Auditory saltation in the vertical midsagittal plane

Susan E Boehnke, Dennis P Phillips
Hearing Research Laboratory, Department of Psychology, Dalhousie University, Halifax, NS B3H 4J1, Canada; e-mail: dennis.phillips@dal.ca
Received 18 August 2003, in revised form 10 June 2004; published online 16 March 2005

Abstract. Auditory saltation is an illusion in which a train of clicks, the first half of which is presented at one location and the other half of which is presented from a second location, is perceived as originating not only from the anchor points, but also from locations between them. That is, intermediate members of the series of clicks have their spatial locations systematically misperceived. In the present study, auditory saltation was examined for the first time in the vertical midsagittal plane. Subjects rated the perceived continuity of motion for 8-click trains systematically varied in inter-click interval (ICI), direction of motion (up, down), and trial type (‘saltation’ versus ‘real’ motion). In all listeners, saltation stimuli supported robust saltation, but only for trials with ICIs less than about 120 ms. Real motion was rated as continuous for all ICIs. These data indicate that the auditory-saltation illusion can exploit monaural stimulus cues for source location in the generation of the illusory motion percept.

1 Introduction
Sensory saltation is an orderly misperception of the spatial location of members of a stimulus train that occurs when the elements of the train are presented repetitively at successive locations with regular, and short, inter-stimulus intervals (after Geldard 1975; Geldard and Sherrick 1986). Auditory saltation has most often been demonstrated with short trains of dichotic clicks, in which the first half of the click train is lateralised to one side, and the second half is lateralised to the opposite side. The resulting percept is of clicks located not only at the ‘sites’ of stimulation, but also at points between them, as if there were a single source moving between the anchor points and emitting clicks periodically as it does so (Hari 1995; Shore et al 1998; Phillips and Hall 2001). Note that this saltatory motion is distinguished from true continuous motion by the discontinuous nature of the stimulus itself, and from visual apparent motion by the discontinuity of the perceived motion (after Geldard and Sherrick 1986; Phillips and Hall 2001). The illusory saltatory motion is also robust in the free field (Bremer et al 1977; Phillips et al 2002).

The illusion is usually demonstrated by the reported change in percept of the click train as a function of the rate [or duration of the inter-click interval (ICI)]. For ICIs less than about 100 ms, listeners typically report the illusory perception—continuous movement of the clicks across spatial location. At longer ICIs (>150 ms) listeners typically rate the saltatory train veridically. For intermediate ICIs (100–150 ms), listeners report an intermediate perception with the middle clicks incompletely displaced from their veridical location. Typically, the middle value of the rating scale has been taken as the ‘threshold’ ICI to generate the illusion. For 8-click trains this value falls near 105 ms (Shore et al 1998), while for 6-click trains this value is slightly higher, near 120 ms (Phillips et al 2002), consistent with a temporal integration window which averages spatial information from the clicks falling within its duration. A sliding temporal window 300–450 ms in width can account for these results (see Phillips and Hall 2001, experiment 3).

§Susan E Boehnke is currently at the CIHR Group in Sensory-motor Systems, Department of Physiology, Queen’s University, Kingston, ON K7L 3N6, Canada; e-mail: susan@eyeml.queensu.ca
¶Author to whom all correspondence should be addressed.
Dichotic studies of auditory saltation have revealed that the illusion is equally supported by clicks lateralised on the basis of either interaural time differences or interaural level differences (Phillips and Hall 2001). Furthermore, switching the lateralisation cue mid train has no impact on ratings (Phillips et al 2002). This raises the question whether the illusion can also exploit monaural stimulus information for sound location. This question is important, because it addresses the level in central auditory processing from which the illusion can arise. The saltation illusion might emerge from the processing specifically of the change in binaural cue information (which is initiated at a low level in the auditory system—Phillips 2001). In such a case, the effect might be attributable to strictly binaural processing phenomena, eg ‘binaural sluggishness’ (eg Akeroyd and Summerfield 1999). Alternatively, the illusion might emerge from a high-level spatial representation, ie one which is insensitive to the origins of the spatial information in binaural and/or monaural stimulus information. In this regard, Geldard (1975) concluded that the effect must be central, because the illusion occurs in all spatial senses (somatosensation, audition, and non-foveal vision).

