Exp Brain Res (2005) 166: 445–454
DOI 10.1007/s00221-005-2384-9

RESEARCH ARTICLE

Roberto Arrighi · David Alais · David Burr

Perceived timing of first- and second-order changes in vision and hearing

Received: 16 July 2004 / Accepted: 21 October 2004 / Published online: 29 September 2005

Abstract Simultaneous changes in visual stimulus attributes (such as motion or color) are often perceived to occur at different times, a fact usually attributed to differences in neural processing times of those attributes. However, other studies suggest that perceptual misalignments are not due to stimulus attributes, but to the type of change, first- or second-order. To test whether this idea generalizes across modalities, we studied perceptual synchrony of acoustic and of audiovisual cross-modal stimuli, which varied in a first- or second-order fashion. First-order changes were abrupt changes in tone intensity or frequency (auditory), or spatial position (visual), while second-order changes were an inversion of the direction of change, such as a turning point when a rising tone starts falling or a translating visual blob reverses. For both pure acoustic and cross-modal stimuli, first-order changes were systematically perceived before second-order changes. However, when both changes were first-order, or both were second-order, little or no difference in perceptual delay was found between them, regardless of attribute or modality. This shows that the type of attribute change, as well as latency differences, is a strong determinant of subjective temporal alignments. We also performed an analysis of reaction times (RTs) to the first- and second-order attribute changes used in these temporal alignment experiments. RT differences between these stimuli did not correspond with our temporal alignment data, suggesting that subjective alignments cannot be accounted for by a simple latency-based explanation.

Keywords Auditory temporal alignments · Audio-visual temporal alignments · Reaction times · Temporal markers

R. Arrighi (✉) · D. Burr
Dipartimento di Psicologia, Università degli Studi di Firenze,
Via S. Niccolò 89, 50125 Firenze, Italy
E-mail: arrighi@in.cnr.it
Fax: +39-50-315-3220

D. Alais
Auditory Neuroscience Laboratory, Department of Physiology
and Institute for Biomedical Research, School of Medical Science,
University of Sydney, Sydney, NSW 2006, Australia

D. Burr
Istituto di Neuroscienze del CNR, Via Moruzzi 1, Pisa, Italy

Introduction

The representation of the external world depends upon the capacity of the brain to localize events correctly in time. However, perception is not an instantaneous process. We perceive events delayed relative to their actual occurrence and many factors could create a discrepancy between actual and perceived time. For example, it is generally accepted that different cortical areas within different neural pathways contribute to the analysis of stimulus attributes such as luminance, color and motion. Thus, differences in latencies for these attributes could create perceptual asynchronies. In addition, many findings have also shown that attention (Stelmach and Herdman 1991; Jaskowski 1993; Spence et al. 2001), adaptation (Fujisaki et al. 2003), stimulus intensity (Roufs 1963) as well as spatial displacement (Zampini et al. 2003) can play a role in temporal judgments. Thus explanations offered to account for perceptual time illusions have usually been couched in terms of neural latencies and/or cortical processing times (Moutoussis and Zeki 1997; Arnold et al. 2001; Viviani and Aymoz 2001; Aymoz and Viviani 2004). One intriguing simultaneity illusion entailing the attributes of color and motion was discovered by Moutoussis and Zeki (1997). They displayed a plaid pattern oscillating smoothly upward and downward at 2 Hz, whose color changed from red to green in synchrony with the direction changes (e.g. red-upward/green-downward). Under these conditions, it was hard to determine which color and direction were paired. However, delaying color transitions relative to motion direction reversals by about 100 ms produced a clear perceptual correspondence of color and direction. The authors interpreted this as reflecting different processing times in the cortical modules specialized for color and motion, citing neurophysiological evidence indicating faster processing times for color than for motion perception.

Nishida and Johnston (2002) have recently proposed a different explanation of the phenomenon. In an elegant series of experiments, they showed that the perceptual asynchrony can be inverted by changing the temporal transformation of the stimuli. If the plaid moves smoothly from red to green and abruptly changes spatial position, color direction reversals had to lead position changes by about 100 ms to produce perceptual synchrony. If the plaid underwent changes in color and spatial position that were either both abrupt or both smooth, no temporal misalignments were perceived between color and spatial change. They suggested that the temporal characteristics of the stimulus transformation determine perceptual alignments because abrupt changes appeared to be processed quickly, and gradual changes slowly. This can be thought of as a difference between first-order and second-order stimulus changes. First-order (abrupt) changes can be detected by monitoring just two locations over time because the decision is binary: is the feature or attribute in position one or two? However, for second-order (smooth) changes, sampling of three points is required: two to determine that the attribute has changed position over time, and a third to determine if it continues to change in that direction or not.

As further evidence that temporal characteristics rather than attribute latencies determine temporal alignment, Nishida and Johnston (2002) replicated their results using luminance instead of color. The authors concluded that because the exact moment of directional change is difficult to determine in a smoothly oscillating motion, it provides less salient information about the moment of phase change than does an abrupt transition. On their analysis, the oscillating motion is a less reliable "temporal marker" than is the abrupt change. Thus, temporal alignments of first- and second-order changes would provide perceptual asynchronies because they require comparison of different types of temporal changes. These illusions could only be annulled if direction reversals (second-order changes) led abrupt transitions (first-order changes), regardless of the stimulus attributes involved. They also predicted that no asynchronies would be perceived if both attributes changed abruptly, or if both changed smoothly, because in these conditions the same kind of temporal changes would be compared.

The aim of the present investigation is to test whether Nishida and Johnston's (2002) hypothesis that the temporal characteristics of stimulus transformation determine perceptual alignments generalizes across modalities. To this end, using a similar technique to theirs, we measured perceptual alignment of auditory and of audiovisual cross-modal stimuli that underwent either the same or different temporal transformations. In the auditory experiment, we used first- and second-order stimuli (i.e. abrupt transitions and direction reversals, respectively) in intensity and in frequency, whilst in the cross-modal version of the experiment we paired stimuli that moved in auditory frequency and in visual space. If Nishida and Johnston's (2002) idea holds across modalities, we expect to find significant perceptual asynchronies in pairing a first- and a second-order change, and little or no asynchrony between stimulus changes of the same order (either both first- or both second-order). Finally, reaction times (RTs) were measured for first- and second-order stimuli to obtain estimates of neural latencies for all of our stimuli. These were used to test whether latency differences provide an explanation of the temporal alignment data. If so, then the latency difference between two attributes would predict the temporal offset required to align those attributes in a temporal alignment task.

