The Ramp Archetype and the Maintenance of Passive Auditory Attention

DAVID HURON
Conrad Grebel College, University of Waterloo

Research on physiological aspects of attention has shown an asymmetry in neurological arousal with respect to the direction of stimulus change. Increases of stimulus intensity level are more effective than equivalent decreases in evoking listeners' attention. Where musical resources limit a continual escalation of intensity level over the course of a work, it is theoretically possible to maintain attention by structuring the work as a sequence of stimulus "ramps"—where intensity increases are small and incremental, but stimulus decreases are large and abrupt. Score-based studies of dynamics in homophonic music and voice entries and exits in polyphonic music show such an asymmetry. It would appear that for a large body of works of western art music, compositional practices are consistent with a theoretical strategy for maintaining passive auditory attention.

Psychologists have traditionally made a distinction between two modes of attention: active and passive. Active attention arises when an individual adopts a self-imposed state of readiness or arousal regarding possible perceptual events. By contrast, passive attention arises when perceptual events themselves capture an individual's attention. In the interaction of humans with the environment, it is doubtful that attention can be so easily divided into passive and active elements, but this conceptual distinction has provided useful bearings for varied research activities. Research concerning active attention focuses on the mental state of the human subject—stressing such factors as motivation, vigilance, and the nature of expectation. By contrast, research on passive attention focuses on those properties or attributes of perceptual stimuli that facilitate the commandeering of consciousness.

The field of attention has attracted considerable study throughout the history of psychology (Bakan, 1966; James, 1890/1950; Kahneman, 1973; Moray, 1969a, 1969b; Parasuraman & Davies, 1984; Underwood, 1976).

Requests for reprints may be sent to David Huron, Conrad Grebel College, University of Waterloo, Waterloo, Ontario, Canada N2L 3G6.
In the area of active auditory attention, various empirical studies have explored vigilance, detection, and focused attention (Dowling, Lung, & Herrbold, 1987; Howard, O'Toole, Parasuraman, & Bennett, 1984; Lowe, 1968; Lukas, 1981; Rhodes, 1987; Scharf, 1988; Scharf, Quigley, Aoki, Peachey, & Reeves, 1987; Squires, Hillyard, & Lindsay, 1973).

In the area of passive auditory attention, one factor known to contribute to attention and arousal is changing loudness. Specifically, existing research suggests that some patterns of dynamic change are better able to arouse listeners and command attentiveness than other dynamic patterns. Of course not all musical compositions are intended to seize the undivided attention of the audience. Some works are intended to evoke a sense of reverie or quiet repose. Other works may rely on novel or subtle manipulations of musical material such that knowledgeable listeners will tend to adopt a state of active attentiveness—whether or not the surface perceptions are overtly attention-grabbing. But in instances where a composer does wish to compose music that directly arouses and commands the attention of listeners, one might expect attention-commanding patterns to predominate.

Physiological research has linked passive attention to the orienting response (Lynn, 1966). The occurrence of the auditory orienting response is most clearly evident when a listener aligns her/his body toward the source of an unexpected sound. But the orienting response rarely includes such gross behaviors. Indeed, its physiological effects are subtle in contrast to related reflexes such as the startle response and the defense reflex. Physiological changes distinctive of orienting response include changes in skin conductance, pupil dilation, heart rate deceleration, cephalic vasoconstriction, peripheral vasoconstriction, as well as electromyographic and electroencephalographic effects (Rohrbaugh, 1984).

In pioneering work by Evgenii Sokolov, orienting response was shown to be induced by changes in the level of stimulation. Sokolov (1960, 1963) showed that orienting response depends on the degree of stimulus change. As the stimulus level departs more from the expected or habituated level, the probability of evoking an orienting response increases. In the domain of auditory stimuli, suitable changes of stimulus may occur with respect to the frequency, intensity, duration, or location of a sound (Apelbaum, Silva, Frick, & Segundo, 1960; Sokolov, 1960).

