Seen in a flash: spatial and temporal aspects in movement related (mis)localization
Maij, F.

2011

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 21. Apr. 2022
Chapter 6

Temporal information can influence spatial localization

Published as:
In order to localize objects relative to ourselves we need to combine various sensory and motor signals. When these signals change abruptly, as information about eye orientation does during saccades, small differences in latency between the signals could introduce localization errors. We examine whether independent temporal information can influence such errors. We asked participants to follow a randomly jumping dot with their eyes and to point at flashes that occurred near the time they made saccades. Such flashes are mislocalized. We presented a tone at different times relative to the flash. We found that the flash was mislocalized as if it had occurred closer in time to the tone. This demonstrates that temporal information is taken into consideration when combining sensory information streams for localization.
Introduction

We combine information from different modalities to make sense of the world. These information sources do not necessarily have the same neuronal latencies. The latencies can differ by tens of milliseconds (Schmolesky et al., 1998). Yet signals that belong together must somehow be related to the same event. It might be that all the different latencies are known and are all taken into account. However, it has been shown that we tolerate artificial differences in timing of up to 100 ms between signals, probably because we know we are uncertain about the timing (Munhall, Gribble, Sacco, & Ward, 1996; van Mierlo, Brenner, & Smeets, 2007).

Tolerating differences in latency will yield errors if the signals are changing fast, as is the case during a saccade. It has been shown that under some conditions flashes presented up to 100 ms before or after a saccade are mislocalized (Matin & Pearce, 1965; reviewed by Schlag & Schlag-Rey, 2002). If this misjudgment is (partly) due to errors in judging time of the flash relative to the eye movement, we should be able to manipulate the errors by changing the judged moment of the flash. To do so, we make use of the expected effect of a tone on the flash. Tones cannot be neglected when judging the number of flashes (Shams, Kamitani, & Shimojo, 2002). Furthermore, it has been shown that a tone can alter the perceived time of a visual stimulus (Morein-Zamir et al., 2003; Vroomen & de Gelder, 2004). We anticipated that people would be unable to neglect a tone presented near the time of a flash when they are required to localize the flash. If the tone and flash are perceived as being one event, we expect the time of that event to be a weighted average of the times of two components (Ernst & Bülthoff, 2004). Therefore, we predict that a tone will influence the (mis)localization of a flash presented near the time of a saccade: presenting a tone just before a flash will have the same effect as presenting the flash earlier (at the same retinal position), and presenting a tone just after the flash will have the same effect as presenting the flash later. This prediction will be quantified with the help of two experiments on explicit temporal judgments and will be tested experimentally in an experiment involving peri-saccadic mislocalization.

Methods

The two temporal judgment experiments will be discussed in the model prediction section. Here we describe the main peri-saccadic mislocalization experiments.

Design and participants

We conducted the experiment in five parts in a normally illuminated room (about 500 lux measured on the table just in front of the participant). The parts only differed in the timing of the tone. Eleven colleagues volunteered to participate in the experiment (including one of the authors). There were four participants in parts 1 and 3 and three in the other three parts. Some participants took part in more than one part of the experiment. Only the author was aware of the specific conditions. All participants had normal or corrected-to-normal vision and normal
hearing. The research in this study was approved by the ethics committee of the Faculty of Human Movement Sciences.

**Experimental setup**

Visual stimuli were presented on a touch screen (EloTouch CRT 19”, 1024x768 pixels, 36 x 27 cm, 85 Hz) using the Psychophysics Toolbox in MATLAB (Brainard, 1997). The visual stimuli were viewed from a distance of 60 cm. Eye movements were registered using an Eyelink II (SR Research Ltd., Mississauga, Ontario, Canada) at a sample frequency of 500 Hz using the Eyelink toolbox (Cornelissen et al., 2002). Participants were asked to follow a 0.5° diameter jumping white dot (108 cd/m²) with their eyes. The dot was presented at a new position every 400 ms (figure 1). It jumped in steps of 7.6° across a gray screen (100 cd/m²). Each jump displaced the dot randomly in one of eight radial directions: horizontal, vertical and diagonal, but never choosing a direction that would bring the dot within 115 pixels (4°) of the edge of the screen.

After a series of 3, 4 or 5 steps (determined at random with equal probabilities) the white dot was removed from the screen. One frame later a 0.5° diameter black dot (7 cd/m²) was flashed (on one frame, which means that it was actually on the screen for less than 1 ms). It was flashed at 2/3 or 4/3 of the 7.6° displacement between the last two positions of the white dot. The flash was thus always 2.5° visual angle from the saccade target. After the flash the screen remained empty (gray) until the participant touched it to indicate where they had perceived the flash. Beside the visual stimuli, we presented a tone (75 dB (A) at the position of the participant; multiple sine-waves with an amplitude that steadily declined to zero in 25 ms) at different moments with respect to the flash. The tone was generated by two loudspeakers that were placed on top of the monitor.