Materials and methods

Subjects

Two of the authors (R.A. and D.A.) and three naive female subjects (mean age 28 years), all with normal hearing and normal or corrected vision acuity, served as subjects and gave informed consent to participate to the study according to the guidelines of the Universities of Sydney and Florence. The tasks were performed in a dimly lit and sound attenuated room.

Apparatus and stimuli

Visual stimuli were generated by a Cambridge VSG 2/3 framestore and presented with a refresh rate of 67 Hz on a Pioneer color plasma monitor subtending 90×57° from the viewing distance of 54 cm with a spatial resolution of 13 pixels per degree. The visual stimulus was a Gaussian luminance blob with a space constant of 2° at 100% contrast that oscillated vertically between two spatial positions 30° apart (±15° relative to the monitor's half-height). First-order oscillation was defined as square wave, and second-order motion by triangular wave (see Fig. 1).

Auditory stimuli were constant pure tones (for first-order changes) or tones sweeping frequency or intensity (for second-order changes), all of them were created using Matlab software and digitized at a rate of 65 kHz. There were two kinds of auditory stimuli: (1) pure tones at 75 dB, which oscillated in frequency between 400 and 600 Hz and (2) pure tones at 500 Hz, which oscillated in intensity between 60 and 90 dB. Auditory stimuli intensity was measured at the sound source and the sound levels expressed above in dB actually refer to absolute levels. It is worth noting that within the chosen frequency range, the same level of sound pressure

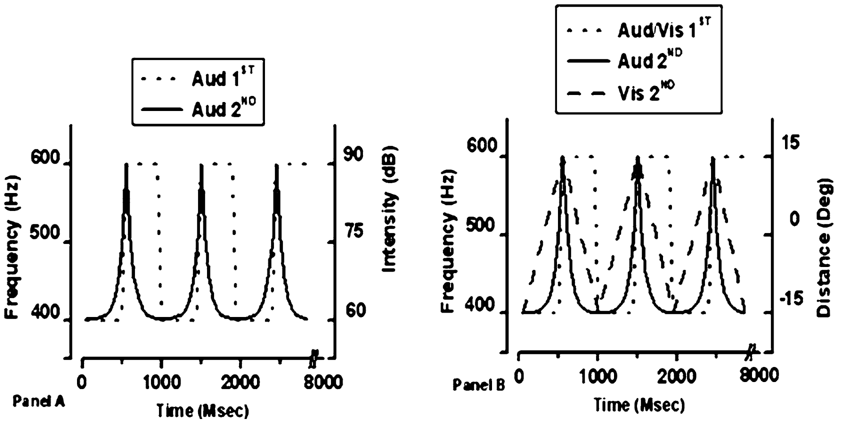


Fig. 1 Illustration of temporal modulations used in these experiments. **a** In the auditory experiment, first-order oscillations were defined by a square-wave temporal profile (*dotted lines*) while second-order auditory oscillations followed a logarithmic ramp (*a triangular wave on log scale*), as indicated by continuous line. **b** In the cross-modal experiment, visual first-order oscillations followed a square-wave temporal profile (*dotted line*) and second-order oscillations followed a triangular profile (*dashed line*). These oscillations were combined with auditory stimuli using the temporal profiles described in (**a**)

produce equal levels of perceptual loudness thus, any intensity cues were provided by oscillations in frequency space (Hartmann 2000). For both types of auditory stimuli, square-wave oscillations were used to produce first-order changes while second-order changes were created using a logarithmic ramp (i.e. a triangular wave but on a log scale for frequency and intensity, see Fig. 1).

In the auditory temporal alignment experiment, the stimuli were presented using professional circumaural headphones (Beyer Dynamic DT990); however, in the cross-modal version of this experiment, they were presented through two high-quality loudspeakers (Yamaha MSP5) flanking the video monitor (center-to-center separation of 100 cm) and lying in the same plane. The loudspeakers were used in the cross-modal condition to displace both the auditory and visual stimuli at the same depth in space, as we did want our subjects to focus their spatial attention on the same location in depth during the experiment (Kopinska and Harris 2004). We then displaced the trajectory of the vertically oscillating blob 10° to the left of the screen center while the sound location was shifted 10° to the right, as it has been shown that presenting auditory and visual stimuli from different spatial locations increase sensitivity for audiovisual temporal judgments (Zampini et al. 2003). Moreover, this arrangement also ensured that our auditory stimulus was not captured by the visual stimulus (Bertelson and Aschersleben 2003; Alais and Burr 2004).

A digital data acquisition system with a temporal resolution of 3000 Hz was used to calibrate audio-visual presentation and to ensure accurate temporal alignment of the stimuli. The auditory channel of the acquisition system received signals directly from the PC sound card whilst the visual channel recorded the plasma screen activations through a high sensitive photocell plugged on top of it. Calibration of visual and auditory signals on a storage oscilloscope revealed a delay of two frames (29.85 ms at 67 Hz) between the visual trigger and the plasma screen activation that was compensated for.

Procedure

Temporal alignments

In the pure auditory experiment, on each block of trials, a sound intensity oscillating in either abruptly (first-order) or in a triangular fashion (second-order), was compared with a sound frequency oscillating in either first- or second-order fashion (four stimulus combinations in total). Intensity modulations were ramped on and off within a raised cosine function of randomized length within the range of 0.5–2 s. In the cross-modal version, the sound oscillating in intensity was replaced by a visual blob oscillating in space as described above. Each stimulus pairing was measured in a separate session. The stimuli were always played for 8 s duration. The oscillation rate (per full cycle) in all the conditions was 1 Hz, resulting in a speed of 60°/s for the second-order visual stimuli. Subjects were required to indicate by key press on a PC keyboard, whether the two stimuli oscillated in-phase or not. In each session, the relative timing of the two stimuli was randomly varied from −250 ms to +250 ms in 50 ms intervals and 25 responses for each relative phase were recorded.

