ELSEVIER

Auditory attention to frequency and time: an analogy to visual local–global stimuli

Timothy Justus[*], Alexandra List

University of California, Berkeley CA, USA

Abstract

Two priming experiments demonstrated exogenous attentional persistence to the fundamental auditory dimensions of frequency (Experiment 1) and time (Experiment 2). In a divided-attention task, participants responded to an independent dimension, the identification of three-tone sequence patterns, for both prime and probe stimuli. The stimuli were specifically designed to parallel the local–global hierarchical letter stimuli of [Navon D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383] and the task was designed to parallel subsequent work in visual attention using Navon stimuli [Robertson, L. C. (1996). Attentional persistence for features of hierarchical patterns. *Journal of Experimental Psychology: General*, *125*, 227–249; Ward, L. M. (1982). Determinants of attention to local and global features of visual forms. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 562–581]. The results are discussed in terms of previous work in auditory attention and previous approaches to auditory local–global processing.

© 2004 Published by Elsevier B.V.

Keywords: Audition; Attention; Priming; Global; Local; Hierarchical; Frequency; Time; Temporal

* Corresponding author. Department of Neurology (127E), University of California/VANCHCS, 150 Muir Road, Research Building 4, Martinez CA, 94553-4668 USA.

E-mail address: tjustus@ebire.org (T. Justus).

1. Introduction

1.1. Attention to local–global structure and spatial frequencies in vision

Information is carried at multiple hierarchical levels in the visual world. These different levels of information may be more or less relevant at different times and under different circumstances. For instance, when observing a landscape, the viewer may wish to examine the details of one figure or form, or consider how these elements sum within the larger scene. Over the past 25 years, a simple but elegant set of stimuli has been used in a variety of research to test hypotheses about visual hierarchical processing. These are the local–global letter stimuli of Navon (1977), which consist of one large global letter made up of smaller local letters (Fig. 1A).

The Navon stimuli have been used in cognitive psychological research to test hypotheses about the way in which the visual system handles information occurring at multiple hierarchical levels in a visual scene, including how these levels are perceived and how attention can be directed to them (e.g. Kim, Ivry, & Robertson, 1999; Lamb & Robertson, 1988; Martin, 1979; Navon, 1977; Robertson, 1996; Sergent, 1982; Shulman & Wilson, 1987; Ward, 1982). The stimuli have also been used in neuropsychological research to examine how the brain implements local–global processing in an asymmetric manner, with the left and right hemispheres more attuned to local and global levels, respectively (e.g. Delis, Robertson, & Efron, 1986; Lamb, Robertson, & Knight, 1989; Robertson & Delis, 1986; Robertson, Lamb, & Knight, 1988). Finally, the stimuli have been used in developmental work to investigate the emergence of local–global processing in normally developing children (e.g. Dukette & Stiles, 1996) and in abnormally developing populations, such as children with Williams Syndrome (e.g. Bihrle, Bellugi, Delis, & Marks, 1989; Farran, Jarrold, & Gathercole, 2003; Pani, Mervis, & Robinson, 1999) and autism (Mottron, Burack, Stauder, & Robaey, 1999; Mottron, Burack, Iarocci, Belleville, & Enns, 2003; Plaisted, Swettenham, & Rees, 1999).

One set of experiments by Robertson (1996) (also see Ward, 1982) attempted to explore the cognitive processes involved in identifying items at local and global levels (Fig. 1A). For each trial, a Navon hierarchical letter stimulus was displayed and the participants' task was to identify which of two target letters (e.g. A and E) was present in the stimulus. Each stimulus contained one target letter (e.g. A) and one distracter letter (e.g. H), and the target letters occurred unpredictably at either the local or global level. Robertson (1996, Experiment 2) found that response times were faster when the target level, local or global, was the same as that of the preceding trial (same level), compared to when the level changed (different level). This occurred even when the target pattern (A or E) changed, suggesting that there was an "attentional persistence" to the level (local or global), and not the form, of the preceding stimulus.

In a subsequent experiment, which explored how levels might be primed independently of perceptual form, level priming failed to occur when primes were constructed with reduced low spatial frequency information (Experiment 3). This supported the idea that priming to local and global levels is normally mediated by a mechanism involving representation of spatial frequency, with local and global

information carried in relatively high and low spatial frequencies, respectively. By eliminating one set of frequency information in the prime, the local–global structure of the prime stimulus could not be parsed into two sets of relatively high- and low-frequency information. Interestingly, not all stimulus dimensions produced a priming effect in these experiments; participants were not any faster when the prime and probe were similar in color, polarity, or contrast. Critical for the spatial frequency hypothesis, the manipulation of these other arguably fundamental dimensions did not account for the absence of the level priming effect when low-frequency information was filtered (Experiment 4).

Our goal in the current studies was to design a set of auditory stimuli that could be used in a paradigm as analogous as possible to that of Robertson (1996), and which could be later used to test a variety of questions about perception and attention across the two modalities of vision and audition. To do this, we first examined what the candidate dimensions were within the auditory system and whether they demonstrated hemispheric asymmetries similar to those found with spatial frequencies in vision.

1.2. Frequency and time as the "indispensable attributes" of audition

Arguably, the most fundamental feature to which the auditory system is attuned is frequency. Several researchers have developed the idea that frequency is the auditory analogue of space in vision, in part because of the way in which sensory transduction occurs; whereas the retina and primary visual cortex are organized spatiotopically, the cochlea and primary auditory cortex are organized tonotopically. One of the most thorough treatments of this analogy from an information processing perspective was provided by Kubovy and Van Valkenburg (2001), who argued that frequency, along with time, are the two "indispensable attributes" of audition, whereas space and time are the two indispensable attributes of vision. These authors referred to a stimulus attribute as indispensable if by distributing elements over the dimension, multiple perceptual objects are perceived. In support of the *Theory of Indispensable Attributes (TIA)*, Kubovy and Van Valkenburg appealed to work in auditory perceptual grouping which has demonstrated effects in frequency–time that are analogous to the visual Gestalt principles operating in space–time (e.g. Bregman, 1990; Handel, 1989). Further, they introduced additional parallels between space–time and frequency–time relating to figure-ground segregation and edge assignment that strengthen the case for a notion of auditory "objecthood" that is defined by frequency–time.