![Figure 1](image)

**Figure 1.** Schematic spatial (A) and temporal (B) overview of a single example trial. The white dot started near the centre of the screen and jumped consecutively in a diagonal, leftward and upward direction. The black dot flashed at 2/3 of the white dot’s last displacement. The flash appeared about one saccadic reaction time after the last white dot appeared on the screen. On this trial a tone was presented 62 ms before the flash.

**Calibration**
To synchronize the eye movement recordings with the images, we measured the moment of the flash with a photo diode and used the signal from the photo diode to drive an IRED that blinded one of the Eyelink cameras. This was done in a separate session and provided the information that we needed to correct for delays between presenting the flash and measuring eye movements. The delay between tone onset and flash was determined using an oscilloscope connected to the above-mentioned photo diode as well as to the audio output. Tone onset was considered as the time of the tone.

Before each session the participant was asked to calibrate the touch screen using a standard nine-point calibration provided by EloTouch. Next, the recording of eye movements was calibrated using the standard nine-point calibration procedure of the Eyelink II.

**Procedure**

Because the mislocalization of the flash only occurs around the moment of the saccade, we wanted to present as many as possible flashes at about that time. We used the saccadic reaction times on previous trials to predict the saccade onset. Dafoe et al. (2007) showed that the saccadic reaction time depends on the saccade direction, so we considered the direction in our predictions. We used the average saccadic reaction time on the five previous trials in which the saccades were in the same direction as that of the trial in question for the prediction. We used Dafoe et al’s (2007) average reaction times for each direction at the beginning of each session. At the predicted saccadic reaction time the black dot was flashed on the screen for one frame at either 2/3 or 4/3 of the last displacement of the white dot. If the predicted reaction time was shorter than the interval between the tone and the flash, the tone was presented immediately after the saccadic target appeared and the flash was delayed accordingly.

The participants were asked to touch the screen at the location at which they saw the black flash. If no new white dots appeared and the participant had not seen a black flash (for instance because he or she blinked), the participant indicated having missed the target by touching the screen in one of the corners. In total there were 480 trials in each session. Participants performed three or four sessions within each part of the experiment (see table 1).

**Table 1. Some experimental details. The conditions are characterized by how long after the flash the tone was presented (in ms).** * One participant performed 4 sessions.

<table>
<thead>
<tr>
<th>Part</th>
<th>Number of participants</th>
<th>Number of sessions</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>3*</td>
<td>-62, 42</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>-62, -4, 42, no tone</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4</td>
<td>-202, -152, -98, -62, -4</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td>-62, 42, 143</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>4</td>
<td>-62, -46, -26, -4</td>
</tr>
</tbody>
</table>
Data Analysis

We used the gaze position data of the right eye to determine the characteristics of the saccades, and the first location at which the finger touched the screen as the perceived position. For an eye movement to be considered to be a saccade, its tangential velocity had to exceed $35^\circ/s$ for at least two consecutive samples (4 ms). We discarded trials in which the touched location differed by more than 200 pixels ($7^\circ$ visual angle) from the actual location of the flash (usually because the participant touched one of the corners, but sometimes they accidently touched the wrong part of the screen; for instance accidently touching the screen with other parts of the hand instead of the planned index finger). We also discarded trials if the direction of the last saccade deviated by more than $22.5^\circ$ from the direction of the last displacement of the white dot, irrespective of saccade amplitude.

We only analyzed the mislocalization in the direction of the saccade: the component of the vector between the touched location and the true location of the flash in the direction of the last displacement of the dot. We plotted these signed errors as a function of the different moments of the flash relative to saccade onset (for individual trials). To draw a smooth curve through these data (for each condition; i.e. each interval between tone onset and flash) we averaged the errors for each participant and condition with a (moving) Gaussian window ($\sigma = 5$ ms). The smooth curve was drawn as long as there were at least five data points within $2\sigma$ of the peak of the Gaussian. We will refer to this curve as the mislocalization curve. We ignored the whole condition for a participant if there was no data for more than 5 ms at any time between 50 ms before and 50 ms after saccade onset because this period is critical for judging the timing of the mislocalization.

To determine whether the pattern of localization errors is shifted in time between the different conditions (which only differed in the timing of the tone), we looked for the shift that would produce the best, single mislocalization curve for each participant in each part of the experiment (figure 2). We determined the time shifts between the conditions that minimize the median squared difference (considering both flashes at $2/3$ and $4/3$ of the last displacement of the white dot) between a single mislocalization curve (for each flash location) and all the data points for that participant in that part of the experiment (see supplementary material for a detailed mathematical description of this method).

This method for finding the temporal shift only yields differences between conditions. In order to align the different parts, and to relate all values to a condition with no tone (while not all parts included a no-tone condition), we combined the above-mentioned differences in three steps. We first aligned the data points of the different participants within each part by minimizing the total between-participant variability across conditions: the average value across conditions was set to the same arbitrary value for all participants. We then aligned the temporal shifts between parts (without shifting the relative positions within each part) on the basis of the overall average values of the common conditions. Finally, the average value of the no-tone condition was considered
to be the origin (i.e. zero; see supplementary material for a detailed mathematical description of the method).