Reaction times

RTs were collected using a high-precision CB3 experiment response box (Cambridge Research Systems) supporting a high-resolution counter with microsecond accuracy. In separate blocks of trials, we measured RTs for one of the stimulus attributes (intensity, frequency or spatial position) undergoing either a first- or second-order change as described above. All changes, whether first- or second-order, were preceded by a variable period. For first-order changes, this period in which the stimulus attribute was constant, was randomly chosen within the range of 0.5–1.5 s. For second-order changes, a rising phase was kept for a random period of 0.25, 0.5 or 1 s before any reversals occurred. In any case, the subject task was to press a button on the response box as soon as she perceived a stimulus attribute change. Fifty trials were collected for each of the attribute/oscillation type combination.

Results

Experiment 1: temporal alignment of auditory stimuli

Data for temporal alignment of auditory stimuli for three of our subjects are shown in Fig. 2. Each point in Fig. 2 represents the percentage of "in-phase" responses at each stimulus onset asynchrony. The dashed lines are the best-fitting Gaussian curves to the data with R^2 values between 0.83 and 0.996. The peak of these curves indicates the temporal asynchrony that produced subjective alignment of the oscillating auditory stimuli. The conditions in which two first-order or two second-order oscillations were paired (the red and the magenta curves, respectively), have peaks near zero, implying that these oscillations were perceived as being aligned when they were physically aligned. The green curve shows the condition in which a first-order oscillation in frequency and a second-order oscillation in intensity were paired. In this condition, the peak occurred when the intensity oscillation occurred about 180 ms earlier than the frequency oscillation. Finally, the blue curve, indicating the condition in which a second-order oscillation in frequency was paired with a first-order oscillation in intensity, peaked when frequency oscillations led intensity oscillations by about 150 ms. Note that the asyn-

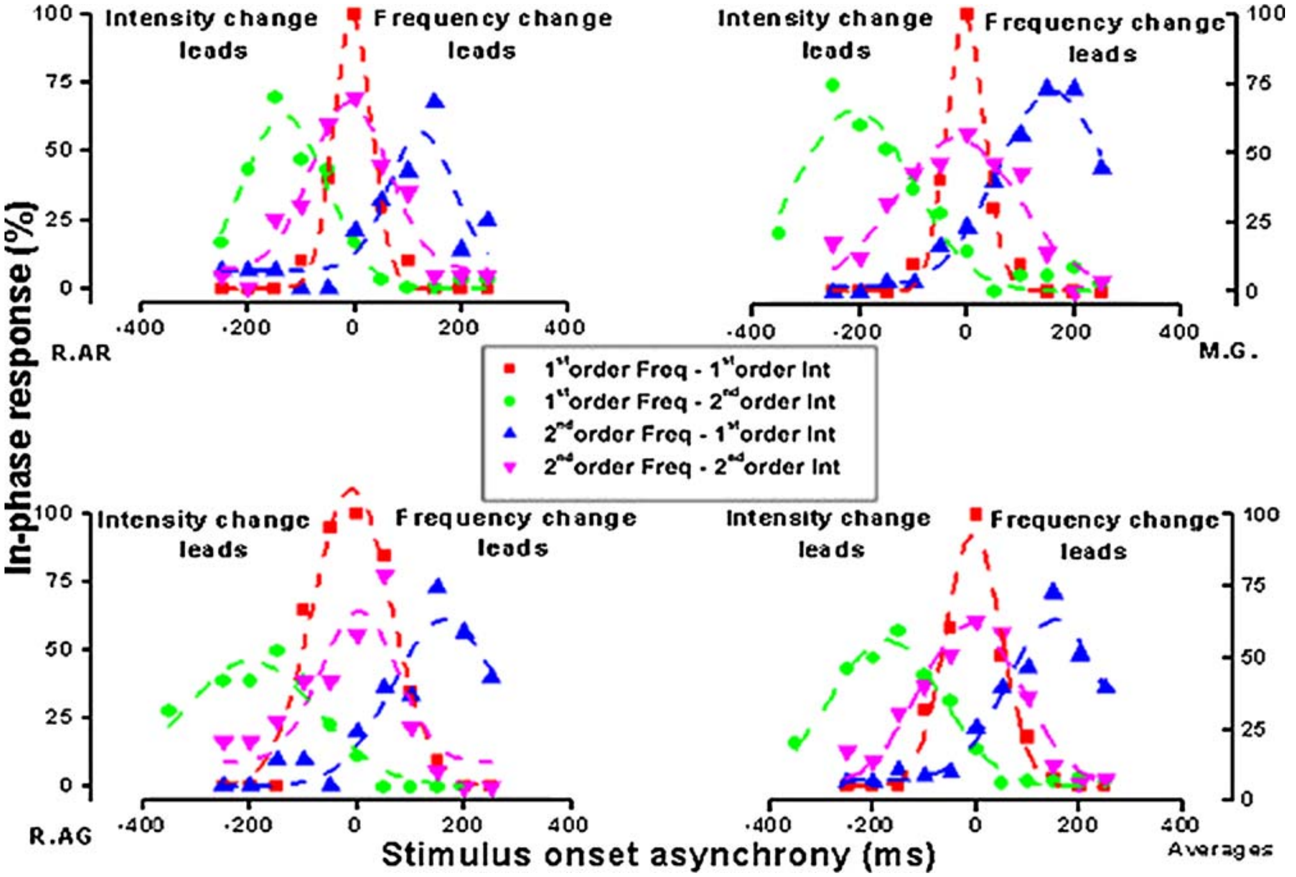


Fig. 2 Data from experiment 1 showing the onset asynchrony required to align auditory stimuli undergoing first- or second-order oscillations in either frequency or intensity. The ordinate shows the percentage of in-phase responses against the onset-time difference between the oscillating stimuli (25 trials/data point). A value of zero on the abscissa indicates the point at which the stimuli are physically aligned in time. The *dashed lines* are the best-fitting Gaussian curves to the data (R^2 values between 0.83 and 0.996) and the peak indicates the temporal asynchrony, which produced subjective temporal alignment. Four conditions are shown: sound intensity, undergoing either a first- or second-order oscillation, paired with sound frequency, undergoing either a first- or second-order oscillation. Conditions in which both stimuli underwent the same type of oscillation (either both first-order or both second-order) have peaks near zero. That is, they were perceived as being aligned when they were actually physically aligned. On the contrary, conditions in which different types of oscillations were paired required the first-order stimulus to be delayed relative to the second-order stimulus in order to obtain subjective alignment

chronies reported here are very large, in the order of 100 ms, and larger than the width of the Gaussian fit. Thus for these first-order, second-order comparisons, when the changes were actually in synchrony they were almost seen always as asynchronous (green and blue curves).

