow/Small condition to be more diatonic, in the main these listeners found all patterns to be rather neutral, i.e., not highly chromatic or highly diatonic.

Next, the same listeners judged the degree to which sequences in each condition seemed to end with a "closed" (i.e., a realized interval by I-R accounts) versus "open" (i.e., implicative) pitch interval. They heard four different versions of each sequence in which 0, 1, 2, or 3 tones were eliminated from the end of each pattern. On a scale of -4 (highly closed) through 0 (neutral) to $+4$ (highly open), they judged how well each ending tone completed the pattern (instructions were readily understood). Patterns were presented in blocked format, beginning with random arrangements of patterns with 2 or 3 missing tones (inter-mixed); then they heard a similar inter-mixed set of patterns with 0 and 1 missing tones. Listeners rated each sequence twice; these ratings were averaged. A similar profile over pattern length emerged for all four types of patterns. Patterns ending on one of the two tones flanking the target tone (0 and 2 missing tones) produced good completion (closed) judgments (mean of -3.35), whereas patterns ending on the target/probe location (1 and 3 missing tones) were judged as open or unresolved (mean of $+2.95$). A one-way ANOVA showed these differences to be significant, $F(3,21) = 39.98$, $p < .001$. According to certain musical expectancy theories (e.g., I-R models), this suggests that the target tone ends an implicative interval and the final pattern tone ends a realized interval.

Taken together, these analyses indicate that: (1) None of the patterns are strongly diatonic; (2) All induce conventional expectancies associated with resolution of the ninth (last) tone i.e., closure; (3) The probe/target location is associated with an unresolved judgment.

2.1.5. Design

The main design was a (2×2) repeated measures design with a total of 24 unique stimulus sequences which crossed two within subjects variables, context (Narrow and Wide) and target (Small and Large). [Counterbalancing variables, such as levels of initial pitch (C#4 and G4) and probe (lower, same, and higher), were randomized over trials and collapsed for data analysis]. The design for control trials similarly included within subjects variables of initial pitch (E4, A#4, and E5) and probe (lower, same, and higher) leading to nine unique pairs of tones randomly ordered.

2.1.6. Procedure

Listeners received recorded instructions supplemented by a diagram indicating pattern cycles and target/probe serial position. They were told that to "pay attention" to the pitch of the target, which was always the eighth tone in a recurring nine-tone pattern. They were asked to judge whether its pitch changed in the last cycle and to respond "higher," "same," or "lower" by pressing a button on a labeled response panel.

Participants received 16 practice trials (with visual feedback). Practice trials provided four instances of each level of context structure and target distance. Practice was followed by three sets of 48 test trials (no feedback) and one set of 54 control trials. All trials were automatically timed; participants had 5 s to respond after the final sequence tone before the next trial began. Presentation order was randomized with the constraint that no more than three instances of any level of a variable could occur consecutively. Listeners heard each unique experimental sequence six times and each unique control sequence six times. The total duration of the experiment was 1.5 h.

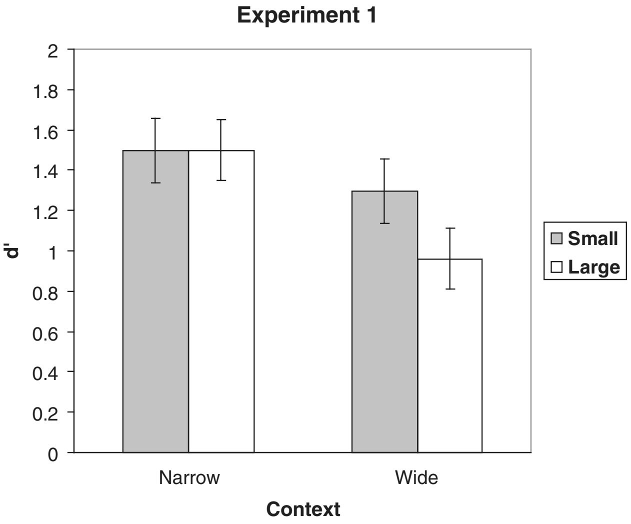

Fig. 5. Mean d' scores as a function of context and target distance in Experiment 1.

2.2. Results and discussion

We computed both d' scores and Proportion Correct (PC). Results from both analyses were similar, differing only in minor details. Accordingly, we report only d' scores were based on hits to probes that were identical to the preceding target pitch ("Same"/SAME) and the mean of false alarm rates to probes that were higher or lower in pitch than the target.[5] We eliminated from full analysis the data from one subject[6] who scored below chance levels (of PC) on the final control block where listeners had to identify the pitch changes based on each of the target pitches, absent any intervening tones (mean PC was .817). The mean d' score for the remaining subjects in this control trial block was 2.63.

Overall the mean d' in Experiment 1 was 1.315 (mean PC was .618). Fig. 5 presents mean d' scores for the four pitch structure conditions, collapsed over initial pitch. A 2×2 repeated measures factorial ANOVA that crossed two levels of context with two levels of target distance indicated that only context significantly affected sensitivity Over the four experimental conditions, mean d' scores ranged from 0.96 (Wide/Large) to 1.50 (Narrow/Large). Overall, people were better with the Narrow than the Wide context, $F(1,32) = 12.92$, $MSe = 4.463$, $p < .001$. The mean d' for sequences with Narrow context intervals was 1.50, whereas the corresponding mean for sequences with Wide context

---

[5] In all three experiments, the d' scores were calculated in two other ways as well. We also computed an average d' based on a d' for Same versus Higher pitch comparison and one for Same versus Lower pitch comparisons; and we computed the d' total scores based on categorical ratings analysis (MacMillan and Creelman, 1991; pages 221–230). All three types of d' scores as well as PC data resulted in the same basic pattern of findings in statistical analyses; we have reported only the simplest one.

[6] In all experiments, chance PC is .33. The mean PC for the eliminated subject in Experiment 1 was .29; in Experiment the two eliminated subjects score .21 and .20, respectively; in Experiment 3.

intervals was 1.13 (Table A1 of Appendix A presents mean d' scores for these conditions for all three experiments; Table A2 presents corresponding Hit and False Alarm rates).

These findings are consistent with approaches that incorporate a role for global, i.e., contextual, pitch proximity. The main finding, favoring patterns with Narrow context intervals, suggests that the pitch relations which initiate a serial pattern can affect selective attending to tones much later in a sequence. This favors Pitch/Time Entrainment and I-R models.

However, other data of Experiment 1 provide little support for theories that hypothesize a dominant role for local pitch proximities (i.e., Small target distances) and they also rule out congruency among pitch intervals as a major influence on selective attending. The failure to find any significant effect of target distance is inconsistent with all of the pitch structure theories (e.g., Pitch/Time Entrainment, Pitch Space, I-R models, Scene Analysis). However, I-R theorists can argue for null results on the basis that the probe tone is not an ending tone (i.e., it completes an implicative, not a realized, pitch interval). As well, listeners were not significantly benefited by the presence of Large target intervals following a Narrow context as predicted by Scene Analysis for the Narrow/Large condition [the interaction of context and target distance yielded an $F(1,32) = 3.242$, $Mse = .287$, $p = .08$]. A trend in the data was consistent with the Scene Analysis prediction that selective attending is favored in the Narrow/Large condition, i.e., versus the Wide/Large condition (a planned comparison showed this difference was significant, $p < .05$). But a critical comparison for Scene Analysis is between Narrow/Large and Narrow/Small, and this planned comparison was not significant. Indeed, these two conditions produced identical performance. In the absence of a strong interaction, indicating singularly best performance for the Narrow/Large condition, we conclude that none of the four pitch structure theories adequately explains performance as function of target distance in these patterns. Finally, the lack of a significant interaction of context with target distance also rules out general memory interpretations based on repetitive similarity which predict benefits due to context-pitch distance congruency.

In sum, Experiment 1 provides two clear conclusions. First, global pitch structure affects selective attending in rhythmically isochronous patterns, with Narrow conditions leading to better performance than Wide conditions. This favors Pitch/Time Entrainment and I-R approaches. Second, variation in local pitch structure (as target distance) has no systematic influence on performance. This favors the Large and Jones (1999) entrainment model (and possibly I-R theory).

