Notes and Comment

Auditory spatial alternation transforms auditory time (again): Comments on Lakatos (1993), "Temporal constraints on apparent motion in auditory space"

GERT TEN HOOPEN

Leiden University, Leiden, The Netherlands

Lakatos (1993) reported interesting data that indeed support "the hypothesis that the extent of spatial separation between successive sound events directly affects the perception of time intervals between these events" (p. 139). The present comment is an attempt to show that, as far as the horizontal plane is concerned, Lakatos's hypothesis was already answered qualitatively by Axelrod and coworkers (Axelrod & Guzy, 1968; Axelrod, Guzy, & Diamond, 1968) in their studies of attention shifting, and by ten Hoopen and coworkers, who quantified the amount of "interaural time dilation" (Akerboom, ten Hoopen, Olierook, & van der Schaaf, 1983; ten Hoopen, 1982; ten Hoopen, Vos, & Dispa, 1982). Nonetheless, Lakatos's study is very worthwhile. It originated from the realm of "apparent motion paradigms," but I will argue that the study rather used an "auditory streaming paradigm," and that the data is a welcome contribution to elucidate how the perceptual processes of auditory stream formation and interaural time dilation interact.

Lakatos (1993) stated that he used an "apparent motion paradigm" to study the perception of time intervals between spatially alternating sounds. In the present comment, I argue that he rather used an "auditory streaming paradigm." First, I will explain this paradigm with reference to van Noorden's (1975) well-known work on the perceptual organization of frequency-alternating tone sequences. Then I will attempt to demonstrate that the perceptual changes that Lakatos found by adding spatial alternation to frequency alternation can be explained in terms of the mechanism of interaural time dilation (e.g., ten Hoopen, Vos, & Dispa, 1982). Finally, I will make some remarks about the interaction between frequency, space, and time in stream formation.

Auditory Streaming by Frequency

Because the phenomenon of auditory streaming is very well documented (cf. Bregman, 1978, 1990, 1993), I only

recap the very core of it with the help of van Noorden's (1975) indispensable Figure 2.9, which is adapted here as Figure 1.¹ Three perceptual areas are delimited by the fission boundary (FB) and the temporal coherence boundary (TCB). Frequency-alternating tone sequences that have combinations of tone intervals and stimulus onset asynchronies (SOAs) in the area above the TCB are inevitably perceived as streaming. Tone sequences in the area below the FB are inevitably perceived as coherent. Tone sequences between these boundaries can be listened to as either coherent or streaming at will. The course of the TCB shows that with decreasing frequency separation, the SOA should be decreased to just maintain inevitable streaming. Thus, we have a tradeoff between frequency separation and tempo.

Adding Spatial Alternation

Lakatos's frequency-alternating patterns comprised four tones—for example, G₃, C#₄, A₄, and D#₄ (392, 554, 440, and 622 Hz, respectively)—which alternated continuously in that order (see Lakatos, 1993, p. 140, Figure 2, pattern 2). His observers were initially presented with a relatively slow tempo of the recurring pattern, so that *one stream* (coherence) could be heard, and the task was to increase the overall rate until one could no longer stick to coherence. Lakatos called this boundary the "critical SOA." In Figure 1, it is illustrated what Lakatos's subjects did: The horizontal dashed arrow, touching the temporal coherence boundary (TCB), stands for the descending trajectory of the adjusting observer, and the perpendicular projection of the touching point with the TCB is the "critical SOA."

The major variable in Lakatos's study was the angle between speakers over which the tones were alternated one by one. He found an increase of the critical SOA when widening the angle from 20° to 110°. With his pattern 2, for instance, this increase was from 160 to 195 msec (estimated by eye from his Figure 3b). In other words, with increasing spatial separation, the SOA can be increased by about 35 msec (for the 90° widening of angle) before the inevitable percept of streaming gets lost and coherence can just be heard at will. Thus, a tradeoff between spatial separation and tempo was established.

