The subjective duration of audiovisual looming and receding stimuli

Massimo Grassi · Andrea Pavan

Published online: 1 June 2012
© Psychonomic Society, Inc. 2012

Abstract Looming visual stimuli (log-increasing in proximal size over time) and auditory stimuli (of increasing sound intensity over time) have been shown to be perceived as longer than receding visual and auditory stimuli (i.e., looming stimuli reversed in time). Here, we investigated whether such asymmetry in subjective duration also occurs for audiovisual looming and receding stimuli, as well as for stationary stimuli (i.e., stimuli that do not change in size and/or intensity over time). Our results showed a great temporal asymmetry in audition but a null asymmetry in vision. In contrast, the asymmetry in audiovision was moderate, suggesting that multisensory percepts arise from the integration of unimodal percepts in a maximum-likelihood fashion.

Keywords Looming · Receding · Audiovisual · Maximum likelihood integration

The traditional approach to perception investigates one sense at time (Fechner, 1889). This approach is useful for understanding the functioning of singular senses; however, it does not take into account that perception is inherently a multimodal process. For example, when looking at a fireplace,

we simultaneously see the fire, hear the noise of the wood burning, and feel the heat of the fire on our skin. The multimodal approach to perception has led to questions that are still largely unanswered, such as how the percepts from the different senses are integrated to form a coherent picture of the sensory world (see Calvert, Spence, & Stein, 2004, for an overview). This is particularly true when studies investigate perceptual continua that do not belong to a particular sense (e.g., a sound's pitch or an object's brightness) but that, on the contrary, are shared by all senses, such as the subjective duration of an event.

We can experience duration with all senses. However, a duration-related phenomenon in one sense may be similar, may disappear, or may be exaggerated or reduced in a different sense. And likewise, it may be similar, may disappear, or may be exaggerated or reduced when all of the senses are stimulated at once. For example, the subjective duration of a sound of increasing intensity (e.g., a stimulus compatible with the proximal stimulation generated by a looming sound source) is longer than the subjective duration of the same sound reversed in time (i.e., a stimulus compatible with the proximal stimulation generated by a receding sound source; see DiGiovanni & Schlauch, 2007; Grassi, 2003, 2010; Grassi & Darwin, 2006; Schlauch, Ries, & DiGiovanni, 2001). A similar temporal asymmetry has been recently observed by van Wassenhove, Buonomano, Shimojo, and Shams (2008). These authors found that the subjective duration of a disc linearly increasing in size (a stimulus compatible with the proximal stimulation of a looming object) is longer than that of a disc linearly decreasing in size (a stimulus compatible with the proximal stimulation of a receding object) when these objects are presented within a repeated temporal string of fixed-size discs (the so-called “oddball paradigm”; see Tse, Intriligator, Rivest, & Cavanagh, 2004). In a subsequent study, Wittmann, van

M. Grassi (✉)
Dipartimento di Psicologia Generale, Università di Padova,
Via Venezia 8,
35131 Padova, Italy
e-mail: massimo.grassi@unipd.it

A. Pavan
Institut für Psychologie, Universität Regensburg,
Universitätsstr. 31,
93053 Regensburg, Germany

Wassenhove, Craig, and Paulus (2010)¹ replicated this result but also observed that the above difference in subjective duration may arise because of a response bias for looming stimuli: Participants simply respond “longer” to looming discs, regardless of the disc’s duration. In sum, although a difference in subjective duration between looming and receding stimuli may exist in vision, too, this difference is not as clear as that observed in audition.

However, the similarities between audition and vision in the perception of looming and receding stimuli are not limited to subjective duration. For example, on the auditory side, looming stimuli are louder (Stecker & Hafter, 2000), change more in loudness (Neuhoff, 1998, 2001; Olsen & Stevens, 2010; Olsen, Stevens, & Tardieu, 2010; Pastore & Flint, 2011; for long durations, see Canévet, 1986; Canévet & Scharf, 1990; Canévet, Teghtsoonian, & Teghtsoonian, 2003; Schlauch, 1992; Susini, McAdams, & Smith, 2002, 2007; Teghtsoonian, Teghtsoonian, & Canévet, 2000, 2005), and have a stronger tonal quality (Patterson, 1994a, 1994b; Patterson & Irino, 1998) than do receding sounds. Similar results can be observed with primates, showing that the asymmetry between looming and receding sounds is shared across species (e.g., Ghazanfar, Neuhoff, & Logothetis, 2002; Lu, Liang, & Wang, 2001; Maier, Neuhoff, Logothetis, & Ghazanfar, 2004). On the visual side, looming (in comparison to receding) visual objects induce fear responses (Schiff, Caviness, & Gibson, 1962) and attract attention more efficiently (Franconeri & Simons, 2003; von Mühlenen & Lleras, 2007). In synthesis, the auditory and visual results converge: A looming stimulus is a perceptually more salient stimulus than a receding one (e.g., Hall & Moore, 2003; van Wassenhove et al., 2008).

How could the auditory and visual subjective durations be integrated when looming and receding stimuli are presented audiovisually? The literature suggests at least three different answers. According to Neuhoff (e.g., 1998, 2001), the marked perceptual asymmetry between looming and receding sounds might reveal an evolutionary advantage. For example, the overestimation of the auditory characteristics of a looming object may increase the margin of safety of the perceiver in response to the approaching sound source (Neuhoff, 1998, 2001). Along this line, the extended subjective duration of looming (in comparison to receding) sounds may contribute, together with greater loudness and the change in loudness, to making these sounds perceptually more salient than receding sounds. Coherently with Neuhoff’s explanation, some literature findings have revealed

further auditory asymmetries with regard to such events, as well as further possible evolutionary explanations, such as occasional sex differences in the perception of these auditory stimuli (Grassi, 2010; Neuhoff, Planisek, & Seifritz, 2009). How would the evolutionary explanation advanced by Neuhoff apply to the subjective durations of looming and receding audiovisual objects? If we assume that a looming overestimation is generally advantageous for the perceiver, because a similar asymmetry exists in vision, we might expect to observe the same perceptual asymmetry (or an exaggerated one) with audiovisual looming versus receding stimuli. Consistent with this prediction, recent findings have shown that audiovisual looming stimuli induce faster motor reactions than do unimodal (i.e., either auditory or visual) looming or receding stimuli (Cappe, Thut, Romei, & Murray, 2009), suggesting a possible additive integration of auditory and visual signals when audiovisual stimuli are presented.

However, the above finding is not necessarily the outcome predicted by current and past theories of multimodal binding. For example, the so-called “modality appropriateness hypothesis” (Rock & Victor, 1964; Welch & Warren, 1980) predicts the dominance of the auditory percept within the audiovisual percept. In its general form, the appropriateness hypothesis states that observers choose to weight the various modality inputs according to their unimodal performance capabilities with respect to a given task. From this understanding, vision should strongly dominate touch in spatial tasks, which itself should dominate audition, because vision is spatially best and audition is spatially worst (Myers, Cotton, & Hilp, 1981; Over, 1966). Conversely, audition should strongly dominate vision on temporal tasks, because these tasks are performed best with audition and worst with vision (e.g., Grondin, Meilleur-Wells, Ouellette, & Macar, 1998; Penney, 2003; Westheimer, 1999). In brief (and caricaturizing): Regardless of the subjective durations of visual looming and receding objects, the audiovisual percept should be driven by audition. In other words, the subjective durations of looming and receding audiovisual stimuli should be identical to the subjective durations of looming and receding auditory objects.