3. Experiment 2: Rhythmic violations and implicit instructions

Experiment 2 builds on Experiment 1 to examine effects of pitch structure on expectancy profiles. We use the same task, stimuli and instructions as those used in Experiment 1. Although we make no explicit mention of probe timing in the instructions, the single difference between the two experiments involved variation in probe times such that in Experiment 2 the probe tones do not always occur at the rhythmically expected time. Instead, equally often a probe occurs on-time, early, or late. Our predictions for on-time probes in Experiment 2 draw on the same pitch structure hypotheses outlined for Experiment 1; for these rhythmically expected probes, we should replicate the Experiment 1 finding of superior performance with Narrow over Wide context conditions in Experiment 2. However, with rhythmically unexpected probes we encounter a new situation. Because neither I-R models nor Scene Analysis theories provide explicit predictions about probe timing in this task, we focus upon the three models that do, namely Large and Jones (1999), Pitch/Time Entrainment, and Pitch Space accounts. These theories offer different answers to the question: "What aspects of a pattern determine performance with on-time versus ill-timed probes?"

The answer provided by Large and Jones (1999) is that rhythmic coherence is the sole determinant of performance with on-time versus ill-timed probes. With a coherent, i.e., isochronous intensity/time pattern, a narrow focus in time is predicted, meaning that early and late probes receive less attention than on-time probes and should be identified less accurately, regardless of pitch structure. The model predicts equally sharp expectancy profiles for all four pitch structure conditions.

Alternative hypotheses feature a role for pitch structure. The Pitch/Time Entrainment model predicts an interaction of target distance with probe time: Small target distances will *degrade* performance with ill-timed probes. Small target distances (versus Large ones) elicit narrow attentional foci, therefore the likelihood is lower that an ill-timed probe occurs within this temporal focus. Together with predictions about anticipatory attending to on-time probes, the Pitch/Time Entrainment account of reactive attending implies that we will find sharp expectancy profiles mainly for patterns in the Narrow/Small condition (Fig. 4, solid lines).

Finally, a third hypothesis comes from the Pitch Space model. Although this theory offers no predictions (or null effects) about global pitch structure, it offers a contrasting account of the interaction of target distance with probe timing, as illustrated in Fig. 4. According to this view, performance with ill-timed probes will be *facilitated,* relative to on-time probes, because these unexpected tones will summon attention when they are proximal in pitch space to a target tone, i.e., when target distance is Small (e.g., Mondor & Bregman, 1994; Mondor & Terrio, 1998). Shallow U expectancy profiles for Narrow/Small and Wide/Small conditions are predicted due to relatively better performance with early and late probes. However, performance should be poorer, overall, in patterns with Large target distances (versus Small ones) leading to flat expectancy profiles (Fig. 4, dashed lines).

In summary, whereas a strict entrainment model predicts no differences among expectancy profiles of the four pitch structure conditions, Pitch/Time Entrainment and Pitch Space models both feature roles for pitch structure. Although the latter two theories both predict an interaction of target distance (local pitch structure) with probe timing, they differ in the nature of this interaction. The most striking difference involves the Narrow/Small condition for which the Pitch /Time Entrainment model predicts a sharp inverted U expectancy profile and the Pitch Space account predicts a shallow U expectancy profile.

3.1. Methods

3.1.1. Participants

Thirty-three non-musicians with normal hearing and fewer than 6 years of formal musical training participated in this experiment for credit in a psychology course at The Ohio State University. Four subjects' data were eliminated as they responded to less than 90% of trials.

3.1.2. Equipment

Identical to Experiment 1.

3.1.3. Stimuli and conditions

The experimental sequences of Experiment 1 were modified for use in Experiment 2 by varying probe time. Probes could occur 20% early (−120 ms), on-time, or 20% late (+120 ms). This was a compensatory time shift, meaning that onset times of the surrounding tones remained the same regardless of the probe time shift. Control trials were identical to Experiment 1.

3.1.4. Design

Within subjects variables included context (Narrow and Wide), target (Small and Large), and probe time (early, on-time, and late) leading to a (2 × 2 × 3) repeated measures design. Counterbalancing variables of initial pitch (C#4 and G4), probe pitch (lower, same, and higher) were combined with the main independent variables to generate a total of 72 unique sequences. Again, control trials included within subjects variables of initial pitch (E4, A#4, and E5) and probe (lower, same, and higher), resulting in a total of nine unique pairs of tones randomly presented over control trials.

3.1.5. Procedure

Listeners received the same recorded instructions/diagram as in Experiment 1. Sessions were conducted over 2 days. On day 1, participants received 16 practice trials (with feedback) followed by two sets of 72 test trials (no feedback). On day 2, participants received the same set of practice trials, followed by one set of 72 test trials and one set of 54 control trials. The order of presentation of trials was randomized with the constraint that no more than three instances of any level of a variable could occur consecutively. As in Experiment 1, listeners heard each unique experimental sequence 3 times and each unique control sequence 6 times. The experiment lasted 2 h on day 1 and 1.5 h on day 2.

3.2. Results and discussion

Again we computed both d' scores and Proportion Correct (PC) and found a similar pattern of findings. Also we eliminated from full analysis the data of two subjects who scored below chance levels (PC = .33) on the final control block where listeners had to identify the pitch changes based on each of the target pitches, absent any intervening tones (cf. footnote 6). The mean d' score for the remaining subjects in this control trial block was 2.23 (mean PC = .752).

Overall, the average d' score for Experiment 2 in experimental sessions was 1.006 (mean PC = .563). Fig. 6 presents mean d' scores for the four pitch structure conditions (collapsed over initial pitch) as a function of the three probe time levels (early, on-time, and late). Mean d' scores over the 12 conditions ranged from a $d' = 0.66$ (Wide/Large, early) to a $d' = 1.51$ (Narrow/Large, on-time). A repeated measures ANOVA combining context, target distance, and probe timing revealed two significant main effects involving, respectively, context and probe timing, but no reliable interactions. Overall, listeners were more sensitive to recurrences of the target tone with the Narrow context (mean $d' = 1.12$) than with the Wide context (mean $d' = 0.89$), $F(1,26) = 11.219$, $MSe = 0.374$, $p < .005$. This replicates our finding in Experiment 1. In addition, in Experiment 2 people generally performed better with on-time probes (mean $d' = 1.29$) than with early (mean $d' = .74$) or late probes (mean $d' = .99$), $F(2,52) = 27.879$, $MSe = 0.291$, $p < .001$. With regard to differences among the three probe timing means within Experiment 2, all pair-wise comparisons (averaged

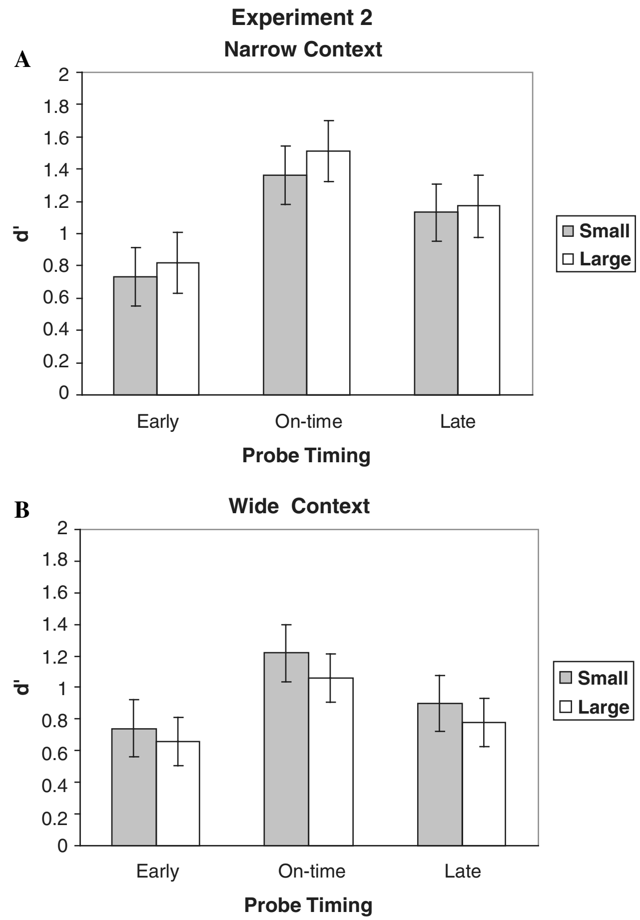

Fig. 6. Experiment 2 data. (A) Mean *d*′ scores as a function of target distance and probe timing in Narrow context. (B) Mean *d*′ scores as function of target distance and probe timing in Wide context.

over pitch structure levels) were significant according to the Tukey HSD, $p < .05$. These data replicate and extend those from experiments in which expectancy profiles have been reported for timing variations of ending tones (Jones et al., 2002).