The special roles of frequency and time as formulated by Kubovy and Van Valkenburg (see also Kubovy, 1981) are not without controversy (e.g. Kubovy, 1988; Handel, 1988; Neuhoff, 2003; Van Valkenburg & Kubovy, 2003). This is partly because in addition to asserting positive claims about frequency and time, *TIA* also makes certain negative claims about the role of space in audition and auditory attention (to which we return in the General Discussion). Nonetheless, all of these authors would likely agree that frequency and time are among the most important organizing dimensions in the auditory modality, if not exclusively so, and that how attention may be directed to auditory frequency and time is a topic that merits further research.
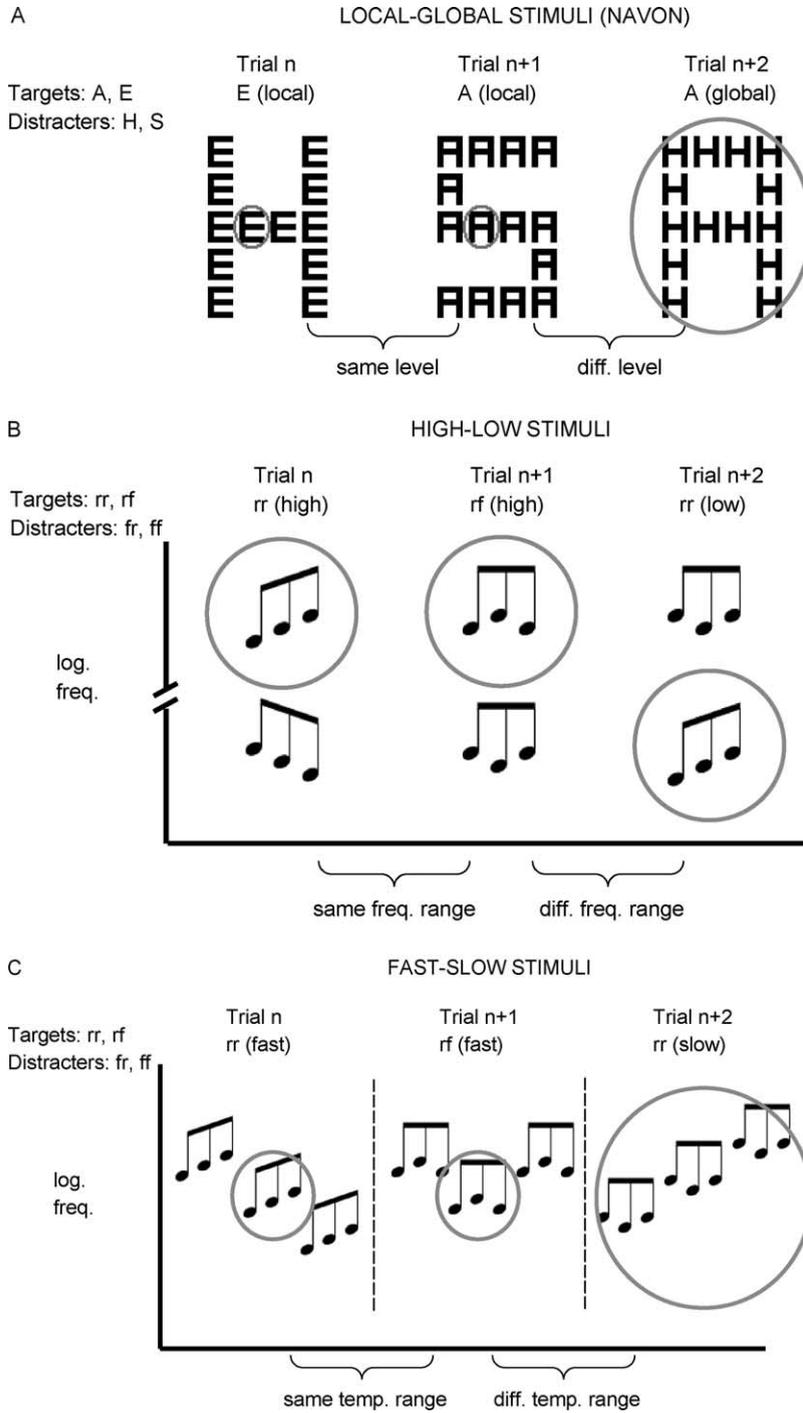
T. Justus, A. List / Cognition xx (2005) 1–21

A   LOCAL-GLOBAL STIMULI (NAVON)

Targets: A, E   Trial n   Trial n+1   Trial n+2
Distracters: H, S   E (local)   A (local)   A (global)

same level   diff. level

B   HIGH-LOW STIMULI

Targets: rr, rf   Trial n   Trial n+1   Trial n+2
Distracters: fr, ff   rr (high)   rf (high)   rr (low)

log.
freq.

same freq. range   diff. freq. range

C   FAST-SLOW STIMULI

Targets: rr, rf   Trial n   Trial n+1   Trial n+2
Distracters: fr, ff   rr (fast)   rf (fast)   rr (slow)

log.
freq.

same temp. range   diff. temp. range

Fig. 1. *Priming design for Navon stimuli and two auditory analogues*. (A) Navon stimuli are constructed such that a global letter is formed from local letters. In a priming design involving Navon stimuli, two letters (e.g. A, E) serve as targets while two other letters (H, S) serve as distracters. In trial $n+1$, the target (indicated with a gray circle) is presented at the same level as in trial n (same level transition), whereas in trial $n+2$, the target is not presented at the same level as in trial $n+1$ (different level transition). Participants are faster to identify the target when it is presented at the same level as in the preceding trial (Robertson, 1996). (B) High–Low stimuli are constructed such that a three-tone pattern in a relatively low auditory frequency range is presented with another pattern in a relatively high auditory frequency range. In a priming design involving High–Low stimuli, two patterns (e.g. rising–rising, rising–falling) serve as targets while two other letters (falling–rising, falling–falling) serve as distracters. In trial $n+1$, the target is presented in the same frequency range as in trial n, whereas in trial $n+2$, the target is not presented in the same frequency range as in trial $n+1$. Participants are faster to identify the target when it is presented in the same frequency range as in the preceding trial. (C) Fast–Slow stimuli are constructed such that a slow pattern covering a relatively wide temporal window contains three fast patterns each covering a relatively narrow temporal window. The target and distracter patterns are the same as for the High–Low stimuli. In trial $n+1$, the target is presented in the same temporal range as in trial n, whereas in trial $n+2$, the target is not presented in the same temporal range as in trial $n+1$. Participants are faster to identify the target when it is presented in the same temporal range as in the preceding trial.

1.3. The lateralization of frequency and time

In addition to being among the more fundamental dimensions of the auditory modality, evidence from neuroimaging and neuropsychology suggests that like spatial frequency, auditory frequency and time may show the requisite hemispheric asymmetries that long have motivated theories of local–global or analytic-holistic processing (e.g. Bever & Chiarello, 1974).

The hemispheric asymmetry within the dimension of frequency was proposed by Ivry and Robertson (1998) in their *Double Filtering by Frequency* (DFF) theory (also see Ivry & Lebby, 1993). These authors argued that many of the asymmetries in both audition and vision result from a preferred role for the left and right hemispheres in processing relatively high and low frequencies, respectively, in the senses of both pitch and spatial frequency. The hemispheric asymmetry within the dimension of time has been developed by many authors, including Ackermann and Riecker (2004), Allard and Scott (1975), Deacon (1997), Ivry and Robertson (1998, Chapter 6), Tallal (1980), and Zatorre, Belin, and Penhune, (2002). We shall refer to this idea as the *Asymmetric Sampling in Time* (AST) theory, the name ascribed to it by Poeppel (2003). Like the DFF theory, the AST theory also suggests that the left and right hemispheres have preferred roles in the processing of high- and low-frequency information, but at much larger time scales such that the relevant auditory dimension has shifted from one perceptual attribute (pitch) to another (time). Thus, the left and right hemispheres are argued to have advantages in processing events occurring at fast and slow temporal scales, respectively.[1]

The two theories, DFF and AST, are not necessarily mutually exclusive; after all, frequency is the inverse of time. The difference between the two emerges from

---

[1] Poeppel (2003) argues that the left and right hemispheres are preferentially tuned to fixed temporal windows of ∼20–40 and ∼150–250 ms, respectively. Thus his hypothesis differs from the DFF theory both in terms of the perceptual attribute (time rather than frequency), and also in that these values are absolute and not relative. We focus here on the distinction between the perceptual attributes.

the different positions on the frequency–time continuum occupied by the psychological experiences of pitch and temporal events. As discussed by Ivry and Robertson (1998), it is quite possible that both asymmetries could co-exist.

1.4. The current investigation: the high–low and fast–slow stimuli

Given that frequency and time are among the more fundamental auditory dimensions to which attention may potentially be directed, and given that both have been suggested to show a hemispheric asymmetry that is potentially analogous to that relating to spatial frequency in vision, we chose these two dimensions as the basis for two sets of stimuli created to parallel the Navon stimuli. In Experiment 1, we introduce the *High–Low stimuli*, designed to explore attentional persistence to auditory frequency range (Fig. 1B), and in Experiment 2 we introduce the *Fast–Slow stimuli*, designed to explore attentional persistence to auditory temporal range (Fig. 1C). These stimulus sets were both devised such that participants can attend to target patterns that may occur in either high or low frequency ranges and in either fast or slow temporal ranges, just as the letters in Navon stimuli may appear at either the local or global levels. To do this, we chose four three-tone sequences as analogues of the four letters. Further, like the letters these four sequences can be manipulated independently at the two levels, such that all possible target patterns at one level can be presented with all possible distracter patterns at the other level.