The Axelrod Effect or Interaural Time Dilation

It occurs to me that this tradeoff is just another example of the phenomenon of interaural time dilation. Axelrod and Guzy (1968), Axelrod, Guzy, and Diamond (1968), and Guzy and Axelrod (1971) reported that the apparent repetition rates of interaural (ear-alternating) sound sequences were underestimated in comparison with those of monaural control sequences.² In addition

I am very grateful to reviewer Stephen Handel, who aptly and amiably rescued me from a logical trap. I am indebted to an anonymous reviewer and to Associate Editor Neil A. Macmillan for helpful comments. Correspondence should be addressed to G. ten Hoopen, Unit of Experimental Psychology, Faculty of Social and Behavioral Sciences, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands (e-mail: tenhoopen@rulfsw.leidenuniv.nl).

Figure 1. A repeating sequence of two alternating tones of different frequencies is perceived as one stream below the fission boundary (FB), but as two streams above the temporal coherence boundary (TCB). In the ambiguous area between FB and TCB, it depends on the set of the observer whether a percept of one or of two streams is heard. The inserted dashed line represents the descending trajectory of the observer in Lakatos's study who adjusts the stimulus onset asynchrony (SOA) until the TCB is reached, yielding the "critical SOA."

to a 180° alternation condition, Axelrod and Powazek (1972) also manipulated the angular separation between alternating sounds (12°, 45°, 90°, 135°, and 168°) and found that the apparent rate of the alternating sequence decreased with increasing angle.

Thus, in a qualitative sense, an answer to Lakatos's (1993) question of whether the subjective tempo of a spatially alternating sound sequence slows down when the azimuthal separation between consecutive sounds increases was already given by Axelrod and coworkers in 1968. Axelrod et al. (1968) attempted to quantify the subjective slowing down of interaural sequences in terms of rate. They plotted the monaural subjective rates against objective interaural rates and found a linear relation: monaural rate $\approx 0.60 \times$ interaural rate (clicks/second). However, the underestimation of interaural rates by about 40% could not be explained. In an unpublished study (ten Hoopen, Akerboom, & Boelaarts, 1985; referred to in Bregman, 1990, p. 82), we pointed out why this equation, relating monaural and interaural rates, might be less informative than it looked superficially.³

In a series of studies (ten Hoopen, 1982; ten Hoopen & Vos, 1979, 1980, 1981; ten Hoopen, Vos, & Dispa, 1982), we solved the quantification problem step by step by concentrating on interaural and monaural *time intervals* rather than on *rates*. We called the mental time template of the SOA *perceptual onset asynchrony* (POA). By convergent operations (absolute estimation, counting, and making stop reactions to sequences of unpredictable length), we determined that an interaural POA was about 25 msec longer than its monaural counterpart. In a subsequent study, we demonstrated that the difference between interaural and monaural POAs was independent of the value of SOA, throughout a wide SOA range of 40–2,130 msec (Akerboom, ten Hoopen, Olierook, & van der Schaaf, 1983).

Thus, when sounds are alternated between left and right ears, the subjective time intervals between their percepts are dilated in auditory memory by an amount correspond-

ing to about 25 objective milliseconds ($POA_{\text{interaural}} = POA_{\text{monaural}} + 25 \text{ msec}$). In Akerboom et al. (1983, Experiment 6), we also varied the amount of spatial separation from 0° to 180° in steps of 30°. We established a linear increase of POA with increasing angle as measured by the stop reaction time paradigm. The title of the Akerboom et al. (1983) study expressed that finding clearly: "Auditory spatial alternation transforms auditory time."

The number of sounds in the sequences that we used in most of our studies ranged between 7 and 30, but in ten Hoopen and Akerboom (1983), we showed that the POA difference between interaural and monaural conditions also appeared when sequences comprised very few sounds. Specifically, in conditions comparable to those of Lakatos—that is, in an SOA range of 50–400 msec and with four sounds in sequence—the POA increase from 0° to 180° amounted to about 30 msec.

Interpretation of Lakatos's Results

The temporal fate in auditory memory of the tone *percepts* in one cycle of Lakatos's pattern 2 is illustrated in Figure 2, where two spatial cases are drawn: monaural presentation (top diagram), and interaural presentation (bottom diagram). According to the mechanism of interaural time dilation, the POAs increase along the dashed lines when the four tone pattern is alternated over increasingly wider angles.