A third possible prediction comes from the recent model suggested by Ernst and Banks (2002; but see also Alais & Burr, 2004). This model would lead us to expect different (and perhaps mixed) findings. Overall, the model states that the perceiver may operate like a maximum-likelihood integrator that minimizes the variance of the multimodal estimates. The model predicts that the multimodal estimate depends on the variance of the unimodal estimates: If the variance is large, the sense will contribute slightly to the multisensory estimate, but if the variance is small, the sense will contribute greatly to the multisensory estimate. In sum, because audition is generally better than vision in estimating

¹ Wittmann et al.’s (2010) study addressed the perception of looming and receding stimuli in audition, vision, and audiovisual. However, the sound stimulus that the authors used was not looming (or receding) compatible: It was an upward (or downward) frequency sweep (Neuhoff & McBeath 1997), which the authors improperly named as a “looming auditory signal” (or a “receding auditory signal”).

duration (Grondin et al., 1998; Penney, 2003; Westheimer, 1999), the audiovisual estimate should be driven by audition. However, since in the present study the participants evaluated the durations of three different stimuli (i.e., looming, receding, and stationary), each with a different (and in several cases unknown) variance, it would be difficult to predict the precise outcome of the audiovisual estimate of each stimulus. For example, the short subjective duration of receding sounds is thought to depend on cognitive (i.e., not on sensory) factors (DiGiovanni & Schlauch, 2007): Although listeners perceive the entire duration of looming and receding sounds, they exclude the ending portion of receding sounds from the computation of subjective duration because they interpret the tail of the sound as an echo. These cognitive factors may introduce a large variance in the estimates of these particular stimuli. If, in contrast, the variance of the estimate of visual receding stimuli is small, the audiovisual estimate of the audiovisual receding stimulus should be driven by vision. After the necessary changes, the same argument can be applied for the prediction of the subjective estimates of audiovisual looming and stationary stimuli. This model enables a point-by-point comparison between predicted and empirical results.

In the present experiments, we investigated which of these hypotheses better predicts the subjective durations of audiovisual looming and receding stimuli. Participants were asked to estimate the subjective durations of looming, receding, and stationary stimuli 0.5, 1, and 2 s in duration. The looming and receding stimuli simulated as closely as possible the proximal stimulation induced by a sound source emitting a steady-amplitude sound that moves at a constant speed toward (or away from) a perceiver. The stationary stimulus was a stimulation that did not change (in sound pressure level or size) over time and served as a control condition with which to assess the estimations of the looming and receding stimuli. The stimuli were either auditory (a tone log-linearly changing in amplitude over time), visual (a disc log-linearly changing in diameter over time), or audiovisual (i.e., a stimulus obtained by presenting simultaneously the auditory and visual stimuli).

Experiment 1

Method

Participants A group of 30 participants (11 males, 19 females) took part in the experiment. The participants sat in a dark and silent (below 35 dBA at the listener's ear) room and were placed 120 cm from the screen and the speakers. Viewing was binocular, and listening was binaural. All of the participants had normal or corrected-to-normal visual acuity and normal hearing. The participants

took part voluntarily, with no compensation, and gave their informed consent prior to inclusion in the experiment. None of the participants was informed about the hypothesis driving the experiment.

Apparatus We wrote our experiments in MATLAB (The MathWorks, Natick, MA) using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). The software was run on a Pentium IV computer connected to a NEC Multisync FP950 monitor (100-Hz refresh rate). The screen resolution was $1,024 \times 768$ pixels, with each pixel subtending ~ 1 arcmin. Sounds had a sample rate of 44.1 kHz and 16-bit resolution. The output of the sound card (M-AUDIO Fast Track Pro) was passed to two M-AUDIO Studiophile AV 30 amplified speakers. The speakers were placed at the left and right of the monitor's sides, and the drivers were aligned with the monitor's horizontal midline, so that the combined sound wave originated from the monitor's center.

Stimuli The stimuli consisted of looming, receding, or stationary auditory, visual, and audiovisual stimuli 0.5, 1, or 2 s in duration. These stimuli were created by simulating the proximal stimulation pattern produced by the motion of an object emitting a steady-amplitude sound and moving at a constant speed toward the perceiver. An object as such generates changes in retinal size and in intensity at the listener's ear that are of an identical magnitude and linear on a logarithmic scale. Note that a sound source as such does not produce any audible frequency shift at the perceiver's ear (Neuhoff & McBeath, 1997). In the present experiment, the change in the object's proximal characteristics over time (i.e., intensity at the listener's ear and retinal size) was set to one hundred. Therefore, the sound's intensity and the size of the visual stimulus changed of 20 dB over the stimulus's duration.

The auditory stimulus was synthesized from a harmonic complex tone of 0.5, 1, or 2 s that was obtained by adding ten sinusoids with frequencies corresponding to Harmonics 2–10 of a fundamental sinusoidal tone of 250 Hz. All of the sinusoids had identical phase and amplitude. The tone's envelope could be either ramped (looming tone), damped (receding tone), or steady (stationary tone) in sound pressure level. Looming sounds were obtained by modulating the amplitude of a steady-amplitude tone with an exponential ramp that decreased in pressure by 40 dB over the tone's duration (thus, 20 dB in intensity, because $I=P^2$). The resulting tone was gated on and off, with two 5-ms raised cosine ramps. The receding tone was a looming tone played backward in time. The stationary tone was a steady tone not modulated in amplitude. The peak pressure level of the looming and the receding tone was of 95 dBA. The pressure level of the stationary tone was 75 dBA—that is, simulating

the proximal stimulation of a sound source placed midway between the starting point and the ending point of the looming (or receding) object.

The visual stimulus was a black disc (0.09 cd/m^2) presented over a white background (67 cd/m^2). The disc's diameter was either increasing (looming disc), decreasing (receding disc), or constant (stationary disc) over a time of 0.5, 1, or 2 s. The starting size of the looming disc was 0.13 deg , the final size was 12.9 deg , and the change in the disc's diameter over time was logarithmic. The receding disc was a looming disc reversed in time. The diameter of the stationary disc was fixed at 1.28 deg —that is, as for the auditory stimulus, simulating the proximal stimulation of an object placed midway between the starting point and the ending point of the looming (or receding) disc.

The audiovisual stimulus was obtained by presenting simultaneously the auditory looming (or receding, or stationary) and the visual looming (or receding, or stationary) stimulus. In order to prevent possible errors in the timing of the stimuli, the custom software that was used to generate the stimuli was identical in all conditions (i.e., auditory, visual, and audiovisual). In the auditory condition, the disc was white so that it could not be distinguished from the background. In the visual condition, the speaker's volume was turned off.