From the alignment data in Fig. 2, we can infer the differential perceptual latencies associated with each of the attribute/oscillation pairings. Figure 3 shows these implied perceptual latencies, for each condition and for each subject (consider only the filled symbols and dotted lines). Similar types of oscillation, whether both first-order or both second-order, have similar perceptual latencies. In contrast, when first- and second-order oscillations were paired, perception of the second-order oscillations clearly has a longer perceptual latency than that for first-order oscillations by an average of 170 ms. This pattern of results indicates that the processing times associated with frequency and intensity perception do not differ, but rather that it is the complexity of the oscillation pattern that determines the perceptual latency. Second-order oscillations appear to enter consciousness 170 ms after first-order oscillations.

Experiment 2: reaction times for first- and second-order auditory stimuli

Experiment 2 measured RTs for the first- and second-order attribute changes used in experiment 1 to obtain estimates of the perceptual latencies associated with each of them. These will be used to test whether latency differences provide an explanation of the temporal alignment data. On the common assumption that the time required for producing a motor response is approximately the same for all the attribute types of change, the RT difference between a pair of them provides a useful estimate of the difference in overall neural processing times. “Overall” in this case includes transduction, transmission and integration times. If differential latencies are the basis for the temporal alignment data obtained in experiment 1, then the RT difference between two attributes should agree closely with the temporal offset required to align those attributes in a temporal alignment task. Experiment 2 tests this prediction.

RTs for perceiving the four attribute types of change used in experiment 1 are shown in Fig. 4. RTs for first-order changes were shorter than those for second-order changes, whether the attribute was frequency or intensity. The RT difference between first-order changes in frequency and intensity averaged about 25 ms. This difference is small but was statistically significant for all observers on a two-tailed unpaired *t*-test (RAR: $t98 = 2.89$*; MG: $t98 = 2.56$*; RAG: $t98 = 3.01$*, $P < 0.05$ for all subjects). The RT difference between second-order changes in frequency and intensity was larger, averaging about 150 ms. For second-order frequency changes, RTs were roughly 80 ms greater than those for first-order changes, and roughly 55 ms longer than those for first-order intensity changes. However, by far the

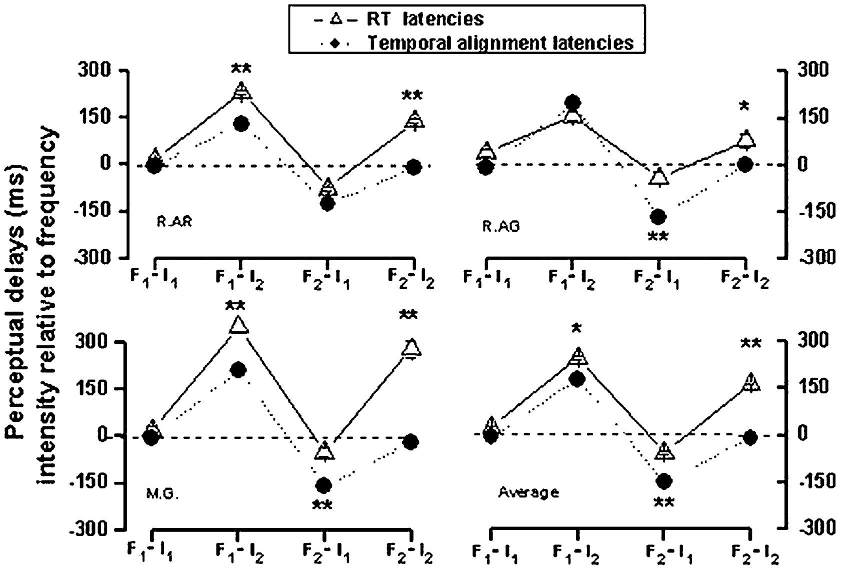


Fig. 3 *Filled symbols:* perceptual delay as derived from the temporal alignment data of experiment 1 for four pairings of attribute and oscillation type. Negative values indicate that intensity was perceived earlier, as was the case when the intensity oscillations were first-order. In the converse condition, when the intensity oscillations were second-order, the reverse was true. Pairing oscillations of the same order, whether first- or second-order, did not yield any perceptual delays. *Open symbols:* perceptual latencies as indicated by RT data derived from the RT latency differences found in experiment 2. The only condition in which the RT data and temporal alignment data do not show a statistically significant difference for all subjects is that in which both stimuli underwent first-order oscillations. In contrast, RT latencies are most wayward from perceptual latencies when both stimuli underwent second-order changes. In this condition, the near-zero differential latency indicated by the temporal alignment data implies approximately equal RTs, whereas the obtained RTs actually differ by nearly 150 ms

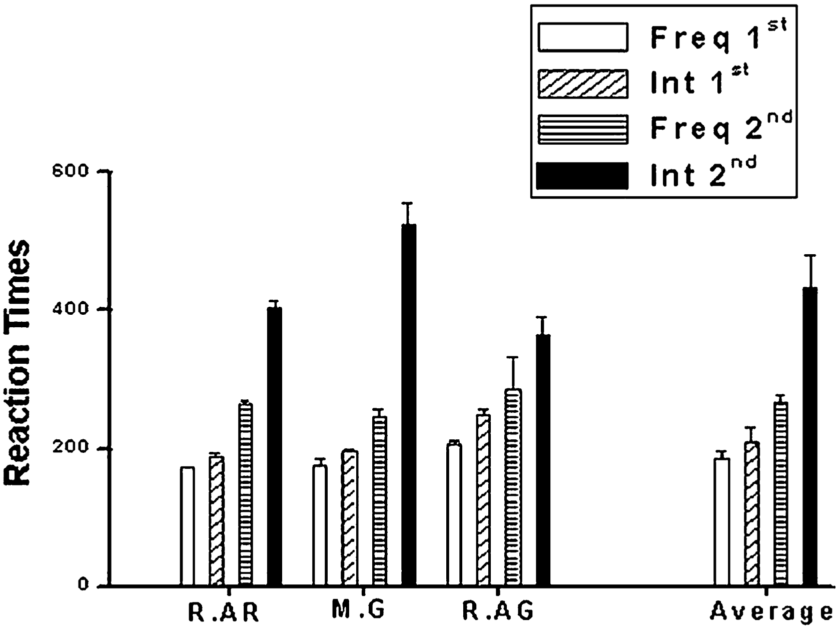


Fig. 4 Data from experiment 2 showing RTs for first- and second-order changes in frequency and in intensity for three subjects and their averaged data. The *bars* indicate the median of the RT distribution and the standard error of the RT distributions of each subject. The data show large and consistent variability in reaction times to the various stimuli, fastest for first-order frequency changes and slowest for second-order intensity changes

longest RTs were recorded for second-order intensity changes, which averaged approximately 430 ms. This pattern of results was consistent across all subjects.