The orienting response can be caused by both increases of stimulation and decreases of stimulation. However, Bernstein (1968) showed that an asymmetry exists between inducing the orienting response by stimulus increase versus stimulus decrease. Specifically, Bernstein showed that with a habituated auditory stimulus, the threshold for loudness-induced orienting response is much smaller for increasing loudness than for decreasing loudness.
The use of physiological measures with long latency periods (such as vascular and electrodermal measures) precludes the examination of attention-capturing stimuli having fast presentation rates. More recent investigations of the orienting response have centered on the measurement of event-related potentials for the central nervous system. Evoked potentials (EP) have significantly shorter latencies and so are more useful for investigating rapidly changing stimuli—such as occurs in music. Evoked potentials also provide greater precision both with respect to the temporal ordering of event and with respect to the neurophysiological loci of these responses.

Using evoked potentials, Näätänen (1986) has stressed the distinction between the neurological bases of the attention and arousal components of the orienting response—that is, between the commandeering of consciousness versus the heightening of motor tonus (in simple terms: the difference between paying attention and being aroused or excited). Specifically, in a series of seminal papers, Näätänen and his colleagues have studied in detail the so-called “mismatch negativity” (MMN) component of event-related potentials (Näätänen, Gaillard, & Mantysalo, 1978; Näätänen & Picton, 1987). The MMN is evoked by stimuli that deviate from habituated stimuli, and its presence appears to be strongly linked with moments of passive auditory attention. Compared with traditional physiological measures of orienting response, the MMN is somewhat more sensitive: where Bernstein showed specific N1 responses to changes of 15 dB, MMNs appear with loudness changes of only 3 dB and with pitch changes of only 2 Hz (at 1000 Hz).

In general, the MMN is symmetrical with respect to increases or decreases of stimulus magnitude. However, an early P300 evoked potential called the P3a component is known to be asymmetrical—being greater for stimulus increments than for equivalent decrements (Näätänen, 1986). Moreover, it is the P3a component that is implicated in indicating when a stimulus change is consciously observed by a listener (Sams, Paavilainen, Alho, & Näätänen, 1985). In short, attention-commanding stimuli are more likely to be elicited by stimulus increments than by equivalent decrements.

The work of Bernstein, and of Näätänen, Sams, Paavilainen, and Alho have a striking parallel in music-related research. In studies by Nakamura

1. Nomenclature for identifying features of event-related potentials can be somewhat confusing since two labeling systems are currently in use. Features of evoked potentials are identified according to changes of polarity and order: hence P1 for the first positive peak, N3 for the third negative peak and so on. Alternatively, numbers are used to indicate the latencies (in milliseconds) for the stimulus onset: hence P300 for a positive peak occurring about 300 msec after stimulus onset. In the case of auditory evoked potentials, N1 and P2 are synonymous with N100 and P175, respectively. See Näätänen and Picton (1987).
(1980, 1982) it was shown that listeners are significantly better able to recognize crescendos than diminuendos for artificially generated stimuli. This finding was replicated and confirmed by Nakamura (1986, 1987) using actual music as stimuli. In a study of the perceptibility of concurrent voices in polyphonic music, Huron (1989) found that voice exits are significantly more difficult to perceive than voice entries. Specifically, it was found that expert listeners are highly prone to fail to identify a reduction in textural density when a single voice is retired from the texture, whereas single voice entries are perceptually much more recognizable. This asymmetry between the perceptibility of voice entries versus voice exits as well as the asymmetry found by Nakamura between the perceptibility of crescendos versus diminuendos appears to be consistent with the observed asymmetry in physiological measures of auditory attention with respect to the direction of stimulus change. We might speculate that the heightened perception of increased loudness coincides with auditory-induced orienting responses, and that single-voice exits or shallow diminuendos fail to reduce the stimulus level by an amount sufficient to evoke an orienting response. This observation prompted three studies examining stimulus increase/decrease asymmetries with respect to voice entries and exits and with respect to musical dynamics.

Consider the problem of maintaining auditory attention over the course of a musical work. Let us assume for the moment, that at least some composers are intent on attracting and maintaining the listener’s attention. As the composer has fairly direct control over the musical stimuli and only indirect influence on the listener’s mental state, it may be that the primary musical tools for capturing and maintaining attention relate to passive rather than active attention.