Procedure The experiment was divided into nine blocks that were performed within a maximum of 5 days by the participants. In each block, the participant was asked to match the duration of a duration-adjustable stationary stimulus (either auditory, visual, or audiovisual) to that of a fixed-duration looming, receding, or stationary target (auditory, visual, or audiovisual). During audiovisual blocks, the participant was explicitly asked to match the duration of the whole audiovisual event. In each block, a trial consisted of the fixed-duration target stimulus (looming, receding, or stationary), always followed after 500 ms by the duration-adjustable stationary stimulus whose duration listeners had to match to that of the target. The duration-adjustable stationary stimulus was similar to the stationary target but had an adjustable duration that the participant could manipulate with the mouse. At the beginning of each trial, the duration of the adjustable stationary stimulus was set randomly within a range of $\pm 80 \%$ of the duration of the target stimulus (this range was also the maximum permitted adjustment during each experimental trial). The target-adjustable pair was then presented to the participant. At this point, the participant could manipulate the duration of the adjustable stimulus with the mouse. A mouse click in the upper half of the screen imposed an increment to the duration of the adjustable stimulus, whereas a click in the

lower half of the screen imposed a decrement to the duration of the adjustable stimulus. After the click, the fixed-duration target was presented to the participant, followed by the duration-adjusted stimulus, so that the participant could test whether or not the target and the adjustable stimulus were similar in duration. If the participant was satisfied, he or she could save the trial and move on to the next trial. If not, the process above was iterated until the participant was satisfied.

Each experimental block contained looming, receding, and stationary targets of the same type (i.e., auditory, visual, or audiovisual) and duration. Within each experimental block, listeners matched the durations of 30 targets presented in random order: ten looming, ten receding, and ten stationary. Listeners encountered the experimental blocks in a nested randomized order. Experimental blocks with targets of the same overall duration were taken consecutively in order to obtain more consistent adjustments (Wright, Buonomano, Mahncke, & Merzenich, 1997), and the order in which participants took the different types of stimulation (i.e., auditory, visual, and audiovisual) was counterbalanced across participants. Each block lasted about 30 min, for a total of about 4 h, or half of the experiment for each participant. The adjustment method used in the present experiment returns a direct estimate of the participant's point of subjective equality (PSE).

Results

Individual participants' PSEs were calculated for each type of motion and separately for type of stimulation (i.e., auditory, visual, or audiovisual) and duration. Successively, outliers (i.e., PSEs exceeding ± 1 standard deviation of the participant's mean) were discarded, and PSEs were recalculated on the basis of the remaining data. PSEs were transformed into estimation errors according to the following equation:

$$E\% = 100 \left(\frac{PSE - d_\phi}{d_\phi} \right) \quad (1)$$

where $E\%$ is the participant's estimation error expressed as a percentage, and d_ϕ is the actual stimulus's duration. This transformation enables us to appreciate in a clear way how the participant's estimation changes across durations (e.g., Grassi & Darwin, 2006). Positive values of the estimation error indicate that the target stimulus is overestimated in comparison to the actual stimulus's duration, whereas negative values indicate that the target stimulus is underestimated in comparison to the actual stimulus's duration. Moreover, because the above transformation is linear, it

does not affect the comparison between the empirical results and the results predicted by the Ernst and Banks (2002) model (see below). Estimation errors are represented in Fig. 1.

Because the literature reports occasional sex differences in the perception of looming and receding sounds (Grassi, 2010; Neuhoﬀ et al., 2009), estimation errors were first subjected to a 3 (sensory modality: auditory, visual, or audiovisual) $\times$ 3 (duration: 0.5, 1, or 2 s) $\times$ 3 (type of motion: looming, receding, or stationary) four-way ANOVA (with Greenhouse–Geisser correction), with Sex of the Participant as a between-subjects factor. The sex of the participant did not affect the estimation error, either as a main effect or in interaction with the other factors of the ANOVA (all $ps > .05$). The sex of the participant was therefore excluded from the analysis, and the ANOVA was recalculated. Overall, estimation error was independent of the sensory modality, $F < 1$. However, it changed across durations, $F(2, 58) = 4.00$, $p = .027$, and, in particular, was different for the three types of motion simulated in the experiment, $F(2, 58) = 35.51$, $p < .0001$ —that is, the estimation error was negative (i.e., underestimation of duration) for receding stimuli, whereas it was null for looming and stationary stimuli. The two-way interactions revealed no difference in the estimation error across durations for the different sensory modalities, $F(4, 116) = 1.67$, $p = .17$. In contrast, we observed an interaction between stimulus duration and type of motion, $F(4, 116) = 5.67$, $p < .0001$, with larger differences in the estimation errors of looming, receding, and stationary stimuli at short durations, and smaller differences at long durations. In addition, the interaction between sensory modality and type of motion was also significant, $F(4, 116) = 7.95$, $p < .0001$, showing that the underestimation of receding stimuli in comparison to looming and stationary stimuli was large in audition, small in audiovision, and null in vision. The three-way interaction was not significant, $F < 1$.

In order to have a closer look at the looming/receding temporal asymmetry the matched durations collected with looming and receding targets were converted into a normalized error, $E_n\%$, which expresses the under- or overestimation of the

looming (or receding) target normalized to the subjective duration of the stationary target:

$$E_n\% = 100 \left(\frac{PSE_{lr} - PSE_{st}}{PSE_{st}} \right) \quad (2)$$

where $E_n\%$ is the participant's normalized estimation expressed as a percentage, PSE_{st} is the PSE of the stationary target, and PSE_{lr} is the PSE of the looming (or receding) target. Positive values of $E_n\%$ indicate that the stimulus was overestimated in comparison to the subjective duration of the stationary control target, whereas negative values indicate that the stimulus was underestimated in duration in comparison to the subjective duration of the stationary control target (see Fig. 2).

The $E_n\%$ s were subjected to a 3 (type of stimulation) $\times$ 3 (duration) $\times$ 2 (type of motion) three-way ANOVA. The ANOVA revealed no main effect of sensory modality, $F(2, 56) = 2.84$, $p = .07$, but main effects of both duration and motion type: respectively, $F(2, 56) = 4.86$, $p = .02$, and $F(1, 28) = 37.96$, $p < .0001$. $E_n\%$ changed as a function of duration, and in particular, it was smaller for receding than for looming stimuli, demonstrating the general underestimation of the receding stimulus in comparison to the looming stimulus. The sensory modality by duration and the type of motion by duration interactions were not significant: both $Fs \leq 2.33$, both $ps \geq .09$. On the contrary, the sensory modality by type of motion interaction was significant: $F(2, 56) = 12.23$, $p < .0001$. This result reveals that the perceptual asymmetry in subjective durations between looming and receding stimuli is large in audition, moderate in audiovision, but null in vision, where both stimuli had similar subjective durations. The three-way interaction was not significant, $F < 1$.