Differential latencies were calculated from the RT data by taking the differences between each pairing of auditory attribute and oscillation type, and plotted together with the differential perceptual latencies obtained in experiment 1 (open symbols of Fig. 3). For all the subjects as well as for the averaged data, a two-tailed and unpaired *t*-test was performed for each condition to determine statistically significant differences between the RT and temporal alignment data. In Fig. 3, conditions in which the differences were significant at $P<0.05$ are marked with an asterisk, with a double asterisk indicating significance at $P<0.01$. In only one of the four conditions (first-order changes for both stimuli) do the perceptual latencies predicted by the RT data closely match the temporal alignment data from experiment 1 for all subjects. In all other conditions, the RT predictions are different from the alignment data. The RT difference between first-order frequency and second-order intensity oscillations (on average about 250 ms) overestimated the temporal alignment differences for this condition which averaged less than 200 ms. In the converse condition, with second-order frequency and first-order intensity oscillations, the RTs underestimated the temporal alignment difference by 50–70 ms. The largest difference between differential RT latencies and perceptual alignments was found for the condition in which both the intensity and the frequency underwent second-order changes. The alignment data imply very similar perceptual latencies for these stimuli (see Fig. 3), yet the RT data show a very large difference averaging almost 150 ms. Clearly, this result poses a major challenge to a simple latency account.

Experiment 3: cross-modal temporal alignments of auditory and visual stimuli

Data for three subjects for cross-modal temporal alignments of visual and auditory stimuli are shown in Fig. 5. The red curve of Fig. 5, indicating the condition in which first-order visual and auditory stimuli oscillations were paired, peaked when visual oscillations lead auditory oscillations by about 35 ms. This difference is statistically significant for all the subjects [RAR: $t(164)=6.83^{*}$; KA: $t(164)=5.05^{*}$; DA: $t(164)=6.11^{*}$; $\alpha=0.05$], showing that in this condition, visual perception turned out to be a little faster than auditory perception. However, the magenta curve, representing the condition in which two second-order oscillations were paired, peaked when auditory changes led visual ones by about 15 ms, and even if this is indeed a small difference, it was found to be statistically significant for two out of three subjects [RAR: $t(164)=-0.65$; KA: $t(164)=-3.71^{*}$; DA: $t(164)=-3.11^{*}$; $\alpha=0.05$], suggesting that in this condition auditory perception is slightly faster. Taken together, these results agree with the purely auditory data (Fig. 2) in that, pairing stimuli with like oscillations yielded near-zero perceptual asynchronies. The green curve indicates the condition in which visual first-order and auditory second-order oscillations were paired, peaking when auditory changes led visual changes by about 120 ms. Conversely, in the condition in which auditory first-order and visual second-order oscillations were paired (blue curve), the peak occurred when visual changes led auditory changes by about 60 ms.

As for experiment 1, we derived from temporal alignment data the perceptual delays for each attribute/oscillation type, shown in Fig. 6 (filled symbols). The pattern of results for cross-modal temporal alignments agrees closely with those for temporal alignments of