With pattern 2, Lakatos found an increase of the critical SOA from about 160 to 195 msec, when widening the angle from 20° to 110°. In Figure 3, the POA is plotted as a function of SOA, and dependent on the smallest (20°) and largest degree (110°) of alternation. Although this figure is a bit simplified,⁴ it suffices for the present explanation. Figure 3 shows that the POA corresponding to the critical SOA of 160 msec at 20° is 163 msec [$160 + (20^\circ/180^\circ) \times 30 \text{ msec}$]. The POA corresponding to the critical SOA of 195 msec at 110° is 213 msec [$195 + (110^\circ/180^\circ) \times 30 \text{ msec}$]. So, over this 90° angular increase, the

Figure 2. The temporal fate of one of Lakatos's four tone patterns in auditory memory when it would have been presented monaurally (top) or interaurally (bottom). As a result of interaural time dilation, depicted by the widening dashed lines, the perceptual onset asynchrony (POA) between the tone percepts is about 30 msec larger in the 180° situation. The tone frequencies are represented by the line lengths.

Figure 3. The perceptual onset asynchrony (POA) as a function of the stimulus onset asynchrony (SOA) between tones, dependent on the degree of angular separation between tones (20° and 110°), as predicted by the mechanism of interaural time dilation (see text). The critical SOAs that Lakatos found for his pattern 2, and the corresponding POAs, are indicated.

critical POA increases by $213 - 163 = 50$ msec. Lakatos's finding that the critical SOA increases approximately linearly with increasing angle is consistent with the mechanism of interaural time dilation: the POA increases linearly with larger angular alternation.

Frequency Alternation, Spatial Alternation, and Time

Finally, some remarks on the combined role of frequency separation, spatial separation, and time with respect to streaming. First, frequency separation and tempo can be traded off in such a way that inevitable streaming is just maintained (van Noorden, 1975). Second, spatial separation and tempo can be traded off similarly (Lakatos, 1993). It seems logical, therefore, to conclude that frequency separation and spatial separation can be traded off too. However, such a conclusion is not warranted without knowing the effect of frequency alternation per se on the POA. In theory it could be possible that the POA is shrunk more and more when the frequency separation increases. If that were the case, we could only predict a tradeoff when we knew the amount of POA shrinking in comparison to the amount of POA dilation caused by spatial separation. Fortunately, we gathered some evidence that the POA is not affected by frequency separation (Akerboom, ten Hoopen, & van der Knoop, 1985; ten Hoopen & Akerboom, 1982; ten Hoopen, van Meurs, & Akerboom, 1982). Hence, it is my supposition that the TCB will shift to larger SOAs when the spatial separation increases. To the best of my knowledge, there exist

no studies that established the TCB of sequences that alternate simultaneously between locations and frequencies.⁵ Van Noorden (1975, p. 22) described that he listened to such sequences, but he did not collect systematic data. Future research, guided by Lakatos's results, should establish such TCBs.