We calculated the looming/receding asymmetry by subtracting the $E_n\%$ for looming from that for receding stimuli, separately for each modality and participant, and independently of the stimulus's duration. This operation provided a net estimate (expressed in the form of a percentage) of the subjective difference between looming and receding stimuli. The differences were subjected to a three-level (sensory

Fig. 1 Experiment 1: Estimation error ($E\%$) as a function of the duration of the target. The horizontal dotted lines represent perfect match (i.e., no error in the estimations). The data points are slightly offset to improve legibility. Vertical bars represent ± 1 standard error of the means

Fig. 2 Experiment 1:

Normalized estimation error ($E_n\%$; see the text for details) as a function of the stimulus's duration. The horizontal dotted lines represent the subjective estimates of the stationary stimulus. The data points are slightly offset to improve legibility. Vertical bars represent ± 1 standard error of the means

modality) one-way ANOVA. The result was significant, with different looming/receding asymmetries in audition ($M = 16.1\%$, $SD = 11\%$), audiovision ($M = 9.3\%$, $SD = 10\%$), and vision ($M = 3.4\%$, $SD = 12\%$): $F(2, 58) = 13.18$, $p < .0001$. The results are represented in Fig. 3.

Testing of the maximum-likelihood integrator Ernst and Banks (2002) suggested that a perceiver may act as a maximum-likelihood integrator that, in the estimation of a multimodal stimulus, minimizes the variance of his or her multimodal estimation. The estimation $\hat{E}$ of a multimodal stimulus would be equal to the sum of the unimodal estimates weighted by a weight w that is proportional to the variances associated with the unimodal estimates. In summary, in the present experiment, the multimodal estimate $\hat{E}_{av}$ would be equal to:

$$\hat{E}_{av} = w_a \hat{E}_a + w_v \hat{E}_v$$

where $\hat{E}_a$ and $\hat{E}_v$ are, respectively, the auditory and visual estimates, together with their respective weights w_a and w_v .

Fig. 3 Experiment 1: Looming/receding asymmetries in audition, vision, and audiovision. The vertical bars represent ± 1 standard error of the means

The weights are proportions that enable us to assess the relative contribution of each sense involved in the multimodal estimate. The weight w ranges from zero to unity. If one unimodal estimate had a null variance, it would uniquely contribute to the final multimodal estimation. On the contrary, if both unimodal estimates had identical variances, both senses would contribute equally to the resulting multimodal estimate. The weights (e.g., w_a) can be calculated as follows:

$$w_a = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_a^2}$$

where σ^2 (for either a or v) is the variance of the (auditory or visual) unimodal estimate.

Equations 1 and 2 were used with the unimodal $E\%$ s and unimodal $E\%$ variances to calculate the unimodal weights w_a and w_v and the predicted audiovisual estimates $\hat{E}_{av}$. The calculation was performed for each participant and separately for each target stimulus (either auditory or visual) presented in the experiment. Successively, the predicted audiovisual estimates $\hat{E}_{av}$ were compared to the empirical audiovisual estimates $E\%$ by means of a 3 (motion type) $\times$ 3 (durations) three-way ANOVA, with Data Source (i.e., predicted or empirical) as a between-subjects factor. The ANOVA revealed no difference between the empirical and predicted $E\%$ s, either as a main effect ($F < 1$) or in all interactions (all F s < 1). The lack of difference between the empirical results and the results predicted by the model was corroborated by a correlation performed on mean estimates, $r(9) = .99$, $p < .0001$. In Fig. 4, empirical audiovisual $E\%$ s are represented as a function of predicted audiovisual $\hat{E}\%$ s.

Discussion

The results of the present experiment show that the subjective duration of receding sounds is shorter than the subjective duration of looming sounds. In contrast, with visual stimuli, the results revealed no difference between the subjective

Fig. 4 Experiment 1: Empirical audiovisual estimates ($E\%$) as a function of the predicted audiovisual estimates ($\hat{E}_{av}$). The vertical bars represent ± 1 standard error of the means of the empirical estimates. The dotted line represents perfect match between the predicted and empirical estimates, and the upper and lower dashed lines represent ± 1 standard error of the predicted estimates

durations of looming and receding stimuli. Finally, the result observed with audiovisual stimuli was intermediate between those observed with auditory stimuli, on the one side, and visual stimuli, on the other: The subjective duration of receding audiovisual stimuli was shorter than the subjective duration of looming stimuli, although the difference was not as large as that observed in audition.

In audition and audiovision, the difference in subjective duration between looming and receding stimuli was dependent on the duration of the stimulus: For short-duration stimuli (e.g., 0.5 s), the difference was large, but for long-duration stimuli (e.g., 2 s), the difference was small/null. This result hides a possible confound that could prevent a correct interpretation of the results. In the present experiment, the stimulus duration covaried with the velocity of the represented motion. For example, the 0.5-s looming stimulus lasted half of the duration of the 1-s looming stimulus. However, the former stimulus also represented the proximal stimulation of an object that traveled at a speed that was twice that of the latter stimulus. In brief, the results do not enable us to disentangle whether the effect observed for stimulus duration can be attributed exclusively to duration or, in contrast, to the represented object's speed, or to both.

In the next experiment, the looming and receding stimuli represented the motion of objects moving at two different velocities (one twice the other) for each of the durations

investigated in Experiment 1. Experiment 2 would enable us to observe separately the effect of the stimulus duration and the effect of the velocity represented by the stimulus.

Experiment 2

The results of Experiment 1 revealed that receding sounds are perceived as shorter than looming sounds. Moreover, receding audiovisual stimuli are perceived as shorter than looming audiovisual stimuli (although the difference was not as large as that observed in audition). In contrast, in vision, looming and receding stimuli had similar subjective durations. In audition and audiovision, the difference in subjective durations between looming and receding stimuli decreased at increasing durations. In Experiment 1, however, stimulus duration and the speed of the object's represented motion covaried. Therefore, the results did not enable us to disentangle whether the effect observed for stimulus duration could be attributed exclusively to duration or, in contrast, to the object's represented speed, or to both. For this reason, in the present experiment, both the stimulus duration and the stimulus's represented speed were manipulated separately, in order to observe separately their effects on subjective duration.

Method

A group of 12 new participants (six male, six female) took part in the experiment. The apparatus was identical to that of Experiment 1. The stimuli were those from Experiment 1, together with one new looming stimulus and one new receding stimulus for each duration and sensory condition (i.e., auditory, visual, or audiovisual). For each duration (i.e., 0.5, 1, or 2 s), the new stimuli simulated the proximal stimulation of a sound-emitting object moving at half of the speed of the stimuli used in Experiment 1. The proximal stimulation pattern of the new stimuli was compatible with the motion of an object emitting a steady-amplitude sound, moving at constant velocity, and traveling half of the distance traveled by the object hypothesized to create the stimuli used in Experiment 1. In addition, the new stimulus simulated an object that, at half of its distance from the perceiver, was placed at a distance identical to that of the looming and receding objects used in Experiment 1. In other words, at half of the stimulus duration, the new and old looming and receding stimuli had identical acoustical and optical characteristics. In detail, the new looming (or receding) sounds changed by 20 dB over the sound's duration: from 65 to 85 dBA, or from 85 to 65 dBA (old stimuli: 55 to 95 dBA or 95 to 55 dBA). The new looming (or receding) disc changed size log-linearly over the stimulus duration from ~ 0.4 to ~ 4.0 deg or from ~ 4.0 to ~ 0.4 deg (old stimuli: 0.13 to 12.9 deg or 12.9 to 0.13 deg). The new

audiovisual stimuli were obtained by presenting simultaneously the new looming sound with the new looming disc or the new receding sound with the new receding disc. The procedure of the present experiment was identical to that used in [Experiment 1](#), with the exception that each stimulus was presented six times to the participant.