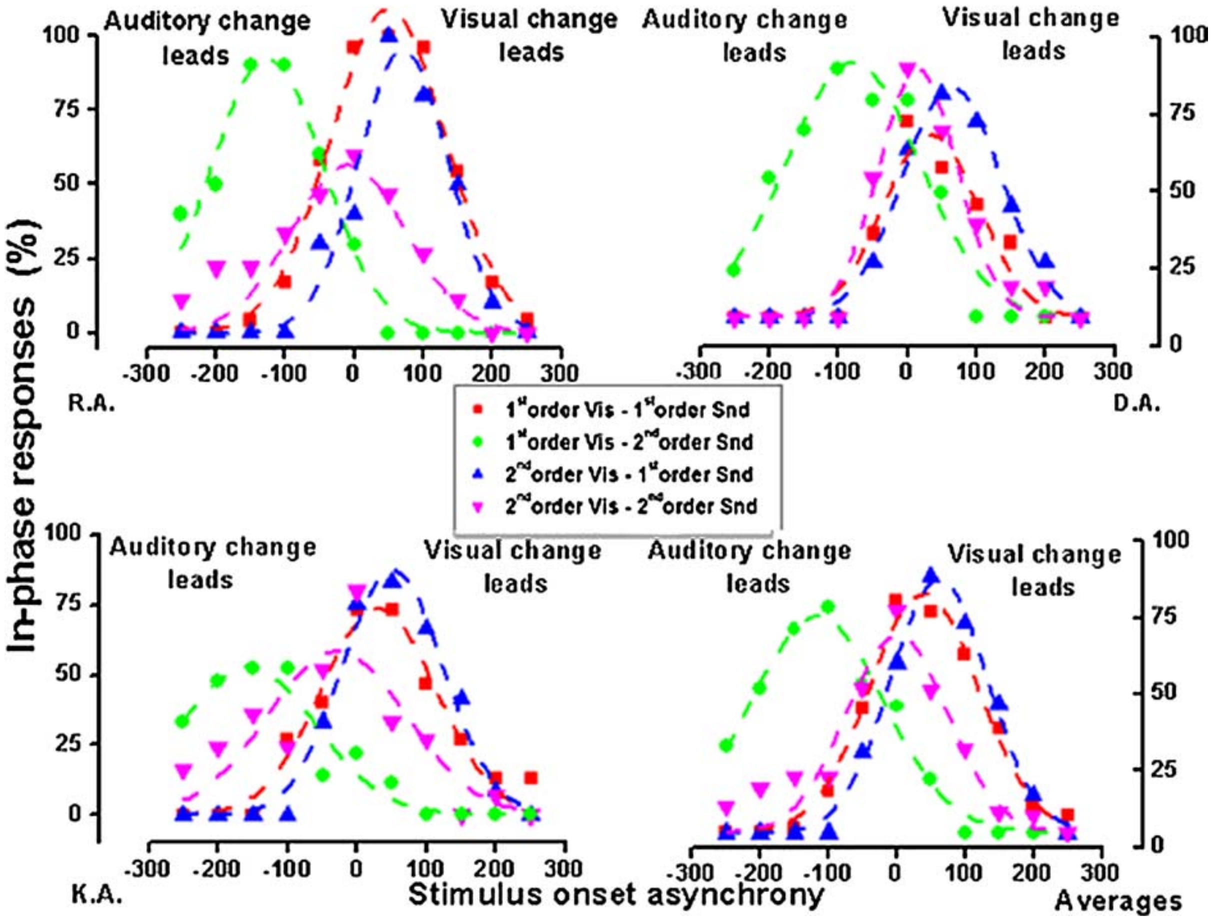


Fig. 5 Data from experiment 3 showing the onset asynchrony required to align auditory and visual stimuli undergoing first- or second-order oscillations. The *ordinate* shows the percentage of in-phase responses against the onset-time difference between auditory and visual oscillations. A value of zero on the *abscissa* indicates the point at which the stimuli are physically aligned in time. The *dashed lines* are the best-fitting Gaussian curves to the data (R^2 values between 0.76 and 0.99) and the peaks indicate the temporal offsets, which produced subjective alignment. The conditions in which both stimuli underwent like oscillations, whether both first- or both second-order, produced relatively small perceptual asynchronies as the near-zero peaks of the *red* and *magenta curves* indicate. This means that these pairs of stimuli were perceived as oscillating in-phase when they were approximately physically aligned. On the other hand, conditions in which different types of stimulus changes were paired required the first-order stimulus to be delayed relative to the second-order stimulus in order to obtain subjective alignment, regardless of the modality involved

purely visual stimuli (Nishida and Johnston 2002) and for purely auditory stimuli (experiment 1). Perceptual latencies of stimuli that underwent the same type of attribute change (both first- or second-order) were similar, while first-order oscillations paired with second-order oscillations created the perceptual latencies for second-order changes were found to be significantly longer (by around 100 ms on average), irrespective of modality. These results clearly suggest that perceptual latencies are determined more by the complexity of the temporal oscillation than by the different processing delays in perceiving the various visual and auditory attributes.

Experiment 4: reaction times for first- and second-order auditory and visual stimuli

RTs for cross-modal stimuli used in experiment 3 are shown in Fig. 7. RTs for first-order changes in auditory frequency were shorter than those for first-order changes in vision by 50 ms on average, in line with many previous findings (Galton 1899; Brebner and Welford 1980; Welford 1980). In contrast, RTs for second-order visual changes were about 100 ms shorter than those for second-order changes in auditory frequency. First-order auditory frequency RTs were slightly shorter than those for second-order visual changes, on average by about 15 ms. An ANOVA repeated-measures analysis confirmed that this small difference was statistically significant across subjects [$F(1,2) = 20.10$, $P < 0.05$] whilst a two-tailed and unpaired t-test reveals that actually this difference was significant for just one of our subjects [RAR: $t98 = 2.69$*; KA: $t98 = 0.55$; DA: $t98 = 1.47$; $\alpha = 0.05$]. Finally, RTs for first-order visual changes (on average, 35 ms longer than second-order changes in the same modality) were about 60 ms shorter than the longest RTs found, those for second-order auditory changes (about 280 ms).

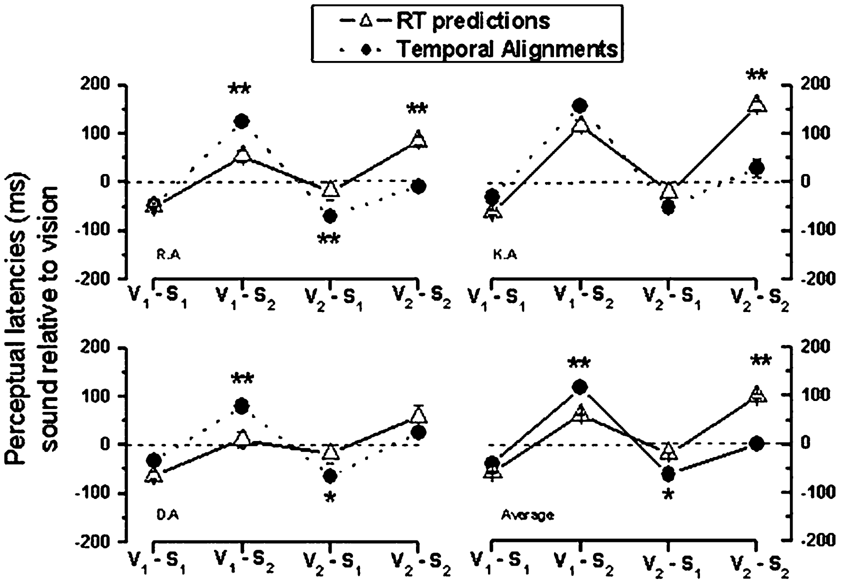
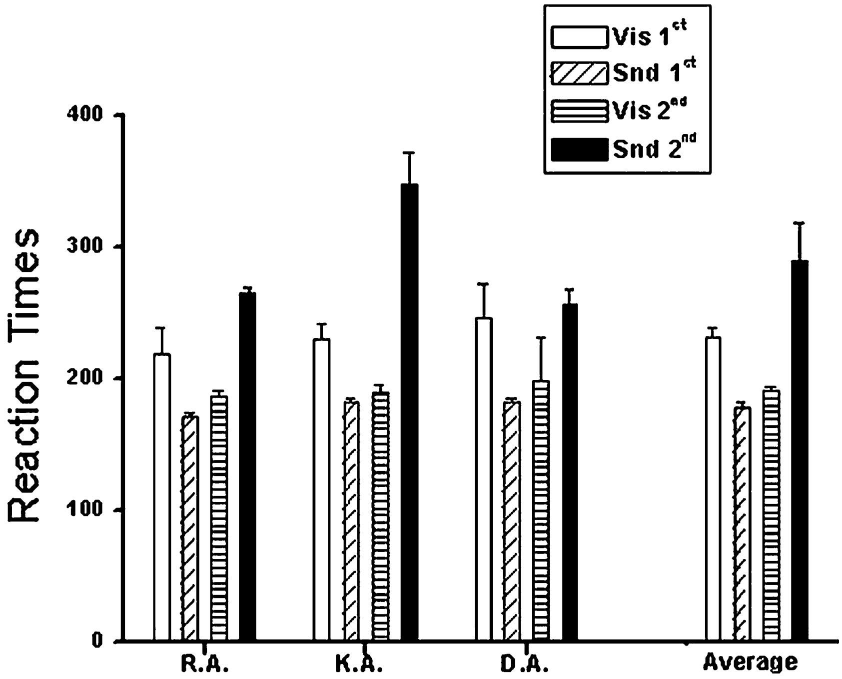