As a first experimental test of the stimuli, in this paper we demonstrate priming of frequency ranges and temporal ranges. The divided-attention priming task was chosen to be as analogous as possible to previous work using Navon stimuli (Robertson, 1996; Ward, 1982; see Fig. 1A–C). In addition to demonstrating two ways in which selection occurs within the auditory modality (frequency and time), the design of the current stimuli as experimental analogues of the Navon local–global visual stimuli suggests a variety of future studies in cognitive, neuro- and developmental psychology exploring whether parallels exist between the two modalities in the representation of spatial frequency, auditory frequency, and time.

2. Experiment 1

The first experiment was designed to demonstrate priming of auditory frequency ranges, using a divided-attention task parallel to that of Robertson (1996, Experiment 2). In this experiment, participants were instructed to identify three-tone target sequences, which could occur in one of two frequency ranges.

2.1. Method

2.1.1. Participants
Sixteen right-handed participants from the Berkeley community participated, with a median age of 20.5 years. The majority of the participants were musically trained, with the number of years of musical training ranging from 0 to 20. Participants were required to

identify at least 14 of 16 trials within four practice blocks before beginning the experiment, taking a mean of 2.2 blocks.

2.1.2. Stimuli and design

The stimuli for each trial in the High–Low version of the experiment consisted of two simultaneous three-tone sequences, each presented in a different frequency range (Fig. 1B). The lower and higher of the two sequences were presented in fundamental frequency ranges of 262–330 Hz (C_4–D_4–E_4) and 371–467 Hz (F#$_4$–G#$_4$–A#$_4$), respectively. Each tone was 150 ms in duration (with 10-ms ramps at onset and offset) and was constructed using five harmonics of inversely proportional amplitude. The rationale both for using frequency ranges separated by a dissonant interval (tritone) and for adding harmonics was to help the participants perceptually segregate the two sequences. The amplitude of both the lower and higher tones and their harmonics were adjusted to equate for loudness (i.e. the lower components were constructed with larger amplitudes). The stimuli were presented at an overall sound level of about 70 dB SPL.

Four possible three-note sequences could occur in either frequency range: a *rising–rising (rr)* pattern, *falling–falling (ff)* pattern, *rising–falling (rf)* pattern, or *falling–rising (fr)* pattern.[2] These stimuli are designed to parallel the Navon (1977) hierarchical letter stimuli: the four patterns are analogous to four letters, and the two auditory frequency ranges are analogous to the two visual hierarchical levels, each proposed to be perceived in terms of its relative spatial frequency information (compare Fig. 1A and B). For each participant, two of the four patterns served as targets—one same–direction sequence (*rr* or *ff*) and one changing-direction sequence (*rf* or *fr*)—and the two remaining patterns served as distracters. Each trial combined one target in one frequency range with one distracter in the other frequency range, and the participants' task was to identify as quickly and accurately as possible which of their two targets had been presented. The use of the four patterns as targets or distracters, and the mapping of the two targets to the left and right hands were counterbalanced across participants.

Pairs of consecutive trials were ordered such that they involved (1) the same or different target frequency range, (2) the same or different target pattern and (3) the same or different distracter pattern. All eight transitions were equally probable, occurring eight times in each of four 65-trial blocks. Each trial simultaneously served as prime and probe, with the first and final trials using identical stimuli to complete a cycle of 64 prime-probe transitions. The trial order in each of the four blocks was fixed, and the order in which the blocks were presented to the participants was counterbalanced.

Each trial began with visual fixation for 1000 ms, followed by the auditory stimulus (450 ms). Participants were given up to 5500 ms after the stimulus offset to make a response. Four trials timed out (for all participants combined) and were removed with the subsequent trial. Errors and the subsequent trial were excluded from response time analyses, as were trials exceeding three standard deviations of each participant's mean. Error rates were calculated as a proportion of non-excluded observations in each cell.

---

[2] *rr*: C–D–E or F#–G#–A#, *ff*: E–D–C or A#–G#–F#, *rf*: C–D–C or F#–G#–F#, *fr*: E–D–E or A#–G#–A#.

2.2. Results and discussion

Two $2 \times 2 \times 2 \times 2$ analyses of variance examined the effects of Frequency Range (same vs. different), Target (same vs. different), Target Frequency Range (high vs. low), and Target Pattern (same–direction: *rr/ff* vs. changing-direction: *rf/fr*)[3] on the participants' response times and error rates. The first two of these variables concern effects of the preceding trial, whereas the latter two concern effects of the target stimulus itself, independently of the preceding trial. Main effects of all of these variables were found. Participants were faster (mean 118 ms) and more accurate when the target was presented in the same frequency range as the preceding trial [RT: $F(1,15)=20.5$, $P<.001$; errors: $F(1,15)=9.6$, $P=.007$] and faster and more accurate when the target pattern *differed* from the preceding trial [RT: $F(1,15)=8.0$, $P=.01$; errors: $F(1,15)=9.2$, $P=.008$]. Participants were also generally faster and more accurate for targets presented in the high frequency range [RT: $F(1,15)=13.0$, $P=.003$; errors: $F(1,15)=5.7$, $P=.03$] and faster for the same–direction (*rr/ff*) target pattern [RT: $F(1,15)=6.0$, $P=.03$; errors: n.s.].

Additionally, the frequency-range priming effect was larger when the target also remained the same [Frequency Range $\times$ Target, RT: $F(1,15)=19.1$, $P=.001$; errors: n.s.]. Despite this interaction, the effect of frequency range was still significant for trials in which the target changed [same target trials, RT: $t(15)=5.7$, $P<.001$; errors: $t(15)=3.1$, $P=.007$; different target trials, RT: $t(15)=2.6$, $P=.02$; errors: $t(15)=2.4$, $P=.03$; Fig. 2A].

The priming effect for frequency range also held for both the high frequency range [RT: $t(15)=4.5$, $P<.001$; errors: $t(15)=4.0$, $P=.001$] and the low frequency range [RT: $t(15)=2.8$, $P=.01$; errors: $t(15)=1.6$, $P=.14$]. This interaction was marginally significant [RT: $F(1,15)=3.5$, $P=.08$; errors: $F(1,15)=4.9$, $P=.04$; Fig. 2B].[4]

We also considered the effects of absolute pitch, the ability to name a tone according to its chroma (e.g. C or do) in the absence of any named reference tone. Seven of the participants were within one semitone of the correct names on a simple identification test at the end of the experimental session. We defined the possession of absolute pitch somewhat generously in order to repeat the $2 \times 2 \times 2 \times 2$ ANOVA with absolute pitch as a between-subjects variable. No main effects of absolute pitch were found, and this variable did not interact with any of the other reported effects, including the priming effect for frequency range.

Experiment 1 demonstrated that when listeners have just identified a target pattern in a given frequency range, they are faster to identify another target pattern in that

---

[3] We grouped the Target Patterns in this way because previous work had suggested that series of three ascending tones or three descending tones are easier to identify than are other combinations of three notes (e.g. Divenyi & Hirsh, 1975, 1978). The main effect along this dimension supports a similar finding in our study. Subsequent comparisons confirmed that there were no significant differences between the two same–direction patterns (*rr* and *ff*) or between the two changing-direction patterns (*rf* and *fr*) in either experiment.