To the extent that a composer pursues the objective of engaging a listener’s attention, we might expect a work to be structured in such a manner that orienting responses would be readily generated. Moreover, we might expect attention-commanding works to elicit a more or less consistent sequence of orienting responses or responses like them. Sokolov’s results suggest that one way to achieve this is by constantly changing the musical dynamics: changes of loudness should be both large and frequent. Unfortunately, the extant physiological research provides only a vague idea of how large and how frequent these changes would need to be to maintain a listener’s attention. It is certainly the case that much music exhibits wide fluctuations in loudness, but it is difficult to relate empirically these changes to theories of auditory attentiveness.

However, Bernstein’s results suggest a further refinement that is easier to investigate. Bernstein’s results suggest that a good way to attract the listener’s attention is by organizing the work as a constant escalation of stimulus level. Informally, this prediction makes some sense. A few works,
such as the well-known Bolero by Maurice Ravel or Pli Selon Pli by Pierre Boulez appear to be structured in a way consistent with this view. Indeed, a large number of musical works appear to be organized as extended crescendos.

On the other hand, Bernstein's results further suggest that a long gradual reduction of stimulus level would generate few orienting responses. Again the musical repertoire informally appears to mesh with this prediction: relatively few musical works begin in a boisterous manner and gradually become more and more subdued.

An obvious difficulty with structuring works as a constant increase of stimulus level is that finite musical resources prohibit any unbounded escalation in loudness—precisely the problem articulated by Boulez in the process of composing Pli Selon Pli. The orchestra provides a wealth of resources by which to engineer a crescendo of some length. But we might expect that other instrumental genres—such as the string quartet or solo piano—would be less capable of sustaining a Bolero-like crescendo. A second difficulty with structuring a work as an extended crescendo is that very long crescendos cease to be perceptually salient. Evidence consistent with an optimum rate of intensity change has been found by Mathews (1979). Mathews found that the perceived "strength" of a crescendo is related to its duration. As the duration of a fixed change of intensity is extended from less than 1 sec to 10 sec, the perceived strength increases. However, as the duration of the crescendo is increased further, the result is a reduction in its perceived strength. This suggests that the strength of the crescendo is related to an optimum rate of intensity increase (dB/sec), and that for a given change in intensity level, there exists a maximum duration over which a crescendo will remain perceptually salient.

However, the crescendo structure is not the only technique for generating orienting responses characteristic of passive auditory attention. Recall that orienting responses can be generated also by reductions in stimulus level, provided that the stimulus decrements are sufficiently large. On the basis of the extant research on auditory attention, we might predict that a sufficient strategy for maintaining a relatively constant stream of orienting responses would be to structure complex events as a sequence of stimulus "ramps." That is, regular increments of stimulus level would be followed by occasionally large decrements of stimulus level. This strategy is schematically illustrated in Figure 1.

![Fig. 1. Schematic illustration of stimulus ramps.](image-url)
The most obvious way of manipulating the stimulus level is by changing the loudness (as through the control of musical dynamics). In a study of the 32 piano sonatas by Beethoven (Huron, 1990b), it was found that a significant asymmetry exists between increasing and decreasing dynamics. Specifically, it was found that (1) crescendos are more frequent than diminuendos, (2) crescendos typically last longer than diminuendos, (3) large changes of dynamics tend to reduce the loudness, and (4) crescendos will follow low dynamic levels more often than diminuendos will follow high dynamic levels. A subsequent study (Huron, 1991) extended and replicated this work using 435 piano works by 14 additional composers. In general, the results were confirmed. Brahms, Chopin, Clementi, Debussy, Dvořák, Franck, Gershwin, Grieg, Ireland, Liszt, Mozart, Scriabin, and Villa-Lobos all use dynamic gestures that are consistent with the ramp pattern. Only one composer studied (Carl Nielsen) failed to exhibit dynamic patterns consistent with the above results. In general, these studies support a “ramp archetype” of musical dynamics in which musical passages tend to build in a gradual way, but tend to subside relatively quickly. Such dynamic gestures are consistent with the conditions for evoking a continuous succession of orienting responses.