Results

Individual PSEs were calculated for each stimulus. Outliers (i.e., PSEs exceeding ± 1 standard deviation of the participant's mean) were discarded, and PSEs were recalculated on the remaining data. As in [Experiment 1](#), the PSEs were then transformed into estimation errors and successively into normalized estimation errors. A 3 (sensory modality: auditory, visual, or audiovisual) $\times$ 2 (type of motion: looming or receding) $\times$ 2 (velocity: fast or slow) $\times$ 3 (duration: 500, 1,000, or 2,000 ms) four-way ANOVA (with Greenhouse–Geisser correction) was calculated. The normalized errors were similar across the three sensory modalities, $F(2, 22) = 2.34$, $p = .14$; across the stimulus velocities, $F(1, 11) = 2.75$, $p = .12$; and across the stimulus durations, $F(2, 22) = 2.62$, $p = .096$. In contrast, they changed across different types of motion, with very different normalized errors for looming and receding stimuli: $F(1, 11) = 49.57$, $p < .0001$. As in [Experiment 1](#), the two-way interaction revealed that the looming/receding asymmetry was large in audition, smaller in audiovision, but null in vision, $F(2, 22) = 11.18$, $p = .003$, and that it decreased with increasing duration, $F(2, 22) = 5.80$, $p = .013$. In addition, the looming/receding asymmetry was larger for high-speed than for low-speed stimuli: $F(1, 11) = 38.65$, $p < .0001$. Finally, the type of motion by sensory modality by duration interaction unfolded the two-way interaction previously observed: The looming/receding asymmetry was large in audition, smaller in audiovision, but null in vision. In addition, the asymmetry (where it existed) decreased at increasing durations: $F(4, 44) = 3.22$, $p = .046$. The remaining interactions were not significant: all $ps > .11$ (see Fig. 5).

As in [Experiment 1](#), we calculated the looming/receding asymmetry by subtracting the $E_n\%$ for looming from that for

the receding stimuli, separately for each modality and participant and independently of the stimulus's duration and velocity. The differences were subjected to a one-way ANOVA with Sensory Modality as the factor. The result was significant, with the largest looming/receding asymmetry for audition ($M = 18.6\%$, $SD = 9\%$), an intermediate one in audiovision ($M = 11.0\%$, $SD = 6\%$), and a null one for vision ($M = 1.6\%$, $SD = 9\%$): $F(2, 24) = 13.21$, $p = .001$.

Testing of the maximum-likelihood integrator As in [Experiment 1](#), unimodal $E\%$ s and unimodal $E\%$ variances were used to calculate the unimodal weights w_a and w_v and the predicted audiovisual estimates $\hat{E}_{av}$. The calculation was performed for each listener and separately for each target stimulus (either auditory or visual) presented in the experiment, as in [Experiment 1](#). The predicted audiovisual estimates $\hat{E}_{av}$ were compared to the empirical audiovisual estimates $E\%$ by means of a 5 (stimulus type) $\times$ 3 (durations) three-way ANOVA with the Data Source (i.e., predicted or empirical) as a between-subjects factor. The ANOVA revealed no difference between the empirical and predicted $E\%$ s, both as main effects ($F_s < 1.71$, all $ps < .05$) and in all interactions (all $F_s < 1$). The lack of difference between the empirical results and the results predicted by the model was corroborated by a correlation performed on the mean estimates: $r(15) = .95$, $p < .0001$. In Fig. 6, empirical audiovisual $E\%$ s are represented as a function of predicted audiovisual $\hat{E}\%$ s.

Discussion

In [Experiment 2](#), we investigated whether the stimulus duration, the (simulated) stimulus speed, or both were responsible for the effects on subjective duration observed in [Experiment 1](#). Overall, the results of [Experiment 2](#) replicated the findings of [Experiment 1](#): namely, a large durational asymmetry between looming and receding stimuli in audition, a smaller asymmetry in audiovision, and a null asymmetry in vision. In addition, the results revealed that the looming/receding

Fig. 5 Experiment 2: Normalized estimation error ($E_n\%$; see the text for details) as a function of the stimulus's duration. The horizontal dotted lines represent the subjective estimates of the stationary stimulus. The data points are slightly offset to improve legibility. Vertical bars represent ± 1 standard error of the means

Fig. 6 Experiment 2: Empirical audiovisual estimates ($E\%$) as a function of the predicted audiovisual estimates ($\hat{E}_{av}$). The vertical bars represent ± 1 standard error of the means of the empirical estimates. The dotted line represents perfect match between the predicted and empirical estimates, and the upper and lower dashed lines represent ± 1 standard error of the predicted estimates

temporal asymmetry decreased at increasing durations (in audition and audiovisual, but not in vision), as well as that it increased at increasing simulated velocities (in audition and audiovisual, but not in vision). Finally, the audiovisual results of Experiment 2 also are consistent with those predicted by the model suggested by Ernst and Banks (2002).

General discussion

In the present study, we investigated how two similar temporal asymmetries in vision and audition merge when vision and audition are stimulated at once. In particular, we tested which of three alternative hypotheses (i.e., the evolutionary hypothesis, the modality appropriateness hypothesis, and the maximum-likelihood integrator hypothesis) of audiovisual binding better predicted the subjective durations of audiovisual looming and receding stimuli.

The subjective durations of auditory and visual looming and receding stimuli

The results gathered with auditory stimuli confirmed the findings presented in the literature (e.g., DiGiovanni & Schlauch, 2007; Grassi, 2003; Grassi & Darwin, 2006; Schlauch et al., 2001): Participants estimated accurately the duration of stationary sounds (errors within about $\pm 10\%$ of the sound's duration). In addition, they estimated

looming sounds to be as long as, or longer than, stationary sounds. On the contrary, they underestimated the duration of receding sounds. The difference in subjective durations between looming (or stationary) and receding sounds was large at short durations and became increasingly smaller at longer durations (Exps. 1 and 2), and it was larger for faster-changing stimuli than for slower-changing stimuli (Exp. 2). The net dimension of the looming/receding asymmetry is variable and depends on the contingent characteristics of the stimuli. For example, in Experiment 1 the asymmetry was about 20 % at the shortest duration and 10 % at the longest duration. In Experiment 2, in contrast, the asymmetry was about 40 % at the shortest duration and 10 % at the longest duration.