REFERENCES

- AKERBOOM, S., TEN HOOPEN, G., & VAN DER KNOOP, A. (1985). Does the octave illusion evoke the interaural tempo illusion? *Perception & Psychophysics*, **38**, 281-285.
- AKERBOOM, S., TEN HOOPEN, G., OLIEROOK, P., & VAN DER SCHAAF, T. (1983). Auditory spatial alternation transforms auditory time. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 882-897.
- AXELROD, S., & GUZY, L. T. (1968). Underestimation of dichotic click rates: Results using methods of absolute estimation and constant stimuli. *Psychonomic Science*, **12**, 133-134.
- AXELROD, S., GUZY, L. T., & DIAMOND, I. T. (1968). Perceived rate of monotic and dichotically alternating clicks. *Journal of the Acoustical Society of America*, **43**, 51-55.
- AXELROD, S., & NAKAO, M. (1974). Apparent slowing of bimanually alternating pulse trains. *Journal of Experimental Psychology*, **102**, 164-166.
- AXELROD, S., & POWAZEK, M. (1972). Dependence of apparent rate of alternating clicks on azimuthal separation between sources. *Psychonomic Science*, **26**, 217-218.
- BREGMAN, A. S. (1978). The formation of auditory streams. In J. Requin (Ed.), *Attention and performance VII* (pp. 63-75). Hillsdale, NJ: Erlbaum.
- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- BREGMAN, A. S. (1993). Auditory scene analysis: Hearing in complex environments. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 10-36). Oxford: Oxford University Press, Clarendon Press.
- BUTLER, D. (1979). A further study of melodic channeling. *Perception & Psychophysics*, **25**, 264-268.
- DEUTSCH, D. (1974). An auditory illusion. *Nature*, **251**, 307-309.
- DEUTSCH, D. (1975, October). Musical illusions. *Scientific American*, **233**(4), 92-104.
- DEUTSCH, D. (1986). Auditory pattern recognition. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 2, pp. 32-1 to 32-49). New York: Wiley.
- DEUTSCH, D., & ROLL, P. (1976). Separate "what" and "where" decision mechanisms in processing a dichotic tonal sequence. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 23-29.
- GUZY, L. T., & AXELROD, S. (1971). Synchronization of unimanual and bimanual responses with monotic and dichotic clicks. *Perception & Psychophysics*, **9**, 161-164.
- GUZY, L. T., & AXELROD, S. (1972). Interaural attention shifting as response. *Journal of Experimental Psychology*, **95**, 290-294.
- HANDEL, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.
- HUGGINS, A. W. F. (1974). On perceptual integration of dichotically alternated pulse trains. *Journal of the Acoustical Society of America*, **56**, 939-943.
- JONES, M. R., & YEE, W. (1993). Attending to auditory events: The role of temporal organization. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 69-112). Oxford: Oxford University Press, Clarendon Press.
- LAKATOS, S. (1993). Temporal constraints on apparent motion in auditory space. *Perception & Psychophysics*, **54**, 139-144.
- MASSARO, D. W. (1976). Perceiving and counting sounds. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 337-346.
- MCADAMS, S., & BREGMAN, A. S. (1979). Hearing musical streams.