Again, the main of effect of target distance was not significant as in Experiment 1 nor did it interact with context ($p = .28$).[7] More importantly, we failed to find an interaction of target distance with probe timing [$F(2,52) = 0.94$]. Planned pair-wise comparisons of quadratic trends within all four pitch structure conditions confirm that expectancy profiles did not differ significantly across the four pitch structure conditions, although a significant

---

[7] We also examined performance as a function of pitch structure by pooling data of Experiment 2 with those of Experiment 1 (both for all timing levels in Experiment 2 and for only on-time probes). We found a significant interaction of Target Distance with Context ($p < .05$) when we included performance on ill-timed probes; this was due to poor performance in the Wide/Large condition. However, when only on-time probe data were analyzed, no significant interaction of Context with Target Distance emerged.

quadratic trend as a function of probe timing was found in each condition [relevant individual quadratic trend $F(1,26)$ ratios were 6.55 (Narrow/Small), 21.22 (Narrow/Large), 10.30, (Wide/Small), and 7.64 (Wide/Large); all $p < .025$].

In Experiment 2, the findings with respect to global and local pitch structure agree with those of Experiment 1. Narrow context (i.e., contextual proximity) again produced better overall levels of pitch identification than the Wide context, a finding consistent with predictions of the Pitch/Time Entrainment account, regarding this variable. Moreover, the fact that performance with on-time probes in Experiment 2 was comparable to performance in Experiment 1 suggests that Narrow context conditions facilitate performance with rhythmically expected probes even when timing variability exists within a session. The absence of an interaction of context with probe timing is consistent with an assumption of entrainment models, namely that reactive attending to expectancy violations (ill-timed probes) is contingent on anticipatory attending to on-time probes. Expectancy profiles are instructive in this regard because they indicate that relative levels of performance to unexpectedly timed probes (versus expectedly timed probes) do not differ with respect to pitch structure. Thus, we infer that absolute levels of performance are due primarily to anticipatory and not reactive attending. We also infer that, consistent with the Pitch/Time Entrainment approach, anticipatory attending is affected by global pitch structure.

Relative performance levels that are evident in expectancy profiles suggest that reactive attending is not directly affected by manipulations of pitch structure. Not only is there an absence of an interaction of target distance with probe timing, but expectancy profiles are very similar across pitch pattern conditions. This finding is not consistent with predictions of either the Pitch/Time Entrainment or the Pitch Space account. Instead it supports predictions of a strict entrainment model (e.g., Large & Jones, 1999).

In sum, we draw two clear conclusions from the data of Experiment 2: first, Narrow contexts facilitate selective attending to rhythmically expected tones within a sequence. Second, neither context (global pitch structure) nor target distance (local pitch structure) affected relative performance levels with ill-timed probes.

4. Experiment 3: Rhythmic violations and explicit instructions

The findings of Experiment 2 are more in line with entrainment explanations (Large & Jones, 1999; Pitch/Time Entrainment) than with a Pitch Space account. However, it can be argued that instructions for the task used in Experiment 2 did not fulfill an important assumption of the Pitch Space theory regarding prioritizing relevant versus irrelevant task dimensions for top-down control of attending. A more appropriate test calls for manipulations designed to directly test voluntary versus automatic attending. Experiment 3 pursues this possibility by introducing instructions that explicitly identify task-relevant (pitch) and irrelevant (time) dimensions.

The neutral (implicit timing) instructions of Experiment 2 may have prevented listeners from developing an effective strategy for responding to unexpected temporal onsets that would enable attentional capture, as predicted by the Pitch Space theory (e.g., see Lamy, Leber, & Egeth, 2004). In Experiment 3, we designed instructions to explicitly encourage voluntary orienting to the pitch dimension and to discourage a listener's surprise with 'irrelevant' timing variations. Following a Pitch Space account, we assumed that instructions would *facilitate* overall performance, perhaps by allowing a voluntary over-ride of any negative influences of sequence structure (e.g., due to unexpected probe times/pitch intervals). In the most extreme case, this could yield uniformly high performance for all probe times, in which quadratic trends of expectancy profiles vanish in all four pitch conditions of Experiment 3.

Although it is plausible that instructions guide voluntary attending, current entrainment theories do not acknowledge this. These approaches imply that attending, particularly reactive attending, is guided by pattern structure in a largely automatic fashion. Thus, advance information about task goals should not affect sequence monitoring. In the extreme case, entrainment accounts predict that the results of Experiment 3 should be identical to those of Experiment 2. Any differences between the outcomes of the two experiments either in performance with on-time probes (anticipatory attending) or with ill-timed probes (reactive attending) will signal the operation of a voluntary component.

We have identified two extreme positions that can be assessed in Experiment 3. According to one position, all expectancy profiles are flat (with high PCs), whereas according to other all profiles are sharp (inverted U's) and unaffected by either pitch structure or instructions. It is unlikely that either of these extreme positions is correct. More interesting (and more likely) are outcomes in which instructions confer a selective influence on attending activities. Attending may be more susceptible to instructions when listeners monitor certain kinds of patterns (e.g., in the Wide/Large condition) than others (e.g., in Narrow/Small). Or, they may have a singular impact on one kind of attending (e.g., anticipatory attending). In short, we hope to identify specific changes in the pattern of outcomes in Experiment 3 relative to Experiment 2 that shed greater light on attending in sequence monitoring.

Of the present theories, only the Pitch Space model offers a reason to expect instructions to differentially impact performance. It predicts that instructions will orient listeners to pitch features of the target/probe and thus promote an interaction of target distance with probe timing. This interaction should reflect attentional capture to ill-timed probes with Small target distance conditions (cf. Fig. 4; dashed lines). Although such predictions were not supported in Experiment 2, Experiment 3 provides a fairer test of this model. If this view is correct, then we expect to find relatively high performance levels for Small target distance conditions with all probe time levels, but especially for early and late probes (due to capture). Shallow U expectancy profiles are predicted for the Narrow/Small and the Wide/Small conditions, whereas flat profiles should occur with Large target distance conditions. Although the Pitch/Time Entrainment model also predicts an interaction of target distance with probe time, the nature of the interaction differs for these two models; furthermore, the Pitch/Time Entrainment model provides no rationale for such an interaction to emerge as a function of instructions (i.e., in Experiment 3).

In sum, all three models (Large & Jones, 1999; Pitch/Time Entrainment, Pitch Space) offer the same predictions for Experiment 3 as they did for Experiment 2, but for different reasons.

4.1. Method

4.1.1. Participants

Thirty-six non-musicians with normal hearing and fewer than 6 years of formal musical training participated in this experiment for credit in a psychology course at The Ohio State University. One subject's data was eliminated as s/he performed below chance (.33) and five subjects' data were eliminated because they responded to less than 90% of trials.

4.1.2. Equipment

Identical to Experiment 1.

4.1.3. Stimuli and conditions

Identical to Experiment 2.

4.1.4. Design

Identical to Experiment 2.

4.1.5. Procedure

The only change made in Experiment 3 is the instructions listeners received. The instructions from Experiment 2 were supplemented with explicit information regarding the timing of the probe. Participants heard recorded instructions and were given a diagram that showed that the eighth tone in the third cycle could sometimes be early or late. They also heard examples of the melody with early, on-time, and late probes. It was emphasized that the main goal of the task was to identify the pitch of this temporally variable tone; they were instructed to ignore these timing variations and to pay close attention to the pitch of the target/probe tone whenever it occurred.

4.2. Results and discussion

Both d' scores and Proportion Correct (PC) produced similar patterns of findings for Experiment 3, hence again we report only d' scores. No subjects were eliminated due to poor performance on the control trial block. In addition, control data were not collected from five subjects due to experimenter error. The mean d' score for the 25 remaining subjects in this control trial block was 2.57 (mean PC was .837).

Overall mean d' for experimental sessions in Experiment 3 was 1.129 (mean PC was .589) which did not differ significantly from the average d' value of Experiment 2. Fig. 7 presents mean d' scores for the four pitch structure conditions in Experiment 3 as a function of the three levels of probe timing (early, on-time, and late). Over the resulting 12 conditions, mean d' scores ranged from 0.53 (Wide/Large, early) to 1.72 (Narrow/Small, on-time). We first report the results for Experiment 3 and then compare these with findings of Experiment 2.