Apart from dynamic shading, another means by which the level of auditory stimulation can be manipulated is by the addition and subtraction of voices in a musical texture. In Huron (1990a), a study of 159 polyphonic works by 25 composers revealed a marked asymmetry between voice entries and exits. Specifically, there is a significant tendency to add voices one at a time, while voice exits tend to occur several at a time. Even when textural changes are measured by the density of successive vertical sonorities rather than by voice-entries or exits, an increment/decrement asymmetry is clearly evident. Again these results are consistent with the conditions for evoking orienting responses.

That both baroque polyphony and romantic homophony would share the prevalence of a ramp archetype is somewhat unusual. The changes in musical language from the end of the baroque period to the subsequent classical/romantic musical language are quite marked. Whereas baroque polyphony tends to ignore dynamics as a significant compositional attribute, dynamics become paramount by the romantic period. Conversely, the baroque preoccupation with polyphonic introductions and retirements of voices is superseded by a generally homophonic or tune-and-accompaniment approach in the subsequent periods. Generalizing somewhat, we may point out a reciprocal relationship between the compositional approaches of polyphonic and homophonic musics. In polyphonic music, voicing tends to be organized according to the ramp archetype, whereas dynamics are organized in a “terraced” fashion. Conversely, in
homophonic music, dynamics tend to conform to the ramp archetype whereas voicing tends to be terraced.

What both types of music share in common is the manner in which musical resources “build” versus the manner in which they “subside.” The music tends to build in a gradual way, whereas it tends to subside relatively quickly.

Conclusion

Research concerning physiological aspects of attention has shown that two factors are paramount in evoking a neurological state of attentiveness. In the first instance, an important factor is the change of stimulus level. In the second instance, increases of stimulus level are known to be more effective than equivalent decreases of stimulus level. The effect of this latter factor has been observed in studies of music where it has been found that single-voice entries are more easily perceived than single-voice exits in multivoiced textures. This effect has also been observed in studies that show that crescendos are more easily perceived than diminuendos.

To the extent that a musical work attempts to maintain the interest of a listener, we might expect the structure of a work to be consistent with neurophysiological principles of attention. Where a musical genre makes available many resources, it is possible to engineer long passages where the amount of stimulation is constantly increasing. However, where a musical genre provides only modest resources, the possibilities for a perpetual escalation of stimulation are soon exhausted. In such circumstances it is possible to maintain attention by rapid and large changes of dynamics, or by structuring the work as a sequence of stimulus “ramps.”

A study of the dynamics in 15 piano composers reveals that changes of loudness tend to be structured as a sequence of passages that build in a gradual way, but subside relatively quickly. A parallel study of sonorities in 25 polyphonic composers reveals a similar asymmetry in the increment and decrement of textural density. Although by no means universal, the “ramp archetype” is clearly evident in the great majority of works studied. It would appear that for a large body of works of western art music, compositional practices are consistent with a theoretically well-adapted strategy for maintaining auditory attention.

In interpreting these results, it is important to bear in mind several caveats. First, there are other dynamic patterns that are in principle consistent with the maintenance of passive auditory attention, yet have not been tested in any of the studies cited. Second, no account has been taken regarding rhythmic, harmonic, melodic, and timbral dimensions of au-
ditor attention or arousal. In particular, the possibility of composers pursuing more sophisticated manipulations to promote active attention strategies on the part of the listener have been ignored here. Third, it is worth reiterating that it is undoubtedly not the purpose of all musical compositions to command the undivided attention of the audience. Erik Satie, for example, once proposed to have his music mingle with the sound of forks and knives at the dinner table—a music that would not attract attention to itself. In short, one must be wary of construing these methods as a critical tool for evaluating composers or compositions.2

References


2 The author extends his thanks to Mark Haggard of the Institute of Hearing Research (Nottingham, U.K.) for offering comments on an earlier draft of this paper. Software support provided by Mortice Kern Systems is gratefully acknowledged. This research was supported in part through funds provided by the Social Sciences and Humanities Research Council of Canada, the Canadian Centennial Scholarship Fund (U.K.), and the U.K. Committee of Vice-Chancellors and Principals.