[4] In an analysis in which each response time was log transformed to correct the skewed distributions that are typical of response times, this interaction was significant [$F(1,15)=5.0$, $P=.04$]. The significance of all other effects, in both this experiment and in Experiment 2, was comparable in the analyses of the transformed data.
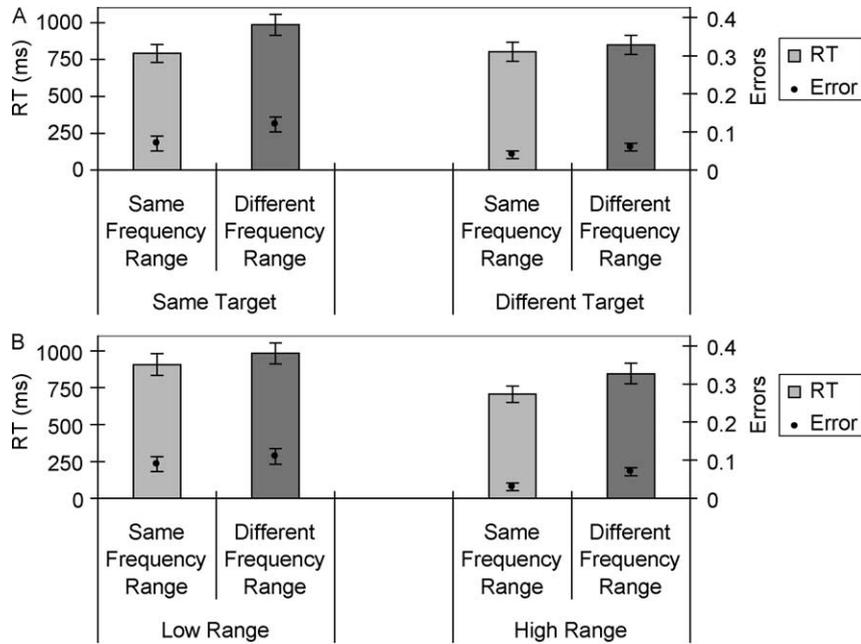
Fig. 2. *Results of Experiment 1 (High–Low Stimuli).* (A) Attending to a target in one frequency range facilitates attention to a subsequent target in the same range (lighter versus darker bars). This is true regardless of whether the target pattern remains the same (left) or changes (right). (B) The frequency priming effect also occurs for both the low range (left) and the high range (right). RTs (bars, left *y*-axis) are depicted in ms relative to stimulus offset. Errors (points, right *y*-axis) are given as proportions. Error bars represent standard error.

same frequency range. The pattern of effects is similar to that reported by Robertson, 1996, Experiment 2), using visual hierarchical letter stimuli in an analogous design. The most important result for our purposes is the main effect of frequency range; participants were (118 ms) faster to identify which of their two targets had been presented when the target was presented in the same frequency range as in the preceding trial. This effect occurred for both the high and low frequency ranges, and regardless of whether the specific target pattern changed or remained the same (Fig. 2). We shall return to the broader implications of this result after presenting the results of Experiment 2, which examined attention to the dimension of time using a precisely analogous method.

3. Experiment 2

The second experiment was designed to demonstrate priming of auditory temporal ranges, using a design parallel to that of Robertson (1996, Experiment 2) and the current Experiment 1. In this experiment, the three-tone target sequences were organized to create a hierarchical temporal structure.

3.1. Method

3.1.1. Participants

A separate group of 16 right-handed participants from the Berkeley community participated, with a median age of 19.5 years. The participants were musically trained, with the number of years of musical experience ranging from 3 to 19. Participants took a mean of 2.3 practice blocks before beginning the experiment.

3.1.2. Stimuli and design

The stimuli for each trial in the Fast–Slow version of the experiment consisted of nine-tone sequences. Each tone belonged to the frequency range 185–330 Hz ($F\#_3$-$G\#_3$-$A\#_3$-C_4-D_4-E_4-$F\#_4$-$G\#_4$-$A\#_4$). Each tone was 100 ms in duration and was constructed as in Experiment 1.

The same four sequence types as in Experiment 1 were used as targets and distracters (*rr*, *ff*, *rf*, *fr*). Rather than being presented simultaneously in two different frequency ranges, the two sequences within each stimulus were present at different temporal scales. The nine-tone sequence was constructed such that one three-item sequence was repeated three times to form another three-item sequence. For instance, in the first Fast–Slow sequence shown in Fig. 1C, the rising–rising pattern occurs three times in the faster temporal range (300 ms) and the falling–falling pattern occurs in the slower temporal range (900 ms). The remainder of the design was exactly the same as in Experiment 1, except that temporal range replaced frequency. Each trial combined one target in one temporal range with one distracter in the other temporal range, and the participants' task was to identify which of their two targets had been presented.

Six trials timed out after 4000 ms (for all participants combined). Errors, their subsequent trials, and outliers were treated as before.

3.2. Results and discussion

As in Experiment 1, two $2 \times 2 \times 2 \times 2$ analyses of variance examined the effects of Temporal Range (same vs. different), Target (same vs. different), Target Temporal Range (fast vs. slow), and Target Pattern (same–direction: *rr/ff* vs. changing-direction: *rf/fr*) on the participants' response times and error rates. The analysis is identical to that of Experiment 1, substituting temporal range for frequency range. Main effects of three of these variables were found. Participants were faster (mean 107 ms), but not significantly more accurate, when the target was presented in the same temporal range as the preceding trial [RT: $F(1,15) = 36.6$, $P < .001$; errors: n.s.]. Unlike Experiment 1, they were neither faster nor more accurate when the target pattern differed from the preceding trial [RT and errors: n.s.]. Participants were also generally more accurate but not significantly faster for the slow temporal range [RT: n.s.; errors: $F(1,15) = 6.3$, $P = .02$]. Finally, as in Experiment 1, participants were generally faster for the same–direction (*rr/ff*) target pattern [RT: $F(1,15) = 5.5$, $P = .03$; errors: n.s.].

As in Experiment 1, the temporal-range priming effect was larger when the target also remained the same [Temporal Range $\times$ Target, RT: $F(1,15) = 27.8$, $P < .001$; errors:
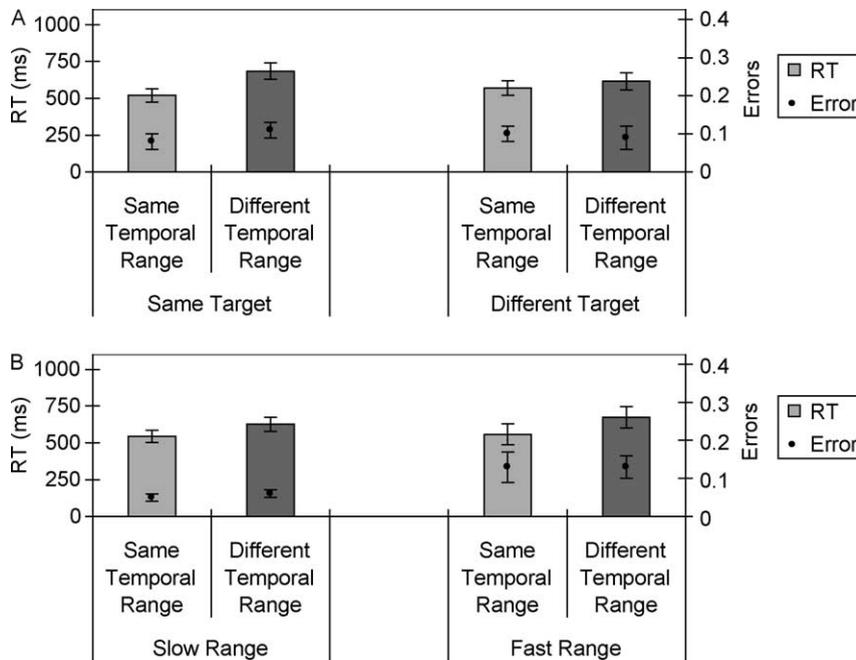
Fig. 3. *Results of Experiment 2 (Fast–Slow Stimuli).* (A) Attending to a target in one temporal range facilitates attention to a subsequent target in the same range (lighter versus darker bars). This is true regardless of whether the target pattern remains the same (left) or changes (right). (B) The temporal priming effect also occurs for both the slow range (left) and the fast range (right). RTs (bars, left *y*-axis) are depicted in ms relative to stimulus offset. Errors (points, right *y*-axis) are given as proportions. Error bars represent standard error.