A repeated measures ANOVA combining context, target distance, and probe timing revealed three significant main effects involving, respectively, context, target distance and probe timing, and an interaction of target distance with probe timing. Overall, listeners were again more sensitive to recurrences of the target tone in the Narrow context (mean $d' = 1.23$) than in the Wide context (mean $d' = 1.03$), $F(1,29) = 7.341$, $MSe = 0.516$, $p < .025$. Once more, probe time affected performance; people were more accurate with on-time probes (mean $d' = 1.29$) than with early (mean $d' = .89$) or late probes (mean $d' = 1.199$), with the timing variable yielding an $F(2,58) = 8.826$, $Mse = 0.585$, $p < .001$. A Tukey post hoc (HSD) indicated that performance on early probes differed significantly from on-time and late ones ($p < .010$).

For the first time target distance exerted a significant influence on performance with higher sensitivity for Small (1.30) than for Large (0.95) target/probe pitch intervals, $F(1,29) = 13.279$, $Mse = 0.815$, $p < .0025$. This was due, in part, to a significant interaction of target distance with probe timing, $F(2,58) = 3.371$, $Mse = 0.432$, $p < .05$. Tukey HSD tests confirmed that listeners were significantly better with Small target distances than with Large ones for on-time probes ($p < .005$). In addition, the average d' scores for early and on-time probes differed for the Small target distances (poorer performance for early probes; $p < .005$), whereas this was not true for Large target distances.

Fig. 7 illustrates the significant interaction of target distance with probe timing. Small target distances produced better performance probes than did Large target distances, but this was especially true for on-time probes. Because theoretical issues center on changes in expectancy profiles we used planned comparisons of trends over probe time to assess these differences. These analyses reveal that, in contrast to Experiment 2, in Experiment 3 different expectancy profiles emerge across the four pitch conditions. Specifically, significant quadratic trends (inverted U's) occurred only in conditions with Small target distances where, $F(1,29) = 8.2$, $Mse = 0.659$, $p < .01$ and $F(1,29) = 5.09$, $Mse = 0.46$, $p = .03$ for Narrow/Small and Wide/Small trend components, respectively. The more pronounced of these two profiles occurs in the Narrow/Small condition. By contrast, non-significant trend components were observed in conditions with Large targets. Collapsing over context and comparing trends in Experiment 3 as a function of target distance, we find a quadratic

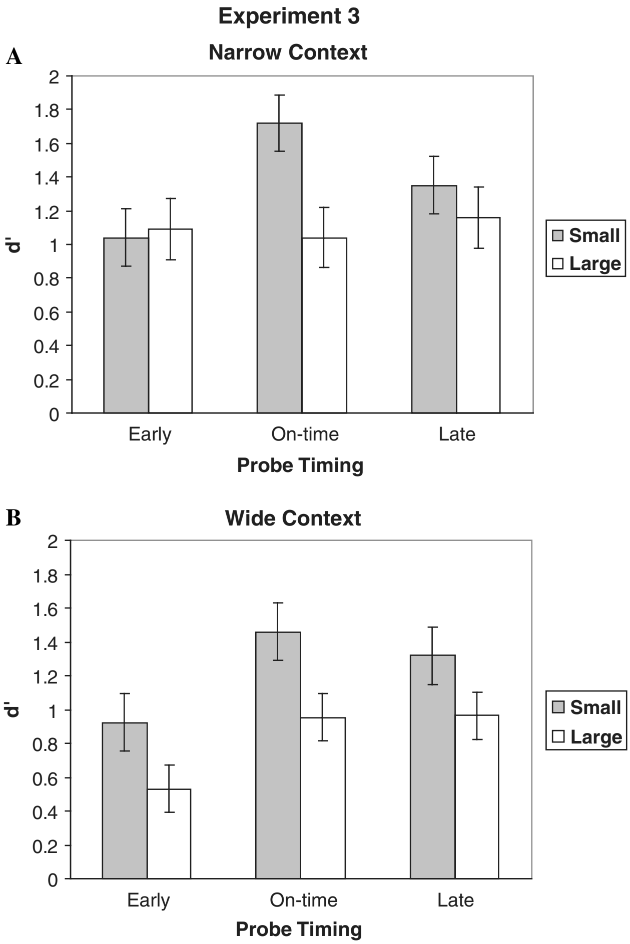

Fig. 7. Experiment 3 data. (A) Mean d' scores as a function of target distance and probe timing in Narrow context. (B) Mean d' scores as function of target distance and probe timing in Wide context.

trend with Small targets [$F(1,29) = 10.50$, $Mse = .707$, $p = .003$], but not with Large targets ($p = .56$). This represents a significant trend change, $F(1,29) = 5.36$, $Mse = 0.516$, $p = .028$. This is due primarily to the fact that instructions facilitated performance with on-time probes embedded in patterns with Small target distances, whereas instructions appeared to dampen responding to on-time probes within patterns with Large target distances. Tukey HSD tests confirm that the only point of significant difference between the two profiles is due to a difference in performance with on-time probes ($p < .001$).

Contributions of the context variable to differences among expectancy profiles in Experiment 3 are modest, but worth mentioning. Although our omnibus test did not reveal a significant three-way interaction of context, target, distance, and probe, planned comparisons of expectancy profiles indicated that context seemed to operate synergistically with target distance. This is evident in Fig. 7: Small target distances lead to different profiles as a function of context. Thus, although a quadratic trend over probe time levels was observed in the Wide/Small condition, $F(1,29) = 5.09$,$Mse = 0.458$, $p < .05$, the profile found in the Narrow/Small condition was significantly stronger, $F(1,29) = 8.2$, $Mse = 0.66$, $p < .01$, than that of the Wide/Small. Thus, the Narrow context appears to slightly enhance effects of Small target distances for on-time probes leading to sharper expectancy profiles in the Narrow/Small condition.

Instructions exerted a selective influence on performance in Experiment 3. Nonetheless, although the Pitch Space model provides a role for the instruction variable, the expectancy profiles of Fig. 7 do not fully conform to predictions of this approach (cf. Fig. 4). The Pitch Space model correctly predicts the target distance main effect and the flat expectancy profile observed in the Narrow/Large condition, but it fails to predict the significant quadratic trends observed in the other three conditions. In all conditions, except the Narrow/Large condition, we find inverted U shaped functions, indicating that ill-timed probes hurt, rather than help, performance. This finding remains problematic for the Pitch Space model which predicts that ill-timed probes will capture attention and thereby facilitate performance in the two pitch conditions that contain small target distances. Instead, we find that in these conditions both early and late probes actually produce *worse*, not better, performance than is evident with rhythmically expected (on-time) probes.

In sum, we can eliminate the two most extreme positions of attentional control in sequence monitoring. Attending is neither entirely driven by pattern structure, nor is it entirely determined by instructions. Rather, the influence of instructions on sequence monitoring appears to be selective; it is qualified by pattern structure. In some conditions, instructions appear to facilitate selective attending whereas in others, instructions have little, or even a slightly negative, influence. These effects are evident mainly in the Experiment 3 finding that listeners are generally better in conditions with Small targets than in those with Large targets, particularly when probes are on-time. To verify this influence, we compare the data of Experiment 3 with those of Experiment 2 in detail.

4.3. Experiment 3 versus Experiment 2

We used a mixed factorial ANOVA with two levels of instructions (Experiment 2, Experiment 3) crossed with three levels of probe timing (early, on-time, and late), two levels of context (Narrow, Wide), and two levels of target distance (Small, Large). As expected, we found significant main effects of context, target distance and probe timing as well as a significant interaction of instructions (experiment) with target distance (no other significant effects were observed). Here, we concentrate only on the latter interaction.

Table 1
Mean d' values for Experiments 2 and 3 as a function of Target Distance

Target Distance	Experiment	
	2	3
Small	1.012 (.34)	1.302 (.32)
Large	1.001 (.32)	0.956 (.30)

Although instructions had no significant influence on overall performance, they did selectively impact performance as a function of target pitch distance. The interaction of experiment (2 versus 3) with target distance (Small versus Large), yielded an $F(1,55) = 7.50$, $Mse = 0.64$, $p < .01$. These data appear in Table 1. Mean d' values were significantly higher for Small target distances than for Large target distances, but only under explicit timing instructions (Experiment 3).