Fig. 6 *Filled symbols:* perceptual delays as implied from cross-modal temporal alignments (experiment 3). The *abscissa* shows the four pairings of attribute and oscillation type, while the ordinate shows the perceptual delay of sound relative to vision. Negative values indicate that sound was perceived faster than vision, as was the case when auditory oscillations in frequency were first-order. However, second-order auditory oscillations paired with first-order visual oscillations implied that sound perception was slower. Pairing oscillations of the same order, whether first- or second-order yielded relatively small perceptual delays. *Open symbols:* perceptual latencies as indicated by RT data derived from the RT latency differences found in experiment 4. The RT data and temporal alignment data do not agree except for the condition in which visual and auditory first-order oscillations were paired. In the condition where the oscillations were both second-order, RT predictions significantly overestimated auditory perceptual delays by about 80 ms. Moreover, RT predictions underestimated second-order auditory perceptual delays (when paired with first-order changes in vision), as well as second-order visual perceptual delays (when paired with first-order auditory oscillations)

Figure 6 plots the RT differences together with the perceptual latencies calculated for alignment (open symbols). As before, for each subject and for each condition, we tested for significant statistical differences between RT and temporal alignment data using a two-tailed and unpaired *t*-test. In line with the findings in the pure auditory condition, RT data does not predict the temporal alignment data except for the condition in which two first-order stimuli were paired. For example, temporal alignments of first-order visual oscillations

Fig. 7 Data from experiment 4 showing RTs for first- and second-order changes in auditory frequency and visual spatial position for three subjects and their average. The *bars* indicate the median and the standard error of the RT distributions. RTs for first-order visual oscillations are longer than those for first-order auditory oscillations. On the other hand, for second-order oscillations, the opposite result is obtained and auditory RTs are significantly longer than those for vision. However, RTs for visual second-order changes were found to be slightly longer than RTs for first-order auditory changes



with second-order auditory oscillations produced an auditory perceptual delay of about 120 ms, while RTs differences predicted 60 ms, about half. Better latency predictions were obtained in the opposite condition, with second-order visual oscillations and first-order auditory oscillations, as here the difference between predicted and actual perceptual latencies was on average just 40 ms. However, as in experiment 2, the widest gap between perceptual delays predicted by RTs and those derived from temporal alignment data, occurred in the conditions pairing two second-order oscillations, as the gap between predictions was around 80 ms. These results cast doubt on the idea that cross-modal temporal asynchronies result solely from differences between neural latencies of visual and auditory first- and second-order stimuli.

Global analysis of reaction times versus temporal alignments data

The data reported here show that both for pure auditory stimuli and cross-modal stimuli reaction-time relative delays do not account for temporal alignment data. To examine this further, we plot in Fig. 8 differences in RTs against differences in temporal alignment judgments for the eight comparisons of this study (four auditory and four cross-modal).

If the two measures correspond, they should fall on the diagonal (dashed line). In practice the only points falling near this line are the first-order/first-order comparisons, both auditory and cross-modal. All the other points are quite distant from the diagonal, in some cases by 200 ms. Clearly latency differences, as estimated by RTs, are not sufficient to account for the temporal alignment data.

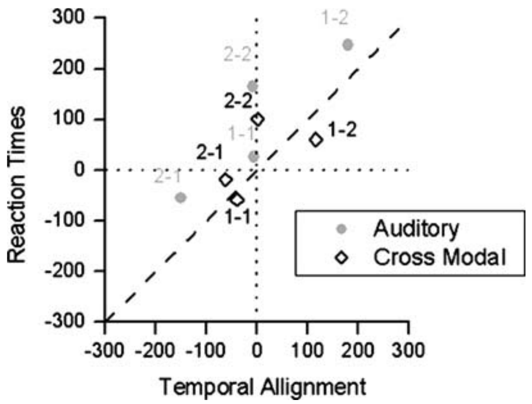


Fig. 8 Differences in reaction times plotted against differences in temporal alignment for the four auditory conditions (*filled circles*, data taken from Fig. 3) and the four cross-modal conditions (*open squares*, data taken from Fig. 6). Standard errors are in all cases smaller than symbol size. The numbers near the data points indicate the comparisons. If the two techniques measure a similar neural property, points should fall near the diagonal (*dashed line*). Only for the two first-order/first-order comparisons is this reasonably true: other points deviate from the diagonal by up to 200 ms

Discussion

The main findings of this study are that perceived synchrony of stimuli oscillating in intensity or frequency does not depend so much on the sensory attribute as on the type of temporal modulation: whether it is first- or second-order. If both intensity and frequency are oscillated abruptly (first-order), then they are perceived to be nearly synchronous. The same is true if both oscillations are second-order. But if one is second-order and the other first-order, then the second-order oscillation must lead by nearly 200 ms to be seen as synchronous, irrespective of whether the oscillation is in intensity or frequency. Similar effects were found in cross-modal comparisons with visually oscillating stimuli: second-order oscillations required a 100–200 ms lead in order to be seen as synchronous. These effects are all very large. Indeed, physically simultaneous first- and second-order oscillations are almost always perceived as asynchronous.

These effects are not readily explained by differential neural latencies derived by RTs, as they do not explain most of the reported effects in perceived synchrony. RTs are a time-honored technique for investigating neural latencies (Galton 1899; Brebner and Welford 1980; Welford 1980). Although they necessarily include a motor component, when RTs to two different stimuli are compared, the motor component is subtracted out. In a previous study we showed that RT estimates agree well with other estimates of neural latencies under similar conditions (Arrighi et al. 2004).