$F(1,15)=4.4$, $P=.05$]. Despite this interaction, the effect of frequency range on response time was still significant for trials in which the target changed [same target trials, RT: $t(15)=5.9$, $P<.001$; errors: $t(15)=2.9$, $P=.01$; different target trials, RT: $t(15)=2.6$, $P=.02$; errors: n.s.; Fig. 3A].

The priming effect for temporal range also held for both the fast range [RT: $t(15)=4.0$, $P=.001$; errors: n.s.] and the slow range [RT: $t(15)=4.0$, $P=.001$; errors: n.s.]. The interaction was non-significant (Fig. 3B).

With regard to absolute pitch, four of the participants were within one semitone of the correct note names on the simple identification test at the end of the experimental session. When absolute pitch was added as a between-subjects variable to the $2\times2\times2\times2$ ANOVA, no main effects of absolute pitch were found, and this variable did not interact with any of the other reported effects, including the priming effect for temporal range.

Experiment 2 demonstrated that when listeners have just identified a target pattern in a given temporal range, they will be faster to identify another target pattern in that same temporal range. These patterns are similar to those of Robertson (1996, Experiment 2) and the current Experiment 1. Again, the most important result for our purposes is the main effect of temporal range; participants were (107 ms) faster to

identify which of their two targets had been presented when the target was presented in the same temporal range as in the preceding trial. This effect occurred for both the fast and slow temporal ranges, and regardless of whether the specific target pattern changed or remained the same (Fig. 3).

4. General discussion

4.1. Local–global, high–low, and fast–slow: similarities and differences

In this paper we have introduced two sets of stimuli that can serve as auditory analogues of the Navon (1977) local–global hierarchical stimuli so often used in visual perception research. Experiment 1 demonstrated a priming effect based on frequency range: attending to a target in one frequency range of the stimulus primes the listener for a target in the same frequency range on the subsequent trial. Experiment 2 demonstrated a priming effect based on temporal range: attending to a target in one temporal range of the stimulus primes the listener for a target in the same temporal range on the subsequent trial.

One of the strengths of these auditory stimuli is that, like the visual stimuli, participants are not explicitly directing attention to the high/low or fast/slow dimensions; they are identifying one of two three-tone target sequences, just as in the visual task participants are often identifying one of two target letters. A difference that emerged between the two current experiments was the effect of same versus different target. In Robertson's (1996, Experiment 2) hierarchical letter experiment and in the current Experiment 1, a main effect of target was found, such that participants were slower to respond to a target pattern that had been presented on the preceding trial (compare the two leftmost bars of Fig. 2A with the rightmost). This effect was not observed in Experiment 2 (Fig. 3A). One reason for this difference may have been the timing of the two experiments. Robertson (1996) found that the main effect of target identity (a slower response to a repeating target) decreased with longer ITIs. In our experiments, the mean trial lengths (given stimulus duration and mean RT) for the High–Low and Fast–Slow stimuli were 3805 and 3998 ms, respectively. If the two tasks were both associated with an inhibition of same target and/or response that diminished over time, the target effect would have been observed more readily in Experiment 1 than Experiment 2.

Another interesting difference between the three stimuli concerns the main effects of spatial frequency range, frequency range, and temporal range. Navon (1977) introduced the concept of *global precedence*, the idea that global information is processed before local information. This idea has been widely disputed, especially based on evidence challenging global RT advantages. The effect can change from an overall global advantage to a local one, depending on the overall size of the stimuli (Kinchla & Wolfe, 1979), their size ratio (Kimchi & Palmer, 1982), and whether the stimuli are presented to the fovea or the periphery (Lamb & Robertson, 1988). In Experiment 1, the High–Low stimuli were associated with an overall advantage for the high frequency range compared to the low frequency range, despite the fact that the frequencies were corrected for loudness (compare the two rightmost bars of Fig. 2B to the leftmost). Although a significant priming effect of frequency range occurred for both the high and low ranges, the effect for the high

range was marginally larger. In Experiment 2, the Fast–Slow stimuli were not associated with an overall advantage for either the fast or slow temporal ranges in terms of response time, and the priming effects of temporal range were highly similar for both ranges (Fig. 3B). However, participants made fewer errors for the slow temporal range, despite the lack of a response time effect. These results suggest, given the parameters of the current studies, that participants may have been biased towards high frequency (local) ranges and slow temporal (global) ranges. Future research might attempt to manipulate the local or global advantage in these stimuli by varying the absolute frequency and temporal ranges chosen. As with the Navon stimuli, the RT advantages seen with the auditory stimuli may prove rather malleable with variations in stimulus parameters.

4.2. Previous work demonstrating attention to auditory frequency and time

Although the High–Low and Fast–Slow stimuli were inspired by the Navon stimuli and, in particular, the priming design of Robertson (1996), the results of the current two experiments are also of interest purely as studies of auditory attention. We briefly review the relevant work on attention to frequency and time in the auditory modality, and discuss how the current studies complement this literature.

4.2.1. Attention to auditory frequency

The primary method that has been used to demonstrate attentional effects on the processing of frequency has been the "probe-signal method" of Greenberg and Larkin (1968) and Larkin and Greenberg (1970). In this method the participant is led to expect a tone of a particular frequency (the "primary") and must choose which of two temporal intervals contains a near-threshold tone embedded in noise. Participants are presented with the expected primary frequency on the majority of trials, but are occasionally presented with unexpected frequencies ("probes") as well. The primary frequency is detected best, followed by probes that are roughly within a half critical band of the primary frequency. More distant probes are detected poorly. In some cases the expected frequency is the same throughout the experiment, and in others is established by cues of various types presented before each trial (Hafter, Schlauch, & Tang, 1993; Hübner & Hafter, 1995; Macmillan & Schwartz, 1975; Scharf, Quigley, Aoki, Peachey, & Reeves, 1987; Schlauch & Hafter, 1991). These cues need not be informative in order for frequency facilitation to occur, which suggests that exogenous or automatic as well as endogenous or controlled mechanisms may be involved (Green & McKeown, 2001).

More recent methods investigating attention to frequency have manipulated cue validity, inter-stimulus interval (ISI) between cue and target, and task, using tones presented suprathreshold. For instance, Mondor and Bregman (1994) found that a valid frequency cue improved performance on a duration discrimination task for ISIs up to 1500 ms. A similar facilitation for validly cued frequencies has been reported for other orthogonal tasks, such as intensity discrimination (Mondor & Lacey, 2001; Ward, 1997; Ward & Mori, 1996) and rise-time discrimination (Mondor, Zatorre, & Terrio, 1998). Attention has also been given to determining the time course of this facilitation and under what circumstances it reverses to inhibition with longer ISIs (Mondor, Breau, & Milliken, 1998; Mondor, Hurlburt, & Gammell, 2003; Prime & Ward, 2002).