Why are receding sounds underestimated in duration, whereas looming and stationary are not? Overall, humans are generally accurate in judging the durations of auditory events. Classic literature studies (Eisler, 1975; Stevens, 1957) have shown that the psychophysical function for auditory duration estimation has an exponent close to unity—that is, a perfect match between physical and subjective durations. The “odd” result is therefore the short subjective duration of receding sounds. Maybe the underestimation arises because participants ignore the ending portion of the sound in the computation of its subjective duration (DiGiovanni & Schlauch, 2007; Grassi, 2003). Stecker and Hafer (2000) suggested that listeners might interpret the tails of receding sounds as echoes, and for this reason, they may exclude this “thought-to-be” echo portion from the computation of the subjective duration of the sound. With the appropriate changes, this simple explanation can also account for the difference in loudness (Stecker & Hafer, 2000) and for the change in loudness (Neuhoff, 1998, 2001) observed with looming and receding sounds. A cognitive explanation such as the echo hypothesis may also explain why the looming/receding asymmetry changes so much in magnitude from one condition to another. For example, we may speculate that here, in Experiment 2, the asymmetry was larger than in Experiment 1 because the presence of various and very different looming and receding stimuli maximized the perceptual differences between them, making the tails of fast-receding sounds more echo-like than the tails of slow-receding sounds (a similar effect can be observed in Grassi & Darwin, 2006). In synthesis, if the asymmetry has a cognitive origin, it may be likely to have a different magnitude according to the contingent characteristics of the stimuli used in the experiment.

A low-level explanation of the variability of the looming/receding asymmetry can be found in the way that the auditory system encodes the onset and offset of auditory stimuli, in particular, when encoding is matched together with the acoustical characteristics of these stimuli. The onset of auditory events is heavily marked in comparison to the offset.

This can be observed either via physiological recordings (see Phillips, Hall, & Boehnke, 2002, for an overview) or via behavioral experiments (e.g., Zera & Green, 1993). Given this fact, the onset of a looming sound is marked because the auditory system is onset-unbalanced, and the offset is clearly perceivable because it ends at a high acoustic intensity. The net result is that the looming stimulus has clear-cut onset and offset temporal edges. In contrast, the auditory system efficiently encodes the onsets of receding sounds but scarcely processes the offsets, because of the lack of effective offset encoding and because receding sounds end at a low acoustic intensity. As a result, the onset temporal edge of receding sounds is clear-cut, whereas the offset is not. A consequence of this asymmetric encoding is that the subjective duration of receding sounds is shorter than the actual duration of the sound. A further, higher-level hypothesis cannot be excluded, however. The looming/receding asymmetry may arise because of a confound: Participants may associate the increasing intensities of looming sounds with a “large” source (or largeness) and the decreasing intensity of receding sounds with a “small” source (or smallness; see Proctor & Cho, 2006, for this type of interpretation). As a consequence, receding sounds would be underestimated because they are associated with smallness, and by the same token, looming sounds would be overestimated in comparison to receding ones because they are associated with largeness.

The looming/receding asymmetry observed in audition disappears in vision, where looming and receding stimuli have similar durations. This result contrasts with the recent findings of van Wassenhove et al. (2008; but see also Wittmann et al. 2010). Several differences, however, do prevent a direct comparison between of the results obtained here and those of preceding studies. These differences are in the stimuli and in the method. On the stimulus side, van Wassenhove et al. and Wittmann et al. investigated only 500-ms-long stimuli. Therefore, it is unknown whether the temporal asymmetry that they observed can be observed at longer durations. Secondly, the present study and the studies above used different types of looming and receding stimuli. In both of the above studies, the change in size of the disc was linear (van Wassenhove, personal communication), whereas here it was logarithmic. The appearances of logarithmic and linear change in size are rather different. On the method side, the present study and the above studies are also different. Here, the subjective duration of the stimuli was estimated directly by means of the so-called “adjustment method” (Fechner, 1889), by providing a fixed-duration target and a duration-adjustable stimulus. In the previous studies, it was estimated indirectly by means of a binary response and by using the oddball paradigm (see Eagleman & Pariyadath, 2009, for an overview of this and of other paradigms in time perception). Van Wassenhove et al. and

Wittmann et al. presented to participants a set of five successive stimuli: four fixed-duration stationary discs, interleaved by a single variable-duration target presented in fourth position. Their participants simply reported whether the target disc was longer (or shorter) than the stationary discs. In the second of the studies above, Wittmann et al. tested whether the response could be affected by a response bias. They found that participants responded “longer” to looming targets (but not to receding and stationary ones), regardless of the target’s duration. Thus, the response to looming discs was affected by a response bias. Along the same line, van Wassenhove et al. reported the case of one participant (see the paper’s “Supporting Information” in the publisher website, Fig. 2 top graph) who systematically responded “longer” to all looming targets presented in the experiment, regardless of the stimulus’s duration. Note also that with the oddball paradigm, Tse et al. (2004) found that, if the fixed-duration discs were looming and the variable-duration disc was stationary, the subjective duration of the stationary disc resulted in being *longer* than that of the looming discs (their Exp. 2).

In brief, the present and the previous results on the subjective durations of looming and receding visual stimuli do not enable us to draw a clear picture. In other words, it is not clear whether the different subjective durations of looming and receding stimuli observed reflect a genuine (and presumably small) difference in the subjective durations or an idiosyncratic response bias, perhaps due to the saliency of a looming stimulus (Franconeri & Simons, 2003; Schiff et al., 1962; van Wassenhove, 2009; von Mühlenen & Lleras, 2007). In any case, we would like to stress one more datum: In the present study, the estimations of visual stimuli (and the estimations of 500-ms-long visual stimuli, in particular) were the most variable. Therefore, it may not be surprising that these stimuli would gather different results in different studies.

The subjective durations of audiovisual looming and receding stimuli

Audiovisual stimuli returned yet another different result. As in audition, the durations of audiovisual looming and stationary stimuli were estimated veridically or almost veridically. Moreover, as in audition, the durations of receding audiovisual stimuli were underestimated in comparison to those of both stationary and looming objects. However, the net size of the looming/receding asymmetry was half of that observed in audition: In audition, the looming/receding asymmetry was about 15 %–20 %, but in audiovision the figure was reduced to 7 %–10 %—that is, halfway between the effects observed in audition and in vision (i.e., no asymmetry). The looming/receding asymmetry was moderate at the shortest duration and at the highest represented object’s speed, but it was null at the longest duration and at the lowest represented object’s speed. Because the estimations did not reveal either auditory or visual