Differences in RTs are plotted in Fig. 3 and 6 together with the stimulus alignment data. It is clear that the RT measures do not in general predict the alignment data. The RT predictions of Fig. 3 and 6 deviate by 50–200 ms from the temporal alignment data, not always in the same direction: RTs overestimate the SOA of subjective alignment in the freq. first/int. second condition of Fig. 3, but underestimate it in the int. first/freq. second condition; for the freq. second/int. second condition, RTs predict a SOA close to 200 ms for perceptual alignment, whereas the empirical value is indeed near zero. This is brought out more clearly in Fig. 8 that plots RTs against temporal alignment for all eight conditions. Only when both frequency and intensity modulations were first-order did the RT prediction agree closely with the empirical data. It is clear that differential neural latencies do not in general provide a good account of the subjective alignment data.

The results of this study agree well with those of Nishida and Johnston (2002), who showed that with oscillating visual stimuli, first-order changes are perceived to precede second-order changes, irrespective of visual attribute. They too observed that their results were not readily predictable by RT estimates of neural latencies. In order to explain their results, Nishida and Johnson advanced the concept of "temporal markers", where second-order oscillations provide a less reliable "temporal marker" than do abrupt changes, and are therefore encoded later. Our results provide broad support for the temporal marker hypothesis, in showing that the nature of the temporal profile—first- or second-order—predicts well the offsets necessary for perceptual temporal alignment. Overall, the temporal marker hypothesis provides the most parsimonious account for these and previous results of temporal alignment (Nishida and Johnston 2002). However, the temporal marker hypothesis remains somewhat poorly defined. It is far from clear why the less reliable temporal markers should be encoded after the reliable ones. It is certainly not just a case of processing time, as the RT data do not predict the alignment data (see Discussion). It seems that salient temporal events receive a precedence in coding, "jumping the cue" to arrive first in consciousness. This hypothesis is reminiscent of the long-standing notion of *prior entry*, which supposes that stimuli that are attentionally selected are processed more quickly than unattended stimuli (Titchener 1908; Reeves and Sperling 1986; Shore et al. 2001; Spence et al. 2001). While attention itself is likely to play little part in the experiments reported here (as subjects were required to pay close attention to both stimuli to make the comparison), but similar mechanisms may apply to visual salience. Salient stimuli, such as first-order oscillations (whose changes are easily localized in time), may benefit from speeded processing with respect to second-order oscillations. However, there is little direct evidence to support the idea at this stage.

A possible explanation for the difference in temporal markers for first and second-order changes may rest with differences of the relative precision with which these changes are encoded. As second-order changes are more subtle (or less salient, as Nishida and Johnston suggest), they could require a longer integration time to achieve a reliable estimate of the moment of the stimulus phase change. This could account for the delay with which second-order changes are perceived relative to first-order changes, in vision, auditory and cross-modal conditions. Our data in fact do suggest that second-order changes are detected with less precision than first-order changes: the best-fitting Gaussians of Fig. 2 and 5 are far more broad for the second-order/second-order comparisons than those for the first-order/first-order comparisons.

Further investigation will be required to determine exactly why first-order changes are perceived earlier than second-order changes. However, this study shows that judgments of temporal order are complex, involving many processes other than simple retinal latencies. One of the key factors affecting temporal order judgments is the temporal profile of the stimulus change. This has not been widely appreciated before, yet as our data show, it is a factor that can outweigh processing latencies as a determinant of subjective alignment.

Acknowledgements We acknowledge support from the Italian Ministry of Universities and Research and an Australian Research Council grant (DP0345797) to D.A.

References

Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. Curr Biol 14:257–262

Arnold DH, Clifford CW, Wenderoth P (2001) Asynchronous processing in vision: color leads motion. Curr Biol 11:596–600

Arrighi R, Alais D, Burr D (2004) Neural latencies do not explain the auditory and audio-visual flash-lag effect. Vision Res (in press)

Aymoz C, Viviani P (2004) Perceptual asynchronies for biological and non-biological visual events. Vision Res 44:1547–1563

Bertelson P, Aschersleben G (2003) Temporal ventriloquism: crossmodal interaction on the time dimension. 1. Evidence from auditory-visual temporal order judgment. Int J Psychophysiol 50:147–155

Brebner J, Welford A (1980) Introduction: an historical background sketch. Academic, New York

Fujisaki W, Shimojo S, Kashino M, Nishida S (2003) Recalibration of audiovisual simultaneity by adaptation to a constant time lag. J Vis 3:34

Galton F (1899) On instruments for (1) testing perception of differences of tint and for (2) determining reaction time. J Anthropol Inst 19:27–29

Hartmann W (2000) Signals, sound and sensation. Springer, Berlin Heidelberg New York

Jaskowski P (1993) Selective attention and temporal-order judgment. Perception 22:681–689

Kopinska A, Harris LR (2004) Simultaneity constancy. Perception 33:1049–1060

Moutoussis K, Zeki S (1997) A direct demonstration of perceptual asynchrony in vision. Proc R Soc Lond B Biol Sci 264:393–399

Nishida S, Johnston A (2002) Marker correspondence, not processing latency, determines temporal binding of visual attributes. Curr Biol 12:359–368

Reeves A, Sperling G (1986) Attention gating in short-term visual memory. Psychol Rev 93:180–206

Roufs JAJ (1963) Perception lag as a function of stimulus luminance. Vision Res 3:81–91

Shore DI, Spence C, Klein RM (2001) Visual prior entry. Psychol Sci 12:205–212

Spence C, Shore DI, Klein RM (2001) Multisensory prior entry. J Exp Psychol Gen 130:799–832

Stelmach LB, Herdman CM (1991) Directed attention and perception of temporal order. J Exp Psychol Hum Percept Perform 17:539–550

Titchener E (1908) Lectures on the elementary psychology of feeling and attention. Ayer, New York

Viviani P, Aymoz C (2001) Colour, form, and movement are not perceived simultaneously. Vision Res 41:2909–2918

Welford A (1980) Choice reaction time: basic concepts. Academic, New York

Zampini M, Shore DI, Spence C (2003) Audiovisual temporal order judgments. Exp Brain Res 152:198–210