The present Experiment 1 extends this work in auditory attention in three respects. First, the high–low stimuli introduced a new dimension (the discrimination of three-tone pitch patterns) that is orthogonal to frequency but nonetheless is affected by a manipulation of same versus different frequency. Secondly, the experiment incorporated a design in which a response to all stimuli was required, as each one served simultaneously as prime and probe. This is in contrast to the probe-signal method and the majority of other studies that have examined attention to frequency (but see Prime & Ward, 2002, Experiment 1; Mondor, Hurlburt, & Thorne, 2003) and allowed for the comparison and interaction of frequency effects and response effects. Finally, the experiment provides additional support for the exogenous component of frequency-based attention; stimuli in our experiment included same-frequency and different-frequency transitions with equal probability. In other words, there was no cue validity of the prime. Thus, the participant gains no advantage by adopting an explicit, endogenous strategy to attend to the primed range (also see Green & McKeown, 2001; Mondor, Hurlburt, & Gammell, 2003).

4.2.2. Attention to time

The majority of the attention literature in both vision and audition includes a temporal attention component; all studies that incorporate a cue and a target, with either a constant or systematically manipulated ISI, introduce temporal expectations about when the target will occur (see Jones, 2001 for discussion). Mechanisms of temporal attention per se have been explored in at least two different genres of experiment. The first has been concerned with rhythmic sequences of auditory events and how the temporal relationships between events affects the perception of subsequent events in the sequence (i.e. Dynamic Attending Theory; Drake et al., 2000; Jones & Boltz, 1989). A second experimental technique employs a temporal analogue of the Posner (1980) spatial cueing task; participants can be effectively cued by a visual symbol to attend to a specific point in time—one of two temporal intervals following the cue (Coull, Frith, Büchel, & Nobre, 2000; Coull & Nobre, 1998; Griffin, Miniussi, & Nobre, 2002; Miniussi, Wilding, Coull, & Nobre, 1999; for review see Nobre, 2001).

In contrast, the present Experiment 2 demonstrated priming of temporal *ranges*, not points in time. Every sound in the experiment began 2500 ms after the response from the preceding trial, and thus any temporal expectancy about when the stimulus would begin was constant. Similarly, every sound in the experiment was an isochronous sequence of nine 100-ms tones, and thus any temporal expectation about the sequence's length and internal temporal structure was also constant. What Experiment 2 demonstrated was priming of temporal scale—the width of the temporal window (300 or 900 ms) over which relevant information would be presented. We are unaware of any previous experiment in which this was demonstrated. Our result complements previous work in music theory and music psychology suggesting that listeners can appreciate musical structure on various levels within a temporal hierarchy, or what Bharucha (1984) referred to as an *event hierarchy* (Cooper & Meyer, 1960; Drake, 1998; Lerdahl & Jackendoff, 1983; Narmour, 1999; Schenker, 1935; Tillmann & Bigand, 2004).

4.3. Previous approaches to auditory local–global structure

Previous approaches to auditory local–global processing have largely been pursued within the music cognition literature, and largely stem from a study by Bever and Chiarello (1974) in which musically trained participants demonstrated better discrimination between old and new melodies when presented to the right ear, whereas the musically naïve were better when the melodies were presented to the left ear. The authors suggested that the musicians adopted a more "analytic" strategy, which resulted in an advantage for the left hemisphere and right ear.

Later work has attempted to map local and global to two hypothesized levels of representation for melody: scale and contour. This distinction was originally proposed by Dowling (1978) and has received some support at a cognitive level in subsequent work on memory for previously heard melodic patterns (see Dowling & Harwood, 1986, for an early review). Scale refers to the specific set of pitches or tones (7 of 12 semitones of the octave) that are used to construct a melody in Western tonal-harmonic music, along with the set of intervals that can be created between these tones. This is sometimes referred to as a more "local" representation of melody. Contour refers to the overall pattern of ups and downs in a melody and is sometimes considered to be more "global." Experiments invoking the representation of scale versus contour typically include the presentation of two melodies in a same–different task. For the different items, one tone of the melody is changed, either such that both interval and contour have been changed, or such that the interval on the scale has been disrupted but the contour preserved. These stimuli have been used extensively in behavioral studies (Peretz, 1987; Peretz & Morais, 1980, 1987), neuropsychological studies (Ayotte, Peretz, Rousseau, Bard, & Bojanowski, 2000; Liégeois-Chauvel, Peretz, Babaï, Laguitton, & Chauvel, 1998; Peretz, 1990; Zatorre, 1985), and studies employing event-related potentials (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Schiavetto, Cortese, & Alain, 1999; Trainor, McDonald, & Alain, 2002), with at best mixed results regarding the proposed hemispheric asymmetry.

One concern about the use of scale and contour as an auditory analogue for visual local–global stimuli is that it remains unclear what cognitive mechanisms are differentially involved when discriminating melodies along these dimensions. Scale and contour are usually defined methodologically in terms of the degree of precision of the pitch representation; a representation of scale and intervals is more precise and "local" than a representation of contour, which could be more fuzzy and "global." On the other hand, descriptions of the strategies that participants are believed to be using when attending to interval- versus contour-change often evoke a temporal mechanism: attending to intervals within a melody may involve a more fine-grained temporal analysis than attending to contour, which may be accomplished by one, wide temporal window. It is noteworthy in this regard that Bever and Chiarello (1974) study demonstrated the right-ear, "local" bias in their musicians for the old and new melody recognition just after they had been asked to identify whether a specific two-tone sequence had been present in the same melody. Further, Peretz (1987) reported that when conditions encourage a self-terminating, analytic search, "different" responses are made before the melodies are complete, based on the first noticeable change, with a right-ear advantage. Correspondingly, under conditions where the "different" responses primarily occur around melody offset, there is a left-ear

advantage. This suggests to us that scale and contour may only be an approximation of a more fundamental attentional mechanism in the auditory system, perhaps related more directly to the temporal scale over which attention is directed.

A second set of concerns about using scale-contour stimuli as a measure of local and global processing is that they are methodologically limited when compared to Navon stimuli. One methodological strength of the Navon stimuli is that experimental tasks need not be explicitly related to the local–global structure of the stimuli. Depending on the specific experimental question, the participant can be instructed to attend explicitly to either the local or global level (*directed-attention tasks*) or to identify specific target letters regardless of whether they occur at the local or global level (*divided-attention tasks*), making the local–global structure of the stimuli incidental from the perspective of the participants. This is not possible with the scale-contour method. One cannot create a stimulus feature (like a letter) to be presented at either the scale or contour level of a melody. The scale-contour method thus cannot be modified into either a directed attention or divided attention task on a single melody; it is only suitable for same–different or old-new discrimination judgments for pairs of melodies, in which case issues of response bias become more critical.

The other methodological strength of the Navon stimuli is that the letters presented at the local and global levels are experimentally independent; the items at each level can be manipulated orthogonally. This makes it possible to test the extent of independence between the two levels with regard to information processing. For instance, although the local and global items are experimentally independent and provide no information about each other, Navon (1977, Experiment 3) found that global items interfered with the identification of local items, but not vice versa. This observation would not have been compelling with a set of stimuli in which the levels were not experimentally independent. The "local" and "global" levels of scale-contour stimuli suffer from this lack of experimental independence; although an interval change with preserved contour can be created, a contour change with preserved interval cannot be. Thus, any local bias on an interval- and contour-discrimination task would not result in reduced performance for the contour discriminations. This lack of experimental independence should be a fundamental concern to researchers attempting to document selective deficits or dissociations between local and global processing in the auditory modality (Foxton et al., 2003; Mottron, Peretz, & Ménard, 2000).

4.4. Relevance to the Theory of Indispensable Attributes (TIA)

The current studies also support the notion of Kubovy and Van Valkenburg that frequency and time are two "indispensable attributes" in audition Both experiments suggest that attention will be directed towards the frequency or temporal scale of a stimulus exogenously, even when the task is orthogonal and there is no benefit to the participants to attend to these dimensions.