The three-way interaction of instructions (Experiment), target distance and probe timing did not attain significance. This indicates that explicit instructions generally improve performance in patterns with Small target distances but not with those containing Large target distances where overall performance (collapsed over probe time) did not change. However, we relied on planned comparisons, suggested by Fig. 4, to pursue theoretical issues that relate to changes in quadratic trends in Experiment 3 versus Experiment 2. Specifically, we considered whether instructions levied significant changes in expectancy profiles in any of the four pitch structure conditions. Only the Narrow/Large condition showed a significant change in the observed quadratic trend from Experiment 2 (Fig. 6A; gray bars) to Experiment 3 (Fig. 7A), $F(1,55) = 9.27$, $MSe = 0.38$, $p = .004$. The flatter Narrow/Large profile in Experiment 3 arises from a combined effect of small *increases* in d' for the two ill-timed Large target probes (relative to Experiment 2) and a large *decrease* in d' for the on-time Large target probes (i.e., from a d' of 1.51 to one of 1.04 in Experiment 3). By contrast, performance changes in the opposite direction occur when Small targets are present in the Narrow context (i.e., the Narrow/Small condition), in Experiment 3, performance with on-time probes *increased* from a d' of 1.36 (Experiment 2) to one of 1.72 (in Experiment 3). Finally, in the two Wide context conditions, although similar differences due to target distances are evident in Experiment 3, the Wide/Large condition did not show a statistically significant flattening of the quadratic trend from Experiment 2 to Experiment 3 [$F(1,55) = .42$]. However, it should be noted that the quadratic trend in the Wide/Large condition although weak, was statistically significant in Experiment 2 but not in Experiment 3 where it vanished altogether.

4.4. Discussion

Instructions appear to exert a selective influence on performance that depends on pattern structure. When people are told to ‘pay attention to pitch and not time’ they comply by performing best in monitoring patterns that contain Small target distances especially when probe tones are rhythmically on-time. Moreover, relative to Experiment 2, where instructions were neutral, there is no real benefit from explicit instructions in monitoring patterns containing Large target distances (in Experiment 3). In fact, with patterns containing Large target distances (Narrow/Large, Wide/Large) listeners perform relatively poorly with on-time probes.

Although the Pitch Space model offers significant potential for addressing the role of instructions, this does not translate into an explanatory advantage when we compare the data of Experiments 2 and 3. Instead of selectively improving performance with ill-timed probes in patterns containing Small target distances, in Experiment 3 instructions generally improved performance in patterns with Small target distances, particularly for on-time probes in these patterns. In Experiment 3, we see that significantly greater benefits accrue to Small targets that are on-time than to ill-timed Small target tones. In other words, although the Target Distance by Probe Time interaction (Fig. 4) is significant in Experiment 3, the particular form of this interaction, predicted by a Pitch Space model, did not emerge. This theory correctly predicts flat expectancy profiles for Large target distances, but its prediction of shallow U shaped profiles for the Small target distances in Experiment 3 is incorrect.

Despite its failure to incorporate a role for instructions, the Pitch/Time Entrainment theory provides a reasonable account of these data. It correctly predicts three of the four expectancy profiles plus the general effect of global pitch context. Assuming that observed expectancy profiles reflect the width of an attentional pitch–time focus, the predictions that attentional focus width varies with pitch structure receive some support in Experiment 3. This model correctly predicts the sharpest profiles for patterns with Narrow/Small conditions and flat profiles for the Wide/Large conditions. For the Wide/Small condition it correctly predicts an expectancy profile but its singular failure is in not predicting the flat profile of the Narrow/Large condition. Both the Narrow/Large and Wide/Large conditions showed significant quadratic expectancy profiles in Experiment 2 and in both cases these trends vanished in Experiment 3. However, the trend change was statistically reliable only in the Narrow/Large condition, mainly because the quadratic trend for the Wide/Large was weaker initially in Experiment 2. To sum up, where the Pitch/Time Entrainment account falls short is in assigning insufficient weight to local pitch structure as a basis for narrowing an attentional focus in the presence of instructions.

Reactive attending appears to be largely unaffected by the joint manipulations of pitch structure and instructions. The main difference between performance in Experiment 3 (relative to Experiment 2) involves overall performance levels in conditions with Small target distances; instructions did not selectively change performance with ill-timed probes in these conditions.

In sum, unexpectedly this set of findings is best described by the Pitch/Time Entrainment account. However, even this model is at a disadvantage because it cannot explain why the observed interaction of target distance with probe time occurs in Experiment 3 but not in Experiment 2. One interpretation of these findings is that instructions selectively affect anticipatory attending, as it operates with small pitch intervals, but not reactive attending. Relative to Experiment 2, the instructions in Experiment 3 appear to cause listeners to weight the contribution of target distance more heavily as a basis for anticipatory attending. As such, our experiments challenge extreme versions of both entrainment theory, which views sequence monitoring as an entirely pattern-driven activity, and Pitch Space theory, which views attending as mainly voluntary. They suggest a middle ground which we consider in Section 5.

5. General discussion

In attending to auditory sequences listeners are strongly influenced by early portions of a sequence, namely by the initial pattern of pitch and time intervals. In addition, they generally perform best in identifying the recurrent pitch of rhythmically expected tones that occur within simple pitch patterns, namely in sequences that contain mainly small pitch intervals. This is true in all three experiments, regardless of instructions. In other pitch patterns, a similar performance profile occurs over timing variations of a to-be-identified (probe) tone, but only when instructions are neutral. If listeners are warned in advance that they should focus on a particular tone's pitch and ignore its timing variations, these instructions primarily benefit performance in patterns that seem to project the simplest overall pitch structure, that based on many small pitch intervals.

In general, this monitoring task is relatively difficult. This is indicated by the fact that identification performance on control target/probe pairs was markedly superior to performance on the same pairs when they were embedded within different pitch time patterns. In part, this difference highlights the fact that when a tone becomes part of a larger sequence, both local (target cell) and global (context cells) pitch differences affect how a listener orients to it. Multiple serial relationships between pitches, as well as the relative timing of tones, complicate selective listening in such tasks (Bregman & Rudnicky, 1975; Johnston & Jones, 2006; Jones et al., 1981; van Noorden, 1975). In light of this, it is noteworthy that the performance which most closely approximates performance levels with control pairs is found with rhythmically expected probes in the most simply structured auditory patterns, those comprising many small pitch intervals (i.e., the Narrow/Small condition). These patterns project a compelling pitch structure: many small, regularly arranged, pitch intervals convey a smooth undulating motion in pitch space.

We assessed hypotheses inspired by five different theoretical perspectives. Three of these (Large & Jones, 1999; Pitch/Time Entrainment, Pitch Space) provide some role for sequence time structure in this task. The emphasis on global (as well as local) sequence timing is prominent in the two entrainment models, which propose that a pattern's rhythm guides anticipatory attending, hence temporal expectancies. One, the Large and Jones (1999) model, assumes that anticipatory attending is induced by rhythmic patterns of intensity and time changes and that reactive attending depends only on local time deviations. The other, the Pitch/Time Entrainment model, posits a role for pitch as well as time structure in both anticipatory attending (global pitch and time structure) and reactive attending (local pitch and time structure). By contrast, the Pitch Space model emphasizes local pitch and time structure. In this view, a global rhythm, rather than eliciting temporal expectancies, lays a foundation for a deviant local time interval (i.e., an ill-timed probe) to automatically capture attention. Selective attention to a particular probe tone is guided by task instructions, but not by pattern structure. Thus, instructions are proposed to activate a top-down voluntary mechanism (e.g., as in feature search mode) that orients attending to a location in pitch space, setting the stage (locally) for attentional capture by a temporally unexpected probe. Finally, the remaining theories concentrate more heavily on the role of pitch than of time structure, especially pitch proximities (I-R theories and Scene Analysis).

No single theoretical view explains all aspects of performance in the three experiments we report, but some do better than others. We discuss the pros and cons of these approaches by considering the following topics: pitch structure, time structure, and voluntary factors in selective attending to patterned sequences.

5.1. Pattern Pitch structure

Because we have orthogonally varied global (context) and local (target distance) pitch structure in these experiments, we can conclude that global pitch structure had a robust effect on performance. Significant main effects of this variable appeared in all three experiments. Listeners' accuracy in identifying recurrences of a target in the penultimate serial position was significantly better when context segments of a sequence comprise relatively small (Narrow) versus large (Wide) pitch intervals. To our knowledge, the finding that early portions of a pitch pattern exert a persisting influence on selective attending to remote sequence tones (probes 3–6 tones removed), has not been previously reported. Our findings indicate that certain initial pitch relationships orient attention better than others for the future "what" and "when" of an unfolding sequence. In this view, pitch and time relationships in the early portions of a recurrent pattern elicit dynamic attunements that project forward in time as anticipatory attending. These findings are in singular accord with the Pitch/Time Entrainment description of anticipatory attending and a variable temporal focus of attending.