dominance, the small looming/receding asymmetry observed in audiovision was the average of the auditory and visual estimates. In fact, if the unimodal weights are equal to .5 (i.e., no difference in weight between unimodal estimates, as in the present experiments), the predicted audiovisual estimate becomes $\hat{E}_{av} = (\hat{E}_a + \hat{E}_v)/2$. This explanation is supported by the comparison between the empirical results and the results predicted by the maximum-likelihood integrator model suggested by Ernst and Banks (2002). In addition, this result further supports the existence of multiple multisensory interactions (van Wassenhove et al., 2008). Since visual information is able to modulate (i.e., reduce) the looming/receding asymmetry obtained in the auditory domain, it is possible that, at some point, auditory and visual signals are integrated. Recent studies have pointed out that the human middle temporal complex (hMT+; a complex of extrastriate areas mainly involved in visual motion perception) is also activated by nonvisual (e.g., auditory and tactile) motion stimuli (Beauchamp, 2005). This point of view is supported by a number of neuroimaging (Beauchamp, Yaser, Kishan, & Ro, 2007; Hagen et al., 2002; Matteau, Kupers, Ricciardi, Pietrini, & Ptito, 2010; Ptito, Schneider, Paulson, & Kupers, 2008; Sani et al., 2010) and psychophysical (Bicchi, Scilingo, Ricciardi, & Pietrini, 2008; Konkle, Wang, Hayward, & Moore, 2009) studies. For example, Sani et al. showed, using fMRI, that in normal-sighted and congenitally blind patients the hMT+ was activated with the passive perception of visual and/or tactile moving stimuli. These results support the notion that the hMT+ has a supramodal organization that can represent the stage at which motion signals from different sensory modalities are integrated (Beauchamp, 2005). In our case, it is possible that in the audiovisual condition, both auditory and visual motion signals combine in order to render more precise and accurate the subjective duration of a stimulation carrying both signals. Indeed, the weak looming/receding asymmetry observed in audiovision is likely to reflect an average (i.e., an integration) between auditory and visual information. Since here we used stimuli that induced apparent motion in depth, it would be worthwhile to investigate, using fMRI and transcranial magnetic stimulation (TMS), whether such integration occurs at the level of the hMT+.

Is looming special?

The visual and audiovisual results of the present study are not interpretable in the light of the evolutionary explanation advanced by Neuhoﬀ (1998, 2001). Nonetheless, we may ask whether a looming overestimation is advantageous across senses. In our opinion, the response is “no.” The reason is that looming and receding proximal stimulations can result not only when an object moves toward or away from the perceiver, but also when the perceiver moves

toward or away from a still sound source. The latter situation occurs mainly with the perceiver moving while front-facing the object (and does not occur with the perceiver walking backward toward the object). Hearing is often considered an alerting sense that can inform us about events that are not visible, either because they are occluded or because they are out of the visual field (Neuhoﬀ, 2004). Under these conditions, an overestimation of a looming sound (in comparison to a receding one) may increase the margins of safety in the case of a motor reaction to the approaching sound source, as suggested by Neuhoﬀ (1998). However, when it comes to vision, such an overestimation would become disadvantageous. Vision needs a more veridical representation of the outside world, because looming objects can result either from the motion of an object toward or away from us or from our motion toward or away from a still object. Distortions in the perception of the durations of looming and receding visual patterns may lead to wrong estimates of the time of contact between the moving and the still object, and this may be highly disadvantageous when the moving object is a human being trying to catch a still prey. After the necessary changes, the results of the present experiment reverse the interpretation suggested by Neuhoﬀ (1998, 2001). As concerns duration, it is the receding sound that is negligible (and not the looming sound that is salient). By the same token, the results of the present study do not completely support the looming (as compared to receding) saliency claimed by other studies. First, here, the looming stimulus was not always overestimated in comparison to the control stationary stimulus (see, e.g., Exp. 1; note that a control condition, as such, is not always included in experiments that compare looming and receding stimuli—e.g., Franconeri & Simons, 2003; Neuhoﬀ, 1998; Schiff et al., 1962). Second, the looming stimulus was not much overestimated (in comparison to the receding one) in all experimental conditions (e.g., in vision, or in audition when the duration of the stimulus was long). Third, when the looming–receding asymmetry occurred, it originated mainly from an underestimation of receding stimuli in comparison to the looming ones. Therefore, the results of the present study suggest an alternative way to interpret the looming–receding asymmetry: It is not looming that is special, but receding that is negligible.

Author note The authors thank Daniela Di Menza, Chiara Gaino, and Silvia Bettin for helping with the data collection. We are also grateful for comments on an earlier version of the manuscript by Angus Gellatly, Steve Franconeri, and an anonymous reviewer.

References

- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14, 257–262.