The purpose of the present experiment was to address this issue by determining whether auditory saltation occurs in the elevation dimension (upwards or downwards) along the midline plane in the free field. Sources in the median sagittal plane produce near-zero interaural stimulus disparities irrespective of elevation; thus, any differences in perceived elevation for sources in this plane are likely caused by differences in the spectral content of the signals reaching the listeners (eg Roffler and Butler 1968; Blauert 1969/1970; Butler and Belendiuk 1977; Middlebrooks et al 1989). The localisation judgment is thought to be based on a comparison of the incoming signal spectrum with an internal template (Middlebrooks 1992), either of the source itself (ie knowledge of the expected sound spectrum), or of the head- and pinna-related transfer function. This process distinguishes vertical from azimuthal localisation, because the latter is based primarily on interaural differences in stimulus timing or level (Wightman and Kistler 1992). The data from the present experiment suggest that auditory saltation does indeed occur in the elevation dimension and suggest, therefore, that monaural spatial cues can produce the illusory saltatory motion.

2 Methods
2.1 Subjects
Six subjects (three male) took part in the experiment. Three listeners were naïve, and the other three (SEB, DPP, SEH) were familiar with auditory-saltation experiments and the theoretical basis for the study. All but one subject had audiometrically normal hearing; the exception had a mild unilateral loss at 8.0 kHz.

2.2 Apparatus and stimuli
Trains of 8-click stimuli were generated digitally and presented by a Power Macintosh G3 computer with Matlab® software (The Mathworks). Each click comprised 2 samples set to an amplitude of one, embedded in zeros with a duration equal to the ICI for that condition. The click trains were generated and played at a sampling frequency of 44100 Hz, and thus each click was 45 µs in duration. The digital signal was passed through an ECHO Darla multichannel sound card using the ASIO protocol, amplified, led into an Eckel sound-attenuating room, and transduced by eight Optimus Pro X44 (40-2080) speakers (0.09–20.0 kHz frequency response), roughly matched for frequency response. The walls and ceiling of the room were lined with 2-inch Auralex foam wedges, and the floor was carpeted to further improve sound absorption. The speakers were stacked on their sides (11 cm high), and separated by 4-cm thick spacers. The total speaker stack including spacers was 120 cm in height. This stack rested on a wooden shelf that was raised 40 cm from the floor. There was a small light-emitting
diode on each speaker; these diodes were lit, all at once, at the beginning of the experiment. Once the listener initiated the first trial (by key press), the LEDs were not presented again. Stimuli were presented at a comfortable listening level. The amplitudes of single clicks were approximately 45 dB SPL (A-weighted) when measured from a distance of 10 cm. The listener was seated opposite the speaker array, at a distance of 130–140 cm with the interaural axis approximately 95–110 cm from the floor, depending on the height of the listener. This placed the head of taller listeners slightly above the centre of the speaker array. The top and bottom speakers were thus at an elevation angle of $+20^\circ$ and $-20^\circ$ respectively, relative to the interaural axis. Listeners used a chin-rest to maintain a constant head position throughout testing. Note that the stimulus arrangement (vertical stacking of the speakers) had the disadvantage that, because the speakers were not located on an arc around the midsagittal plane, the speakers at the limits of the array were slightly further ($\leq 14$ cm) from the listener than were those in the centre of the array. On the other hand, this system had the virtue that the speakers were located in the line of gravity.

2.3 Design and procedure

There were 3 independent variables: the type of motion (‘saltation’ or ‘real’ stimulus position change), the direction of motion (up or down), and the inter-click interval (ICI) (30, 60, 90, 120, 240, 480 ms). Thus, there were 24 types of click trains, and listeners were presented 10 instances of each in a fully randomised order. In the ‘saltation’ trials, 4 clicks were presented from a speaker at one end of the array, followed by 4 clicks from the speaker at the other end of the array. In the ‘real’ motion trials, 1 click was presented from each of the speakers in sequential order (top through bottom, or bottom through top, depending on the direction of motion). Note that both trains are ‘saltatory’ motion (jumping) and are not to be confused with continuous motion. The direction of motion was included as a variable primarily to prevent the listener from anticipating which speaker would be activated first, and therefore from selectively attending to one end of the speaker array at train onset. There was no a priori reason to expect a directional asymmetry, though it would be interesting if ‘falling’ auditory motion was more salient than upwards motion, simply because of the former’s greater rate of occurrence in the natural environment.

Listeners were tested individually. The listener initiated a trial by pressing the ‘enter’ button on a keyboard that rested on her/his lap. The train of 8 clicks was presented by the computer, and the listener rated the strength of the illusion (“rate the continuity of the perceived motion”, see below) on a 5-point scale, by pressing a button from 1 to 5 on the keyboard. Listeners were permitted to re-listen to a given trial repeatedly before registering their responses. 1 s after the subject entered a response, the next trial was presented automatically.