Local pitch structure, by contrast, exerts a much smaller influence on performance than global pitch structure in the first two experiments. And, although the influence of global pitch structure is consistent with the Pitch/Time Entrainment model, the null influence of local pitch structure is not. According to a Pitch/Time entrainment view, smaller pitch intervals surrounding target/probes (e.g., as in the target distance variable) should contribute to anticipatory attending and to narrowing the pitch–time focus. This implies that pitch proximity will *lower* performance levels with ill-timed probes. This did not happen in Experiment 2 (or Experiment 3). Instead, relative differences in performance levels for on-time versus ill-timed probes did not differ as function of pitch structure (local or global). Such findings are more in line with predictions that the corrective response to expectancy violations (i.e., reactive attending) depends only on local time changes. That is, in an entrainment view such expectancy corrections should behave like simple phase corrections and hence are largely unaffected by the local pitch intervals (large or small) carried by a target/probe tone. In this respect, a strict entrainment model as outlined by Large and Jones (1999), finds support.

The pitch structure manipulations used in this research also allow for assessments of theories concerned with pitch relationships, particularly pitch proximity. The fact that smaller pitch intervals enhance performance in all three experiments agrees with I-R theories. However, it should be noted that we found these advantages of smaller pitch intervals to hold even when the patterns involved failed to embed the smaller pitch intervals in serial positions that elicit what I-R theory defines as *good closure* target/probes (cf. Section 2.1). Often in previous research the realized intervals rated as most effective by listeners have been ones with small pitch intervals that appear in pattern-final serial positions. This has been interpreted as supported for Implication–Realization theory. However, for the present set of experiments these theories imply that even when pitch proximity is pronounced, the target/probes in sequences used in Experiments 1–3 may function as unexpected pitch intervals. This is because in our patterns, the non-final target/probes end implicative, rather than realized, intervals. Nevertheless, we find that in all three experiments global pitch proximities benefit performance for these probe tones, at least when they are rhythmically expected. From this, we conclude that the construct of a pattern-based expectancy is not necessarily conditional upon closure status of a particular pitch intervals (i.e., open or closed in I-R terms), or upon serial locations (final or non-final) or certain tasks (ratings or pitch identification). Rather, we suggest that in auditory sequences an expectancy refers to a listener's anticipation about the future occurrence of an element in pitch space and time. Within this broader view, an important theoretical goal is to explain certain determinants of expectancies (probabilistic, structural) and the experience of these (as a continuation or a closure).

Scene Analysis shares with I-R theory an important role for pitch proximity. However, we find little support for its account of selective attending to slow pitch patterns. Unlike the fast patterns used by Mondor and Terrio (1998), the slow patterns used in the present research should elicit schema-driven attending.[8] According to Bregman (1990), selective attending operates more efficiently in sequences where a target/probe tone "stands out" from surrounding (pre-attentively grouped) tones. However, in the present task the relevant interactions of global with local context did not support this prediction. In no experiment did performance levels in the Narrow/Large condition significantly exceed those in the Narrow/Small. Although this critical comparison proved significant in Experiment 3, its direction was opposite that predicted by Scene Analysis: listeners performed *significantly worse* in the Narrow/Large condition than in the Narrow/Small.

An alternative possibility that does not figure prominently in any of the five target theories of this research should be mentioned. It appeals to a simple memory comparison algorithm. It is possible that people remember target tones better when they appear in certain pitch patterns and this accounts for observed differences in probe judgments due to pitch structure. However, a simple memory model (for target tones), which is based mainly on retrieval of a tone's memory trace and a related comparison with a probe tone, cannot predict the present findings unless it also specifies precisely how a memory code and its retrieval are systematically affected by instructions and by surrounding serial context (i.e., pitch and time structure) *in the absence of attention.* That is, in sequence monitoring, any memory account must engage working memory and a precise working memory explanation becomes difficult to distinguish from a dynamic attending one (cf. Barnes & Jones, 2000). A more sophisticated conventional memory model might address pitch context effects by assuming that pattern repetition strengthens working memory traces for a particular pitch structure which, in turn, primes the listener for a certain (similar) target intervals. However, this interpretation implies that retention should be best for serial patterns with congruent pitch structures (Narrow/Small, Wide/Large). But, this was not found. Finally, to our knowledge no current memory theory spells out predictions for our manipulations as precisely as do the five theories of central interest in this research.

In a Pitch/Time Entrainment approach, the dynamics of attentional entrainment are postulated to lead to more or less efficient attunements to a sequence (i.e., a driving rhythm) as a result of serial pitch and time relationships. Both the width of constituent pitch intervals and the coherence of successive time intervals contribute to the precision of attentional targeting in real time, hence to anticipatory attending and related expectancies. The fact that global pitch structure has a marked effect on performance is consistent with this view. Without specific instructions, it appears that people implicitly weigh opening portions of a sequence relatively heavily. We conjecture that in early portions of a recurrent sequence, the initial period and phase adjustments of one (or more) entraining oscillators result in a pacing of attending which, in turn, plays an elementary role in implicit learning of pattern relationships. Probes that "fit' an anticipated pitch/time trajectory are more likely to be accurately identified (hence, serially integrated, i.e., learned) when they fall within focus of attending associated with that trajectory. A difficulty with this explanation is that the Pitch/Time Entrainment approach also assumes that local pitch structure contributes to anticipatory attending such that smaller target distances further narrow the attentional focus set motion by context por-

---

[8] Bregman (1990) does not provide guidelines for differentiating fast rates from the slow ones proposed to guide schema-driven attending. However, conventionally fast rates are taken to be ones based on IOIs around 250 ms or less.

tions of a pattern. Although the data from Experiments 1 and 2, taken together, indicate that overall performance was best in the Narrow/Small condition and worst in the Wide/Large, these interactions fell short of significance (cf. footnote 7). It is this finding that is problematic for the Pitch/Time Entrainment account. However, to more fully address this and related predictions about local pitch structure, we must consider responding to ill-timed probes in more detail in the context of reactive attending. We consider this next.

5.2. Pattern time structure

Clearly, our findings suggest that the pitch of rhythmically expected tones is easier to identify than that of ill-timed probes. We propose that this finding is related to the width, as well as the temporal locus, of an attentional focus which develops during the monitoring of a temporally coherent (isochronous) pattern. Indeed, using simulations of oscillator entrainment, Large and Jones (1999) showed that an attentional focus narrows over time during entrainment to rhythmic patterns that contain few or no temporal variations. Accordingly in the present task, we assume that a listener, who is engaged in moment-to-moment monitoring of a rhythmically coherent pattern, is event-paced, and relies on periodically targeted attending energy that becomes increasingly concentrated about expected points in time; in turn, this improves performance with on-time probes. By contrast, in rhythms with less coherence, this focus widens. Thus, both entrainment models predict flatter expectancy profiles whenever a driving rhythm is incoherent, i.e., contains substantial timing irregularities. This entrainment approach applies to designs of Experiments 2 and 3 where sequence rhythms are formed by highly coherent intensity/time patterns (only probe time changes sometimes). Consequently, application of the original Large and Jones (1999) entrainment model to a pitch judgment task provides the best explanation of the main effect of probe timing in Experiment 2; it correctly predicts that, regardless of local pitch structure, a narrow attentional focus will develop leading to a relatively sharp expectancy profile with best performance for on-time probes. In this experiment, where listeners were not forewarned about probe timing irregularities, robust quadratic trends occurred over levels of probe time in all four pitch structure conditions. In sum, we found that the strongest and most consistent finding in this research was associated with the probe timing effect.

The fact that quadratic trends associated with probe timing did not differ as a function of pitch structure in Experiment 2 is important. It is consistent with a strict entrainment interpretation of reactive attending. Performance levels are relatively low to ill-timed probes regardless of local pitch relationships between a target/probe tone and surrounding pitches. Therefore, in Experiment 2, we can infer that reactive attending is generally inefficient, i.e., constrained in time, and unaffected by local pitch proximities. Had reactive attending been effective in shifting the focus of attention to early or late occurrences of the target/probe, a capture effect would be apparent in this experiment, indicating superior performance with ill-timed probes. A Pitch Space model predicts such a capture effect for pitch-proximal targets/probes, but this prediction was not supported in Experiment 2 (or Experiment 3).

In summary, when listeners are not specifically instructed about pitch and time, they rely heavily on global pitch structure and rhythm. As a result they show clear expectancy profiles featuring relatively good performance for patterns with Narrow contexts and for probes that are rhythmically expected along with relatively poor performance with ill-timed probes. Finally, local pitch structure has little impact on levels of responding to early versus late probes, indicating that reactive attending is generally unaffected by the presence of pitch proximities.