In addition to its positive positions regarding the importance of frequency and time, *TIA* also makes negative positions about the role of space in the perception of auditory objects. This is not to say that there is no role for auditory spatial information in Kubovy and Van Valkenburg's work; they argue that spatial information does play a role in an auditory

"where" system, which is in their view subservient to the visual "where" system. It is with regard to the auditory "what" system that these authors argue for the importance of frequency and time, and not space.

A strong version of *TIA* would predict that auditory attention can be directed only along an indispensable attribute. In fact, Kubovy and Van Valkenburg suggest an interesting reinterpretation of early work in auditory selective attention (e.g. Broadbent, 1958; Cherry, 1953) in terms of frequency–time rather than space: although people are clearly capable of attending to one voice in a crowded room, the mechanisms that permit this selection may not be inherently spatial, but related to the spectral and temporal differences between the attended speaker (the figure) and everyone else (the ground). It is also noteworthy that early extensions of visual attention designs to the auditory modality failed to demonstrate auditory attention operating in space (e.g. Posner, 1978; Scharf et al., 1987, Introduction).

More recent work suggests that a weaker version of this prediction might hold: perhaps attention is directed to the indispensable attributes of frequency and time automatically, even when they are not relevant to the listener's goals (as was found in the current experiments). In contrast, it may be the case that attention to auditory space depends on the presence of either *informative* spatial cues or a task that requires the *explicit* representation of space. For example, Spence and Driver (1994) found that auditory spatial cues affected an auditory spatial discrimination task regardless of whether the cue was informative, suggesting that the mechanism is at least in part unconscious or exogenous. However, spatial cues affected a pitch discrimination task only when the cue was informative, suggesting a more conscious or endogenous mechanism. This suggests that frequency may be automatically encoded during any auditory attention task, whereas the encoding of auditory spatial information is dependent on the specifics of the experimental procedure (see also Buchtel & Butter, 1988; Mondor & Zatorre, 1995; Rhodes, 1987; Woods, Alain, Diaz, Rhodes, & Ogawa, 2001). To date, this hypothesis has not received unequivocal support, however; Mondor, Zatorre, & Terrio (1998, Experiment 2) found that even when performing a task on an orthogonal dimension, both frequency and location cues influenced performance, even when they were uninformative.

We remain agnostic on the issue of auditory spatial attention, as the current experiments did not contrast auditory space with frequency and time. Nevertheless, the similarities between our results and the study they were designed to parallel (Robertson, 1996), are consistent with the positive positions of *TIA*. In the auditory modality, attention can be oriented within the dimensions of frequency and time. This appears to happen automatically or exogenously, i.e. when there is no predictive value of target frequency or temporal range from trial to trial.

5. Conclusion

We have introduced two new kinds of auditory stimuli, one based on a frequency manipulation (the High–Low stimuli) and the other based on a temporal manipulation (the Fast–Slow stimuli) in order to examine selective attention to these two dimensions in the auditory modality. Starting with the goal of creating auditory analogues of the Navon

visual hierarchical letter stimuli, these auditory dimensions were chosen because they are arguably fundamental to the auditory system, and like spatial frequency, have been associated with hemispheric asymmetries. The stimuli were used to demonstrate exogenous attentional persistence to the dimensions of frequency and time, using a task that required a response on an orthogonal variable to every prime and probe. The High–Low and Fast–Slow stimuli may prove useful in future studies investigating parallels between the auditory and visual systems with regard to local–global processing and hemispheric asymmetry, as well as studies within the auditory modality examining attention to what are arguably its two most fundamental dimensions.

Acknowledgements

References

Ackermann, H., & Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: A review and a new hypothesis. *Brain and Language*, *89*, 320–328.

Allard, F., & Scott, B. L. (1975). Burst cues, transition cues, and hemispheric specialization with real speech sounds. *Quarterly Journal of Experimental Psychology*, *27*, 487–497.

Ayotte, J., Peretz, I., Rousseau, I., Bard, C., & Bojanowski, M. (2000). Patterns of music agnosia associated with middle cerebral artery infarcts. *Brain*, *123*, 1926–1938.

Bever, T. G., & Chiarello, R. J. (1974). Cerebral dominance in musicians and nonmusicians. *Science*, *185*, 537–539.

Bharucha, J. J. (1984). Event hierarchies, tonal hierarchies, and assimilation: A reply to Deutsch and Dowling. *Journal of Experimental Psychology: General*, *3*, 421–425.

Bihrle, A. M., Bellugi, U., Delis, D., & Marks, S. (1989). Seeing either the forest or the trees: Dissociation in visuospatial processing. *Brain and Cognition*, *11*, 37–49.

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

Broadbent, D. E. (1958). *Perception and communication*. Elmsford, NY: Pergamon Press.

Buchtel, H. A., & Butter, C. M. (1988). Spatial attention shifts: Implications for the role of polysensory mechanisms. *Neuropsychologia*, *26*, 499–509.

Cherry, E. C. (1953). Some experiments upon the recognition of speech with one and two ears. *Journal of the Acoustical Society of America*, *25*, 975–979.

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*. Chicago, IL: University of Chicago Press.

Coull, J. T., Frith, C. D., Büchel, C., & Nobre, A. C. (2000). Orienting attention in time: Behavioural and neuroanatomical distinction between exogenous and endogenous shifts. *Neuropsychologia*, *38*, 808–819.

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: The neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience*, *18*, 7426–7435.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: Norton.

Delis, D. C., Robertson, L. C., & Efron, R. (1986). Hemispheric specialization of memory for visual hierarchical stimuli. *Neuropsychologia*, *24*, 205–214.

Divenyi, P. L., & Hirsh, I. J. (1975). The effect of blanking on the identification of temporal order in three-tone sequences. *Perception and Psychophysics*, *17*, 246–252.

Divenyi, P. L., & Hirsh, I. J. (1978). Some figural properties of auditory patterns. *Journal of the Acoustical Society of America*, *64*, 1369–1385.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 341–354.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. San Diego, CA: Academic Press.

Drake, C. (1998). Psychological processes involved in the temporal organization of complex auditory sequences: Universal and acquired processes. *Music Perception*, *16*, 11–26.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*, 251–288.

Dukette, D., & Stiles, J. (1996). Children's analysis of hierarchical patterns: Evidence from a similarity judgment task. *Journal of Experimental Child Psychology*, *63*, 103–140.

Farran, E. K., Jarrold, C., & Gathercole, S. E. (2003). Divided attention, selective attention and drawing: Processing preferences in Williams syndrome are dependent on the task administered. *Neuropsychologia*, *41*, 676–687.

Foxton, J. M., Stewart, M. E., Barnard, L., Rodgers, J., Young, A. H., O'Brien, G., & Griffiths, T. D. (2003). Absence of auditory global interference in autism. *Brain*, *126*, 2703–2709.

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, *16*, 1010–1021.

Green, T. J., & McKeown, J. D. (2001). Capture of attention in selective frequency listening. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 1197–1210.

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

Griffin, I. C., Miniussi, C., & Nobre, A. C. (2002). Multiple mechanisms of selective attention: Differential modulation of stimulus processing by attention to space or time. *Neuropsychologia*, *40*, 2325–2340.

Hafter, E. R., Schlauch, R. S., & Tang, J. (1993). Attending to auditory filters that were not stimulated directly. *Journal of the Acoustical Society of America*, *94*, 743–747.

Handel, S. (1988). Space is to time as vision is to audition: Seductive but misleading. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 315–317.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.

Hübner, R., & Hafter, E. R. (1995). Cuing mechanisms in auditory signal detection. *Perception and Psychophysics*, *57*, 197–202.