The task of the listener on each trial was to rate the ‘evenness’ or ‘continuity’ of the perceived motion. The listener was instructed to rate as ‘1’ a percept in which the 8 clicks were equidistantly spaced in elevation, and as ‘5’ a percept of 4 clicks at one (top or bottom) speaker followed by 4 clicks at the other end of the speaker stack. The listeners were instructed to assign intermediate values on the rating scale according to the degree to which they perceived the eccentricities of the clicks to be unevenly spaced, ie to be intermediate between purely even, continuous motion (‘1’) and no continuous motion (‘5’). Listeners who had not used this scale in previous studies of auditory saltation practised using the rating scale with real and saltatory dichotic click trains that were perceived to move across intracranial azimuth.
3 Results
Ratings of the perceived motion continuity were collected on each trial. These values were averaged over the 10 repetitions, and then plotted as a function of ICI, with trial type (`saltation' or `real') and direction (up or down) as the parameters. These data are shown in figure 1, separately for each listener in the upper six panels, and averaged across listeners (with standard error bars) in the bottom left panel. Ratings for ‘real’ trials (open symbols) were almost uniformly near 1, and therefore independent of ICI, for all listeners. Ratings for the ‘saltation’ trials (filled symbols) showed a sigmoidal shape as a function of ICI: ratings were usually near 5 for long ICIs (120 – 480 ms),

![Graph showing motion ratings](image)

**Figure 1.** Upper six panels show mean (±SE) ratings of motion continuity (1, clicks spatially continuous over time; 5, clicks spatially discontinuous) separately for each listener, plotted as a function of inter-click interval (ICI), with trial type (`saltation' versus `real'-motion) and direction of location change (up versus down) as the parameter. Grand mean data, collapsed across listeners, are shown in the panel at bottom left.
indicating poor saltation, and near 1 for short ICIs (30–90 ms), indicating good saltation. Ratings were a graded function of ICI at intermediate values of that variable. Intra-individual variability was small, as shown by the standard error bars for the individual listeners’ data depicted in figure 1.

These qualitative observations were confirmed by a 3-factor (‘type’, ‘direction’ and ICI), repeated-measures, multivariate analysis of variance. There was a significant main effect of ‘type’ of motion ($F_{1,5} = 50.87$, $p < 0.001$), but not ‘direction’ of motion ($F_{1,5} = 1.62$, $p < 0.26$, ns) or ICI ($F_{5,1} = 56.13$, $p > 0.1$). Motion ‘direction’ did not interact with any other variable (all $p$s > 0.4). As expected, there was a significant interaction between the ‘type’ of motion and ICI ($F_{5,1} = 689.764$, $p < 0.03$), consistent with ratings for saltation stimuli being dependent on ICI and those for real-motion trials being independent of ICI.

4 Discussion
Auditory saltation is a misperception of the spatial location of stimuli. It occurs when there are short intervals between the members of a stimulus train in which the early members of the train occur at one location, and the later elements occur at a second location. The form of the misperception is that the clicks are perceived as originating not only from the actual stimulus locations, but also from points between them. The mechanisms that give rise to the percept are unclear. Following Shore et al (1998), one hypothesis is that under the temporal processing demands of short ICIs, the clicks are less likely to be heard as separate events and assigned their own (veridical) spatial locations; rather, the clicks are ‘grouped’ or ‘streamed’ as a single auditory object whose spatial location shifts between the beginning and the end of the stimulus train (after Bregman 1990; Phillips et al 2002). Perceptual interpretation of the click train as a single moving auditory object prompts the attribution of erroneous locations to the middle members of the train. That is, when the click train is heard as a single continuous stream or object, it is more likely to be heard as moving smoothly from one location to another. Conversely, when each click is heard as an individual event (as occurs at longer ICIs), each click may be heard at a unique location, uninfluenced by the location of the later, independent clicks.

Short ICIs provide the situation in which, during the course of the perceptual elaboration of one click in the train, the initiation of the processing of the next click is most able to influence or contaminate the processing of the earlier click [following the multiple-drafts model of Dennett (1991)]. The window width for this contamination varies between listeners, but usually is less than 400 ms (Phillips and Hall 2001). It is when successive clicks fall within this window that clicks late in the saltation train may ‘draw’ earlier clicks in the same train towards them. These considerations suggest that it is the interaction between the processing of temporally proximate events caused by the temporal processing demand that in turn causes the perceptual grouping, rather than the perceptual grouping driving a top–down assignation of erroneous locations to the middle clicks of the trains. In this regard, click stimuli may be particularly prone to this kind of spatial mislocalisation, because their extreme brevity makes them somewhat difficult to localise in the vertical plane in the first place (after Hofman and Van Opstal 1998).