5.3. Limits on the voluntary control of attending?

A final issue concerns the determinants of voluntary versus automatic control of attending in this task. Typically, voluntary control is associated with a top-down, deliberate, goal-directed search for target-relevant features (e.g., in visual attention), whereas automatic control entails a reactive shift of attending to a sudden or unexpected stimulus. In entrainment theories, where the serial structure of a stimulus pattern figures prominently in definitions of a driving rhythm, reactive attending is most likely to correspond to an auditory/temporal correlate of an automatic, i.e., stimulus-driven, attending activity. However, our grounding of reactive attending in the time domain, namely as a *temporal* shift of attending, distinguishes it from more conventional construals of stimulus-driven attending that involve spatial shifts (Jones, 2001). By contrast, the degree to which anticipatory attending fulfills the conventional role of a purely voluntary activity is more debatable. This is because entrainment accounts have assumed that anticipatory attending is also pattern driven, namely that it depends upon periodically recurrent structural relationships within an unfolding sequence. In this role, we speculate that anticipatory attending contributes to implicit learning of pitch/time serial relations (Reber, 1989). However, the degree to which anticipatory attending is susceptible to explicit voluntary control remains an open question.

In Experiment 3, we tackled this question. We considered whether instructions might differentially affect anticipatory (versus reactive) attending. In contrast to entrainment theories, other approaches (e.g., Pitch Space theories, Scene Analysis) do not link expectancies directly to pattern structure. Rather it is more common to link voluntary attending to instructions, task goals and/or domain-specific learning. Indeed, an advantage of Pitch Space theories is that they provide a basis for predicting that instructions (in Experiment 3) will explicitly focus attention on pitch (to the exclusion of time). In this case, if instructions elicit goal-directed attending, then the pitch of tones neighboring the target/probe tone in pitch space (i.e., the target distance variable) becomes task relevant. Indeed, the fact that target distance does affect performance in Experiment 3 but not in Experiment 2 testifies to the power of such instructions in this respect.

Experiment 3 juxtaposed predictions from an extreme version of pattern-driven attending, which assumes the dominance of pattern structure over instructions, with an equally extreme version of the top-down approach, which assumes dominance of instructions over pattern structure. By introducing explicit instructions about the relevance of probe pitch and the irrelevance of probe timing, we asked: “Can instructions, in some interesting way, over-ride the influence of pattern structure on performance?”

It turns out they can, but in an unexpectedly interesting way. Rather than *generally* facilitating performance in a pitch-relevant task, instructions had a *selective* influence on performance. Performance levels improved only for patterns with small target distances in Experiment 3. Follow up analyses in this study, based on quadratic trend predictions of the probe time variable, revealed that this improvement was primarily confined to performance with on-time probes in patterns with small target distance intervals. In fact, performance with on-time targets embedded in patterns with large target distances declined noticeably relative to the levels observed in Experiment 2. As well, it is significant that instructions had no obvious effect on performance with ill-timed probes. Taken together, these results suggest that instructions differentiated anticipatory from reactive attending. They also indicate that anticipatory attending is susceptible to limited voluntary control whereas reactive attending is not.

We now speculate on theoretical implications of these findings. The puzzle we address is this: the Pitch/Time Entrainment view does not incorporate a voluntary component, yet it fares better than other approaches in explaining Experiment 3 results! Why? We think the answer lies in identifying determinants of attentional focus width. The Large and Jones model fails in Experiment 3 because it predicts that focus width varies only as a function of sequence timing, not as a function of either pitch structure or instructions. The Pitch Space approach fails in Experiments 2 and 3 because, in spite of its assumption that instructions orient an attentional focus in pitch space, its focus is neither temporal in nature nor variable in width. For these reasons, neither the Large and Jones model nor the Pitch Space model predict the changes in expectancy profiles observed in Experiment 3 (versus 2). The Pitch/Time Entrainment model is the most promising of the five target theories examined here, but it faces two problems. First, unlike the Large and Jones model, it assumes that pitch structure can affect reactive attending. Second, unlike the Pitch Space model, it does not feature a role for instructions. We offer a modification of the Pitch/Time Entrainment approach that addresses both problems.

We begin by assuming, first, that only anticipatory attending is affected by pitch structure and instructions. Reactive attending, following Large and Jones (1999), depends only on local time structure (i.e., target cell). Next, we assume that instructions can influence anticipatory attending, but not reactive attending. Anticipatory attending is susceptible to voluntary control because it is an activity associated with a future goal point (in time) whereas reactive attending is not goal-directed. Note that in practice voluntary control usually refers to the fact that people follow instructions by relying upon a goal-oriented expectancy that is *narrowly focused* on task-relevant information. In the present situation, we provided task-relevant information by telling listeners that their goal was to attend to a given tone's pitch, as this relevant, and not to time, as this was irrelevant (Experiment 3). Accordingly, we argue that as a sequence unfolds, such explicit information facilitates the goal of tracking a pitch target/probe, but that this goal is most readily accomplished with the on-time probes that are embedded in patterns with small pitch intervals. Thus, we claim that instructions 'correct' a listener's tendency to attentionally weight early portions of a sequence relatively heavily in sequence monitoring in rhythmically paced attending. Without such explicit instructions, global pitch/time relationships within recurrent cycles of a sequence mainly shape anticipatory attending. However, with explicit instructions, where the impact of local pitch/time relationships is more evident, it is possible to infer that people heighten attending to the temporal location of an anticipated target/probe thus weighting target distance more heavily than otherwise. Theoretically, this implies that the width of a listener's attentional focus, carried by anticipatory attending, becomes more sensitive to local pitch structure under the influence of instructions; small target distances tend to narrow a listener's focus whereas large target distances tend to increase focus width. These two modifications of the Pitch/Time Entrainment model clarify the role of instructions in selectively mediating the width of an attentional focus as function of local pitch structure.

In conclusion, the simplest story is this: when listeners normally listen to auditory events, they are implicitly driven, more or less efficiently, by global pattern relationships (global pitch and time structure) which induce paced anticipations about expected elements. Pattern structure tacitly dominates performance in this case. However, if listeners are explicitly encouraged to "pay attention' to the pitch of a particular tone within such events (but not its timing), they comply by selectively adjusting attending in anticipation of a future pitch/time point that is associated with this goal; moreover, they are most effective in achieving this goal with rhythmically expected probes embedded in relatively simple pitch patterns (e.g., Narrow/Small). But it is important to observe that a listener's attending in real time remains strongly influenced by a pattern's pitch and time structure in both situations, although in the latter it is clearly contingent, as well, upon explicit instructions. We infer that anticipatory attending to rhythmically expected tones is influenced by instructions as well as pitch/time structure, whereas reactive attending to rhythmically unexpected tones is influenced only by local time structure. In sum, voluntary factors combine with aspects of a pattern's pitch structure to directly affect anticipatory attending but they do not alter reactive attending. We conclude that dynamic attending in sequence monitoring can be selectively shaped by instructions that alert listeners to critical aspects of future targets.

Appendix A

See Tables A1 and A2.

Table A1
Mean d' values for all three experiments

		Small	Large	
Experiment 1. d' Values and Standard errors in parentheses				
	Narrow	1.49 (.16)	1.50 (.17)	1.50 (.21)
	Wide	1.29 (.15)	0.96 (.16)	1.13 (.19)
		1.40 (.19)	1.23 (.22)	
Experiment 2. d' Values and Standard errors in parentheses				
Early	Narrow	0.73 (.18)	0.82 (.20)	0.78 (.22)
	Wide	0.74 (.18)	0.66 (.14)	0.70 (.13)
		0.73 (.20)	0.74 (.17)	0.74 (.23)
On time	Narrow	1.36 (.18)	1.51 (.19)	1.43 (.19)
	Wide	1.22 (.19)	1.06 (.15)	1.14 (.20)
		1.29 (.20)	1.29 (.19)	1.29 (.26)
Late	Narrow	1.13 (.17)	1.17 (.19)	1.15 (.23)
	Wide	0.90 (.18)	0.78 (.16)	0.84 (.22)
		1.01 (.21)	0.97 (.23)	0.99 (.29)
Experiment 3. d' Values and Standard errors in parentheses				
Early	Narrow	1.04 (.18)	1.09 (.19)	1.06 (.24)
	Wide	0.92 (.17)	0.53 (.13)	0.73 (.20)
		0.98 (.24)	0.81 (.22)	0.89 (.30)
On time	Narrow	1.71 (.17)	1.03 (.18)	1.37 (.21)
	Wide	1.46 (.18)	0.95 (.14)	1.21 (.19)
		1.59 (.23)	0.99 (.20)	1.29 (.26)
Late	Narrow	1.35 (.17)	1.16 (.18)	1.26 (.21)
	Wide	1.32 (.17)	0.96 (.16)	1.14 (.18)
		1.33 (.21)	1.06 (.16)	1.20 (.24)