NOTES

- Computer Music Journal*, 3, 26-43. Reprinted in C. Roads & J. Strawn (Eds.) (1985). *Foundations of computer music* (pp. 658-698). Cambridge, MA: MIT Press.
- NAKAJIMA, Y. (1987). A model of empty duration perception. *Perception*, 16, 485-520.
- NAKAO, M. A., & AXELROD, S. (1976). Effects of bilateral alternation on perceived temporal uniformity of auditory and somesthetic pulse trains. *Perception & Psychophysics*, 20, 274-280.
- SMITH, J., HAUSFELD, S., POWER, R. P., & GORTA, A. (1982). Ambiguous musical figures and auditory streaming. *Perception & Psychophysics*, 32, 454-464.
- TEN HOOPEN, G. (1982). *The perceptual organization of alternating sound sequences*. Unpublished doctoral dissertation, Leiden University.
- TEN HOOPEN, G. (1985). The detection of anisochrony in monaural and interaural sequences. In J. A. Michon & J. L. Jackson (Eds.), *Time, mind, and behavior* (pp. 140-150). Berlin: Springer-Verlag.
- TEN HOOPEN, G., & AKERBOOM, S. (1982). The perceived tempi of coherent and streaming tone sequences: II. *Perception & Psychophysics*, 32, 481-485.
- TEN HOOPEN, G., & AKERBOOM, S. (1983). The subjective tempo difference between interaural and monaural sequences as a function of sequence length. *Perception & Psychophysics*, 34, 465-469.
- TEN HOOPEN, G., AKERBOOM, S., & BOELAARTS, L. (1985). *Streaming by locus or interaural time dilatation?* Unpublished manuscript, Leiden University, Unit of Experimental Psychology.
- TEN HOOPEN, G., BOELAARTS, L., GRUISEN, A., APON, I., DONDERS, K., MUL, N., & AKERBOOM, S. (1994). The detection of anisochrony in monaural and interaural sound sequences. *Perception & Psychophysics*, 56, 110-120.
- TEN HOOPEN, G., VAN MEURS, G., & AKERBOOM, S. (1982). The perceived tempi of coherent and streaming tone sequences. *Perception & Psychophysics*, 31, 256-260.
- TEN HOOPEN, G., & VOS, J. (1979). Attention-switching and grouping in counting interaurally presented clicks. *Acta Psychologica*, 43, 283-297.
- TEN HOOPEN, G., & VOS, J. (1980). Attention switching is not a fatigable process: Methodological comments on Axelrod & Guzy (1972). *Journal of Experimental Psychology: Human Perception & Performance*, 6, 180-183.
- TEN HOOPEN, G., & VOS, J. (1981). Attention switching and patterns of sound locations in counting clicks. *Journal of Experimental Psychology: Human Perception & Performance*, 7, 342-355.
- TEN HOOPEN, G., VOS, J., & DISPA, J. (1982). Interaural and monaural clicks and clocks: Tempo difference versus attention switching. *Journal of Experimental Psychology: Human Perception & Performance*, 8, 422-434.
- VAN NOORDEN, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Unpublished doctoral dissertation, Institute of Perception Research, Technische Hoogeschool, Eindhoven.
1. It is not so easy to lay one's hand on the thesis of van Noorden (1975), but adaptations from his Figure 2.9, with good discussions, can, for instance, be found in Deutsch (1986), Handel (1989), Jones and Yee (1993), and McAdams and Bregman (1979).
2. The "Axelrod effect" is very robust. I have listened innumerable times to my classroom audiodemo, which switches back and forth between interaural and monaural sequences, having the same objective tempi, and I am still surprised by the impressive subjective tempo difference. The Axelrod effect also strongly influences performance on other tasks, such as counting numbers of sounds and the detection of anisochrony. Guzy and Axelrod (1972), Massaro (1976), and ten Hoopen and Vos (1981) reported that subjects undercount the number of sounds in interaural sequences as compared with monaural controls. Nakao and Axelrod (1976), van Noorden (1975), ten Hoopen (1985), and ten Hoopen et al. (1994) demonstrated that the threshold for detecting anisochrony is higher in interaural sound sequences than in monaural ones. The Axelrod effect is not restricted to the auditory modality: Axelrod and Nakao (1974) showed that the temporal sluggishness of spatial alternation holds for the tactile modality as well. In sum, the Axelrod effect reflects how our perceptual system temporally maps the succession of spatially alternating events.
3. Although the objective interaural rates were reasonably divided between about 1 and 40 events per second, their inverse values—that is, the time intervals between sounds—were not. Values between 130 and 970 msec were not covered, and hence the reliability of the equation should be doubted. Similar criticisms led Huggins (1974) to perform an improved replication of Axelrod et al. (1968). However, he also quantified interaural temporal processing in terms of rates, which, in our opinion, did not yield a satisfying solution either. Three subjects adjusted the rate of a nonalternating pulse train to 16 different rates of an interaural pulse train. The 16 corresponding SOAs covered a range of 21–667 msec. When the SOA was larger than 120 msec, subjects successfully matched the interaural rate. When the SOA was smaller than 50 msec, subjects matched the rate of the pulses in one ear. I am indebted to Professor A. W. F. Huggins for providing his original data.
4. The monaural POA and SOA, although linearly related, differ by a constant. Nakajima (1987) established the psychophysical relation between subjective and objective empty time intervals by convergent operations: $\tau(t) = k(t + 80 \text{ msec})$. In our terminology: $\text{POA} = k(\text{SOA} + 80 \text{ msec})$.
5. However, experiments in which location alternation and frequency alternation in tone sequences are put in competition (instead of cooperation) to study principles of auditory organization are abundant (e.g., Butler, 1979; Deutsch, 1974, 1975; Deutsch & Roll, 1976; Smith, Hausfeld, Power, & Gorta, 1982).

(Manuscript received January 10, 1994;
revision accepted for publication November 6, 1994.)