The present study showed that robust saltation can occur in the elevation plane, and that there was no directional (up versus down) asymmetry in responses to the saltation stimuli. The absence of a directional asymmetry may be interesting from the standpoint that one might have anticipated a greater salience of, or sensitivity to, stimuli falling in the line of gravity than rising against it. This is because the natural world provides far more examples of the former than of the latter. The present study confirmed the effect of ICI on the ratings of motion continuity for saltation trials,
consistent with that observed for studies of this illusion in the horizontal plane, or across intracranial space. In the present study, it was the shortest ICIs (30–90 ms) which most strongly supported the illusion. In our previous study, the range of ICIs that supported the illusion was slightly longer (30–120 ms; Phillips and Hall 2001). This difference likely reflects that the present study used 8-click trains, while the earlier study used 6-click trains. Note that the present results are consistent with those of Shore et al (1998), who used 8-click trains. The temporal window over which this distortion of spatial perception occurs shows some individual differences but is usually less than about 350–400 ms wide (Phillips and Hall 2001). It follows that the temporal smearing need include only 1 click at the disparate location in order to smear the location of the preceding clicks, i.e., inclusion of the fifth click (first click at a different location) in the window will affect the earlier clicks, assuming the temporal window is sliding in time.

The fact that listeners gave similar ratings to ‘real’ and ‘saltation’ stimuli for short ICIs does not necessarily mean that listeners could not discriminate between the two trains at those ICIs. The rating-scale approach to describing the illusory percept is a useful but subjective method. Development of objective measures (e.g., signal-detection-theory approaches) to study the illusion is required to determine whether the ‘real’ and ‘saltation’ stimuli are discriminable. To date there has been one objective measure of the auditory-saltation illusion (Kidd and Hogben 2004; see Cholewiak and Collins 2000 for a tactile example), and this is an important future direction for studies of this illusion. Failure to discriminate between the two trains would suggest that neural representations of the two stimuli to which the perceptual processor had access are indistinguishable.

This raises the question whether the ratings were biased by factors other than the continuity of the motion percept. The fact that ratings for ‘real’-motion trials were independent of ICI (figure 1) suggests that the listeners were capable of separating the motion quality of the percept from overall train duration or click rate. Listeners may have employed simple rules based on the rate and activation of the middle speakers. For example, if the rate is fast, the listener responds 1 or 2; but if the rate if slow, the listener checks for the activation of the middle speakers. If activation is detected, the listener responds 1 for real motion; otherwise the listener responds 5. Note however, that selective scrutiny of the activation of the middle speaker would require the clicks to be heard as separate events, and that this is likely only possible at ICIs long enough that they fail to support the illusion. It is possible that the pitch or timbre of clicks influenced listeners’ ratings. There is little doubt that monaural localisation is based on spectral filtering of the source stimulus by the head and pinnae (Wightman and Kistler 1997; see also Blauert 1969/1970; Middlebrooks 1992) and specifically that localisation in the vertical plane is associated with orderly shifts in perceived spectral quality as source elevation changes (Roffler and Butler 1968; Butler and Belendiuk 1977). Because spectral filtering is highly correlated with source elevation, it is arguable whether perceived pitch (or timbre) can be isolated from perceived elevation in an experimental design like that used here. In this regard, however, we do not know of any evidence that auditory salutation-like phenomena (i.e., orderly misperceptions, within a stimulus train, of element location along some physical dimension) occur in continua other than spatial location. Future studies could examine whether a saltatory-like illusion occurs for stimuli of different pitches (either spectral or temporal) which might provide evidence for the supramodal characteristic of the effect in the auditory domain. The fact that salutation occurs also in somatosensation and in non-foveal vision, i.e., in all spatial senses, suggests that the illusion acts on a high-level spatial representation rather than emerging from constraints of neural coding at the periphery (Geldard 1975).
The major contribution of the present study is the demonstration that auditory saltation can occur for stimuli in a spatial dimension that provides only monaural cues for source location. This suggests that (i) the illusion either operates similarly on low-level spectral cues and on low-level binaural cues, or it arises at a central level of spatial processing after both the monaural and binaural spatial cues are combined, (ii) saltation should be demonstrable in monaural listeners, and (iii) saltation in dichotic studies is not a simple consequence or epiphenomenon specific to the processing of the change in interaural disparity that occurs mid-train.

Acknowledgments. This work was supported by grants from NSERC of Canada to DPP and fellowships from NSERC of Canada and the Killam Trust to SEB. Special thanks to Susan E Hall for excellent technical assistance, and to three reviewers for thoughtful commentaries on previous versions of the manuscript.

References
Dennett D C, 1991 *Consciousness Explained* (Boston, MA: Little, Brown)
Geldard F A, Sherrick C E, 1986 “Space, time and touch” *Scientific American* **255**(1) 90 – 95
Conditions of use. This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.