Table A2
Hit and False Alarm rates for all three experiments

		Small	Large	
Experiment 1. Hit Rates and False Alarm rates in parentheses				
	Narrow	.72 (.23)	.59 (.15)	.65 (.19)
	Wide	.68 (.25)	.60 (.28)	.64 (.27)
		.70 (.24)	.60 (.22)	.645 (.23)
Experiment 2. Hit Rates and False Alarm rates in parentheses				
Early	Narrow	.48 (.25)	.40 (.16)	.44 (.20)
	Wide	.41 (.20)	.44 (.23)	.43 (.21)
		.44 (.22)	.42 (.20)	.43 (.215)
On time	Narrow	.84 (.37)	.77 (.27)	.81 (.32)
	Wide	.79 (.36)	.74 (.37)	.77 (.36)
		.81 (.36)	.76 (.32)	.79 (.34)
Late	Narrow	.57 (.20)	.52 (.15)	.54 (.18)
	Wide	.50 (.21)	.51 (.25)	.50 (.23)
		.53 (.21)	.51 (.20)	.52 (.205)
Experiment 3. Hit Rates and False Alarm rates in parentheses				
Early	Narrow	.66 (.31)	.54 (.20)	.60 (.26)
	Wide	.57 (.27)	.51 (.34)	.54 (.31)
		.61 (.29)	.53 (.27)	.57 (.28)
On time	Narrow	.81 (.24)	.61 (.26)	.71 (.25)
	Wide	.73 (.25)	.67 (.34)	.70 (.30)
		.77 (.24)	.64 (.30)	.71 (.27)
Late	Narrow	.64 (.21)	.57 (.18)	.61 (.20)
	Wide	.64 (.21)	.56 (.25)	.60 (.23)
		.64 (.21)	.57 (.22)	.60 (.21)

References

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology, 41*, 254–311.

Besson, M., & Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance, 21*(6), 1278–1296.

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other nonlinguistic contexts. *Psychophysiology, 24*, 14–25.

Boltz, M. (1989). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 749–761.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge: MIT Press.

Bregman, A. S., & Rudnicky, A. I. (1975). Auditory segregation: Stream or streams? *Journal of Experimental Psychology: Human Perception and Performance, 1*(3), 263–267.

Cherry, E. C. (1957). *On human communication: A review, a survey, and a criticism*. New York/Cambridge: John Wiley/Technology Press, MIT.

Cuddy, L. L., & Lunney, C. A. (1995). Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity. *Perception & Psychophysics, 57*, 451–462.

Dowling, W. J., Lung, K. M., & Herrbold, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics, 41*(6), 642–656.

Egeth, H., & Yantis, S. (1997). Visual attention. *Annual Review of Psychology, 48*, 269–297.

Garner, W., & Gottwald, R. L. (1968). The perception and learning of temporal patterns. *Quarterly Journal of Experimental Psychology, 20*, 97–109.

Goschke, T. (1998). Implicit learning of perceptual and motor sequences. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning*. Thousand Oaks, California: Sage Publications.

Greenberg, G. Z., & Larkin, W. D. (1968). Frequency-response characteristic of auditory observers detecting signal of a single frequency in noise: The probe-signal method. *Journal of the Acoustical Society of America, 44*, 1513–1523.

Hoffman, J., & Koch, I. (1998). Implicit learning of loosely defined structures. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 161–199). Thousand Oaks, CA: Sage.

Johnston, H. M., & Jones, M. R. (2006). Higher order pattern structure influences auditory representational momentum. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 2–17.

Jones, M. R. (1974). Cognitive representations of serial patterns. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 187–229). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Jones, M. R. (1981). A tutorial on some issues and methods in serial pattern research. *Perception & Psychophysics, 30*, 492–504.

Jones, M. R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology, 9*, 193–228.

Jones, M. R. (2001). Temporal expectancies, capture and timing in auditory sequences. In C. F. B. Gibson (Ed.), *Attraction, distraction, and action: Multiple perspectives on attentional capture*. Elsevier Science, B.V.

Jones, M. R. (2004). Attention and timing. In J. A. P. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 49–85). Academic Press.

Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance, 7*(5), 1059–1073.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*, 313–319.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics, 43*, 346–354.

Klein, J. M., & Jones, M. R. (1996). Effects of attentional set and rhythmic complexity on attending. *Perception & Psychophysics, 58*, 34–46.

Koch, I., & Hoffman, J. (2000). The role of stimulus-based and response-based spatial information in sequence learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 863–882.

Lamy, D., Leber, A., & Egeth, E. (2004). The effects of task relevance and stimulus-driven salience in feature-search mode. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 1019–1031.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119–159.

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 1102–1125.

Mondor, T. A., & Bregman, A. S. (1994). Allocating attention to frequency regions. *Perception & Psychophysics, 56*, 268–276.

Mondor, T. A., & Terrio, N. A. (1998). Mechanisms of perceptual organization and auditory selelctive attention: The role of pattern structure. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1628–1641.

Moray, N. (1969a). *Listening and attention*. Baltimore: Penguin Books.

Moray, N., Fitter, M., Ostry, D., Favreau, D., & Nagy, V. (1976). Attention to pure tones. *Quarterly Journal of Experimental Psychology, 28*(2), 271–283.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication–realization model*. Chicago: University of Chicago Press.

Narmour, E. (1991). The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception, 9*, 1–26.

Narmour, E. (1992). *The analysis and cognition of melodic complexity: The implication–realization model*. Chicago: University of Chicago.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*, 1–32.

Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review, 77*, 481–495.

Restle, F. (1972). Serial patterns: The role of phrasing. *Journal of Experimental Psychology, 92*, 385–390.

Restle, F., & Brown, E. R. (1970). Organization of serial pattern learning. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 4, pp. 249–331). New York: Academic Press.

Scharf, B. (1998). Auditory attention: The psychoacoustical approach. In H. Pashler (Ed.), *Attention*. Hove, East Sussex, UK: Psychology Press.

Scharf, B., Quigley, S., Aoki, C., Peachey, N., & Reeves, A. (1987). Focused auditory attention and frequency sensitivity. *Perception & Psychophysics, 42*, 215–223.

Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication–realization model. *Cognition, 58*(1), 75–125.

Schellenberg, E. G., Adachi, M., Purdy, K. T., & McKinnon, M. C. (2002). Expectancy in melody: Tests of children and adults. *Journal of Experimental Psychology-General, 131*(4), 511–537.

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception, 7*, 109–150.

Shin, J. C., & Ivry, R. B. (2000). Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 445–457.

Snyder, J. S., & Large, E. W. (2004). Tempo dependence of middle- and long-latency auditory responses: Power and phase modulation of the EEG at multiple time-scales. *Clinical Neurophysiology, 115*(8), 1885–1895.

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

Thompson, W. F., Cuddy, L. L., & Plaus, C. (1997). Expectancies generated by melodic intervals: Evaluation of principles of melodic implication in a melody-completion task. *Perception & Psychophysics, 59*(7), 1069–1076.

Todd, R., Boltz, M., & Jones, M. R. (1989). The MIDILAB auditory research system. *Psychomusicology, 8*, 17–30.

Unyk, A. M., & Carlsen, J. C. (1987). The influence of expectancy on melodic perception. *Psychomusicology, 7*(1), 3–23.

van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Eindhoven: University of Technology.

Ward, L. M., & Mori, S. (1996). Attention cuing aids in auditory intensity resolution. *The Journal of the Acoustical Society of America, 100*, 1722–1727.

Wearden, J. H., & Bray, S. (2001). Scalar timing without reference memory? Episodic temporal generalization and bisection in humans. *Quarterly Journal of Experimental Psychology Section B—Comparative and Physiological Psychology, 54*(4), 289–309.

Winfree, A. T. (2000). *The geometry of biological time* (2nd ed., Vol. 12). New York: Springer-Verlag.

Wright, B. A., & Fitzgerald, M. B. (2004). The time course of attention in a simple auditory detection task. *Perception & Psychophysics, 66*, 508–516.

Yantis, S., & Egeth, H. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 661–676.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 601–621.