Ivry, R. B., & Lebby, P. C. (1993). Hemispheric differences in auditory perception are similar to those found in visual perception. *Psychological Science*, *4*, 41–45.

Ivry, R. B., & Robertson, L. C. (1998). *The two sides of perception*. Cambridge, MA: MIT Press.

Jones, M. R. (2001). Temporal expectancies, capture, and timing in auditory sequences. In C. L. Folk, & B. S. Gibson (Eds.), *Attention, distraction and action: Multiple perspectives on attentional capture* (pp. 191–229). Amsterdam: Elsevier.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*, 459–491.

Kim, N., Ivry, R. B., & Robertson, L. C. (1999). Sequential priming in hierarchically organized figures: Effects of target level and target resolution. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 715–729.

Kimchi, R., & Palmer, S. E. (1982). Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 521–535.

Kinchla, R. A., & Wolfe, J. M. (1979). The order of visual processing: Top-down, bottom-up or middle-out. *Perception and Psychophysics*, *25*, 225–231.

Kubovy, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy, & J. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, NJ: Lawrence Erlbaum.

Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 318–320.

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, *80*, 97–126.

Lamb, M. R., & Robertson, L. C. (1988). The processing of hierarchical stimuli: Effects of retinal locus, locational uncertainty, and stimulus identity. *Perception and Psychophysics*, *44*, 172–181.

Lamb, M. R., Robertson, L. C., & Knight, R. T. (1989). Attention and interference in the processing of global and local information: Effects of unilateral temporal–parietal junction lesions. *Neuropsychologia*, *27*, 471–483.

Larkin, W., & Greenberg, G. Z. (1970). Selective attention in uncertain frequency detection. *Perception and Psychophysics*, *8*, 179–184.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Liégeois-Chauvel, C., Peretz, I., Babaï, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, *121*, 1853–1867.

Macmillan, N. A., & Schwartz, M. (1975). A probe-signal investigation of uncertain-frequency detection. *Journal of the Acoustical Society of America*, *58*, 1051–1058.

Martin, M. (1979). Hemispheric specialization for local and global processing. *Neuropsychologia*, *17*, 33–40.

Miniussi, C., Wilding, E. L., Coull, J. T., & Nobre, A. C. (1999). Orienting attention in time: Modulation of brain potentials. *Brain*, *122*, 1507–1518.

Mondor, T. A., Breau, L. M., & Milliken, B. (1998). Inhibitory processes in auditory selection attention: Evidence of location-based and frequency-based inhibition of return. *Perception and Psychophysics*, *60*, 296–302.

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

Mondor, T. A., Hurlburt, J., & Gammell, L. (2003). Enhancement, extension, and reversal of the frequency selectivity effect. *Psychonomic Bulletin and Review*, *10*, 480–487.

Mondor, T. A., Hurlburt, J., & Thorne, L. (2003). Categorizing sounds by pitch: Effects of stimulus similarity and response repetition. *Perception and Psychophysics*, *65*, 107–114.

Mondor, T. A., & Lacey, T. E. (2001). Facilitative and inhibitory effects of cuing sound duration, intensity, and timbre. *Perception and Psychophysics*, *63*, 726–736.

Mondor, T. A., & Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 387–409.

Mondor, T. A., Zatorre, R. J., & Terrio, N. A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 66–79.

Mottron, L., Burack, J. A., Iarocci, G., Belleville, S., & Enns, J. T. (2003). Locally oriented perception and intact global processing among adolescents with high-functioning autism: Evidence from multiple paradigms. *Journal of Child Psychology and Psychiatry*, *44*, 1–10.

Mottron, L., Burack, J. A., Stauder, J. E. A., & Robaey, P. (1999). Perceptual processing among high-functioning persons with autism. *Journal of Child Psychology and Psychiatry*, *40*, 203–211.

Mottron, L., Peretz, I., & Ménard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond central coherence? *Journal of Child Psychology and Psychiatry*, *41*, 1057–1065.

Narmour, E. (1999). Hierarchical expectation and musical style. In D. Deutsch (Ed.), *The psychology of music*, 2nd ed. (pp. 441–472). San Diego, CA: Academic Press.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383.

Neuhoff, J. G. (2003). Pitch variation is unnecessary (and sometimes insufficient) for the formation of auditory objects. *Cognition*, *87*, 219–224.

Nobre, A. C. (2001). Orienting attention to instants in time. *Neuropsychologia*, *39*, 1317–1328.

Pani, J. R., Mervis, C. B., & Robinson, B. F. (1999). Global spatial organization by individuals with Williams syndrome. *Psychological Science*, *10*, 453–458.

Peretz, I. (1987). Shifting ear differences in melody comparison through transposition. *Cortex*, *23*, 317–323.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, *113*, 1185–1205.

Peretz, I., & Morais, J. (1980). Modes of processing melodies and ear asymmetry in non-musicians. *Neuropsychologia*, *18*, 447–489.

Peretz, I., & Morais, J. (1987). Analytic processing in the classification of melodies as same or different. *Neuropsychologia*, *25*, 645–652.

Plaisted, K., Swettenham, J., & Rees, L. (1999). Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. *Journal of Child Psychology and Psychiatry*, *40*, 733–742.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as asymmetric 'sampling in time'. *Speech Communication*, *41*, 245–255.

Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25.

Prime, D. J., & Ward, L. M. (2002). Auditory frequency-based inhibition differs from spatial IOR. *Perception and Psychophysics*, *64*, 771–784.

Rhodes, G. (1987). Auditory attention and the representation of spatial information. *Perception and Psychophysics*, *42*, 1–14.

Robertson, L. C. (1996). Attentional persistence for features of hierarchical patterns. *Journal of Experimental Psychology: General*, *125*, 227–249.

Robertson, L. C., & Delis, D. C. (1986). "Part-whole" processing in unilateral brain-damaged patients: Dysfunction of hierarchical organization. *Neuropsychologia*, *24*, 363–370.

Robertson, L. C., Lamb, M. R., & Knight, R. T. (1988). Effects of lesions of temporal–parietal junction on perceptual and attentional processing in humans. *Journal of Neuroscience*, *8*, 3757–3769.

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

Schenker, H. (1935). *Der freie Satz*. Vienna: Universal Edition.

Schiavetto, A., Cortese, F., & Alain, C. (1999). Global and local processing of musical sequences: An event-related potential study. *Neuroreport*, *10*, 2467–2472.

Schlauch, R. S., & Hafter, E. R. (1991). Listening bandwidths and frequency uncertainty in pure-tone signal detection. *Journal of the Acoustical Society of America*, *90*, 1332–1339.

Sergent, J. (1982). The cerebral balance of power: Confrontation or cooperation? *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 253–272.

Shulman, G. L., & Wilson, J. (1987). Spatial frequency and selective attention to local and global information. *Perception*, *16*, 89–101.

Spence, C. L., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 555–574.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*, 182–198.

Tillmann, B., & Bigand, E. (2004). The relative importance of local and global structures in music. *The Journal of Aesthetics and Art Criticism*, *62*, 211–222.

Trainor, L. J., McDonald, K. L., & Alain, C. (2002). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience*, *14*, 430–442.

Van Valkenburg, D., & Kubovy, M. (2003). In defense of the theory of indispensable attributes. *Cognition*, *87*, 225–233.

Ward, L. M. (1982). Determinants of attention to local and global features of visual forms. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 562–581.

Ward, L. M. (1997). Involuntary listening aids hearing. *Psychological Science*, *8*, 112–118.

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

Woods, D. L., Alain, C., Diaz, R., Rhodes, D., & Ogawa, K. H. (2001). Location and frequency cues in auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 65–74.

Zatorre, R. (1985). Discrimination and recognition of tonal melodies after unilateral excisions. *Neuropsychologia*, *23*, 31–41.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6*, 37–46.