- Beauchamp, M. S. (2005). See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology*, 15, 145–153. doi:10.1016/j.conb.2005.03.011
- Beauchamp, M. S., Yaser, N. E., Kishan, N., & Ro, T. (2007). Human MST but not MT responds to tactile stimulation. *Journal of Neuroscience*, 27, 8261–8267.
- Bicchi, A., Scilingo, E. P., Ricciardi, E., & Pietrini, P. (2008). Tactile flow explains haptic counterparts of common visual illusions. *Brain Research Bulletin*, 75, 737–741.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436. doi:10.1163/156856897X00357
- Calvert, G., Spence, C., & Stein, B. E. (2004). *The handbook of multisensory processes*. Cambridge, MA: MIT Press.
- Canévet, G. (1986). Estimation de sonie pour de sons purs à variation monotone de niveau. *Acustica*, 62, 256–264.
- Canévet, G., & Scharf, B. (1990). The loudness of sounds that increase and decrease continuously in level. *Journal of the Acoustical Society of America*, 88, 2136–2142.
- Canévet, G., Teghtsoonian, R., & Teghtsoonian, M. (2003). A comparison of loudness change in signals that continuously rise and fall in amplitude. *Acta Acustica United With Acustica*, 89, 339–345.
- Cappe, C., Thut, G., Romei, V., & Murray, M. M. (2009). Selective integration of auditory–visual looming cues by humans. *Neuropsychologia*, 47, 1045–1052.
- DiGiovanni, J. J., & Schlauch, R. S. (2007). Mechanisms responsible for differences in perceived duration for rising-intensity and falling-intensity sounds. *Ecological Psychology*, 19, 239–264.
- Eagleman, D. M., & Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philosophical Transactions of the Royal Society B*, 364, 1841–1851.
- Eisler, H. (1975). Subjective duration and psychophysics. *Psychological Review*, 82, 429–450.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415, 429–433. doi:10.1038/415429a
- Fechner, G. T. (1889). *Elemente der Psychophysik* (2nd ed.). Leipzig: Breitkopf & Härtel.
- Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, 65, 999–1010. doi:10.3758/BF03194829
- Ghazanfar, A. A., Neuhoﬀ, J. G., & Logothetis, N. K. (2002). Auditory looming perception in rhesus monkeys. *Proceedings of the National Academy of Sciences*, 99, 15755–15757.
- Grassi, M. (2003). Differenza nella durata percettiva di suoni crescenti o calanti in intensità: Permanenza o decurtamento? *Giornale Italiano di Psicologia*, 30, 659–663.
- Grassi, M. (2010). Sex difference in subjective duration of looming and receding sounds. *Perception*, 39, 1424–1426.
- Grassi, M., & Darwin, C. J. (2006). The subjective duration of ramped and damped sounds. *Perception & Psychophysics*, 68, 1382–1392. doi:10.3758/BF03193737
- Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short time-intervals. *Psychological Research*, 61, 261–268.
- Hagen, M. C., Franzen, O., McGlone, F., Essick, G., Dancer, C., & Pardo, J. V. (2002). Tactile motion activates the human middle temporal/V5 (MT/V5) complex. *European Journal of Neuroscience*, 16, 957–964.
- Hall, D. A., & Moore, D. R. (2003). Auditory neuroscience: The salience of looming sounds. *Current Biology*, 13, 91–93.
- Konkle, T., Wang, Q., Hayward, V., & Moore, C. I. (2009). Motion after-effects transfer between touch and vision. *Current Biology*, 19, 1–6.
- Lu, T., Liang, L., & Wang, X. (2001). Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. *Journal of Neurophysiology*, 85, 2364–2380.
- Maier, J. X., Neuhoﬀ, J. G., Logothetis, N. K., & Ghazanfar, A. A. (2004). Multisensory integration of looming signals by rhesus monkeys. *Neuron*, 43, 177–181.
- Matteau, I., Kupers, R., Ricciardi, E., Pietrini, P., & Ptito, M. (2010). Beyond visual, aural and haptic movement perception: hMT + is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. *Brain Research Bulletin*, 82, 264–270.
- Myers, A. K., Cotton, B., & Hilp, H. A. (1981). Matching the rate of concurrent tone bursts and light flashes as a function of flash surround luminance. *Perception & Psychophysics*, 30, 33–38.
- Neuhoﬀ, J. G. (1998). Perceptual bias for rising tones. *Nature*, 395, 123–124.
- Neuhoﬀ, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, 13, 87–110.
- Neuhoﬀ, J. G. (2004). Ecological psychoacoustics: Introduction and history. In J. G. Neuhoﬀ (Ed.), *Ecological psychoacoustics* (pp. 23–36). San Diego, CA: Elsevier Academic Press.
- Neuhoﬀ, J. G., & McBeath, M. K. (1997). Overcoming naive mental models in explaining the Doppler Shift: An illusion creates confusion. *American Journal of Physics*, 65, 618–621.
- Neuhoﬀ, J. G., Planisek, R., & Seifritz, E. (2009). Adaptive sex differences in auditory motion perception: Looming sounds are special. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 225–234.
- Olsen, K. N., & Stevens, C. J. (2010). Perceptual overestimation of rising intensity: Is stimulus continuity necessary? *Perception*, 39, 695–704.
- Olsen, K. N., Stevens, C. J., & Tardieu, J. (2010). Loudness change in response to dynamic acoustic intensity. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 1631–1644.
- Over, R. (1966). A experimentally induced conflict between vision and proprioception. *British Journal of Psychology*, 57, 335–341.
- Pastore, R. E., & Flint, J. (2011). Magnitude judgments of loudness change for discrete, dynamic, and hybrid stimuli. *Attention, Perception, & Psychophysics*, 73, 886–907. doi:10.3758/s13414-010-0056-8
- Patterson, R. D. (1994a). The sound of a sinusoid: Spectral models. *Journal of the Acoustical Society of America*, 96, 1409–1418.
- Patterson, R. D. (1994b). The sound of a sinusoid: Time-interval models. *Journal of the Acoustical Society of America*, 96, 1419–1428.
- Patterson, R. D., & Irino, T. (1998). Modelling temporal asymmetry in the auditory system. *Journal of the Acoustical Society of America*, 104, 2967–2979.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442. doi:10.1163/156856897X00366
- Penney, T. B. (2003). Modality differences in interval timing: Attention, clock speed, and memory. In W. H. Meck (Ed.), *Functional and neural mechanisms of interval timing* (pp. 209–233). Hoboken, NJ: CRC Press.
- Phillips, D. P., Hall, S. E., & Boehnke, S. E. (2002). Central auditory onset responses, and temporal asymmetries in auditory perception. *Hearing Research*, 167, 192–205.
- Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, 132, 416–442. doi:10.1037/0033-2909.132.3.416
- Ptito, M., Schneider, F. C., Paulson, O. B., & Kupers, R. (2008). Alterations of the visual pathways in congenital blindness. *Experimental Brain Research*, 187, 41–49.
- Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, 143, 594–596.
- Sani, L., Ricciardi, E., Gentili, C., Vanello, N., Haxby, J. V., & Pietrini, P. (2010). Effects of visual experience on the human MT +

- functional connectivity networks: An fMRI study of motion perception in sighted and congenitally blind individuals. *Frontiers in System Neuroscience*, 20, 1–9.
- Schiff, W., Caviness, J. A., & Gibson, J. J. (1962). Persistent fear responses in rhesus monkeys to the optical stimulus of “looming. *Science*, 136, 982–983.
- Schlauch, R. S. (1992). A cognitive influence on the loudness of tones that change continuously in level. *Journal of the Acoustical Society of America*, 92, 758–765.
- Schlauch, R. S., Ries, D. T., & DiGiovanni, J. J. (2001). Duration discrimination and subjective duration for ramped and damped sounds. *Journal of the Acoustical Society of America*, 109, 2880–2887.
- Stecker, G. C., & Hafter, E. R. (2000). An effect of temporal asymmetry in loudness. *Journal of the Acoustical Society of America*, 107, 3358–3368.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64, 153–181. doi:10.1037/h0046162
- Susini, P., McAdams, S., & Smith, B. K. (2002). Global and continuous loudness estimation of time-varying levels. *Acta Acustica United With Acustica*, 88, 536–548.
- Susini, P., McAdams, S., & Smith, B. K. (2007). Loudness asymmetries for tones with increasing and decreasing levels using continuous and global ratings. *Acta Acustica United With Acustica*, 93, 623–631.
- Teghtsoonian, R., Teghtsoonian, M., & Canévet, G. (2000). The perception of waning signals: Decruitment in loudness and perceived size. *Perception & Psychophysics*, 62, 637–646. doi:10.3758/BF03212115
- Teghtsoonian, R., Teghtsoonian, M., & Canévet, G. (2005). Sweep-induced acceleration in loudness change and the “bias for rising intensities. *Perception & Psychophysics*, 67, 699–712. doi:10.3758/BF03193526
- Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics*, 66, 1171–1189. doi:10.3758/BF03196844
- van Wassenhove, V. (2009). Minding time—An amodel representational space for time perception. *Philosophical Transactions of the Royal Society B*, 364, 1815–1830.
- van Wassenhove, V., Buonomano, D. V., Shimojo, S., & Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS One*, 3, e1437. doi:10.1371/journal.pone.0001437
- von Mühlenen, A., & Lleras, A. (2007). No-onset looming motion guides spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 1297–1310. doi:10.1037/0096-1523.33.6.1297
- Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88, 638–667. doi:10.1037/0033-2909.88.3.638
- Westheimer, G. (1999). Discrimination of short time intervals by the human observer. *Experimental Brain Research*, 129, 121–126.
- Wittmann, M., van Wassenhove, V., Craig, A. D., & Paulus, M. P. (2010). The neural substrates of subjective time dilation. *Frontiers in Human Neuroscience*, 4, 1–9. doi:10.3389/neuro.09.002.2010
- Wright, B. A., Buonomano, D. V., Mahncke, H. W., & Merzenich, M. M. (1997). Learning and generalization of auditory temporal-interval discrimination in humans. *Journal of Neuroscience*, 17, 3956–3963.
- Zera, J., & Green, D. M. (1993). Detecting temporal onset and offset asynchrony in multicomponent complexes. *Journal of the Acoustical Society of America*, 93, 1038–1052.