Figure 2. Schematic illustration of how the temporal shift of the mislocalization patterns ($\Delta(t)$) was determined. **A.** Hypothetical data with a clear shift in time between the mislocalization in the two conditions (both with a flash at 2/3 and 4/3 of the last displacement of the white dot). The black curve is the smoothed average of all data points. **B.** The data points of condition 2 have been shifted in time to find the best fit (the value of $\Delta(t)$ that minimizes the median squared deviation of the points from the black curves for both flash locations).

**Model prediction**
We used a model based on the following reasoning to predict the temporal shift that a tone will induce in the mislocalization pattern. If there is some uncertainty about the timing of the flash, and a tone is presented near the moment of the flash, the participant may judge the flash and the tone to have resulted from the same event. If the interval between the flash and the tone is long, they will be considered to arise from separate events. We describe the probability of considering the flash and the tone to have occurred simultaneously by a normal distribution with parameters that will be estimated later (thick line in figure 3A). If the tone and flash are perceived to originate from the same event, the moment of the flash can best be judged from a weighted average of the timing of the two signals. Such weighted averaging will cause a shift in the judged moment of the flash towards the time of the tone (figure 3B).

Our prediction for the influence of the timing of the tone on the timing of the mislocalization of the flash ($\Delta(t)$) is the product of the probability of considering the flash and tone to have occurred simultaneously and the
influence that the tone has when they are considered to occur simultaneously:

\[ \Delta(t) = wt \cdot e^{-\frac{(t-b)^2}{2\sigma^2}} \]  

(1)

where \( t \) is the time difference between the flash and the tone, \( w \) is the weight given to the tone when determining the time of the combined event from the estimates of the timing of the flash and tone (slope in figure 3B), \( b \) is the offset between the flash and the tone for which the two are most likely to be considered synchronous, \( \sigma \) is the width of the distribution of times for which flash and tone are considered to arise from the same origin, and \( c \) is the peak likelihood of considering the tone and flash to be simultaneous (figure 3A). Figure 3C shows our prediction for the temporal shift of the mislocalization curves (the product of 3A and 3B).

In order to estimate reasonable values for this interpretation (and thus to be able to draw curves in figure 3) we performed two temporal judgment experiments. Both were performed after the main experiment. To determine the moment of simultaneity we showed 14 participants the same displays as in the main experiment and instructed them to follow the white dots in the same manner, but rather than asking them to localize the flash we asked them to report whether or not the tone and the flash occurred simultaneously. The participants touched the screen in one of the two lower corners to indicate whether the flash and tone occurred simultaneously or not. We fit a normal distribution to each participant’s data by using psignifit version 2.5.6 (see http://bootstrap-software.org/psignifit/), a software package which implements the maximum-likelihood method described by Wichmann and Hill (2001). The median values of the parameters of these fits were an amplitude (\( c \)) of 0.82, an uncertainty (\( \sigma \)) of 82 ms and a peak when the tone is presented 31 ms before the flash (\( b = -31 \); figure 3A).

To determine the weight given to the tone in the perception of the time of the flash (slope of figure 3B) we showed 14 participants a sequence of three black flashes (same specifications as in the main experiment) at the screen centre. The third flash was 1200 ms after the first whereas the second was between 400 ms and 800 ms after the first (steps of 50 ms). The participants were asked to judge whether the first or the second interval was longer by pressing the lower left or right corner of the touch screen. The second flash was accompanied by a tone that was presented either 46 ms before or 42 ms after the flash. We fit each participant’s data for each interval between tone and flash with a psychometric function to estimate the points of subjective equality. Again psychometric functions were fit using psignifit. The difference between the points of subjective equality for the two intervals between tone and flash gives us a measure of the influence of the tone (for each participant). By dividing the median difference between the points of subjective equality by the difference between the times of the tones (-46 ms and 42 ms), after correcting for the likelihood of considering the tone to arise from the same events (on the basis of the values from the simultaneity judgment experiment at those times), we estimated that the median weight (\( w \)) given to the tone was 0.53 (figure 3B).
Temporal information can influence spatial localization

Figure 3. Model predicting the temporal shift of the mislocalization pattern as a result of presenting a tone at various times. A. Probability of the flash and tone being considered to arise from the same event. The thick curve is based on the median values from fits of normal distributions to 14 participants’ simultaneity judgments (dots connected by thin lines). B. The influence of the tone on the judged time of the flash if the two are considered to arise from the same event is a weighted average of the time of the flash and the tone. The thin lines are individual participants’ data and the thick line is the median value based on 14 participants’ interval judgments. C. Predicted temporal shift as function of the time (t) of the tone relative to the flash: product of the values in A and B. Since interpreting the retinal stimulation by the flash as if it had occurred later means that the eye orientation signal will influence earlier flashes, the sign is inverted.

Results
Eye movements
We obtained useful localization judgments from 21,227 trials. This is 70% ± 10% (mean ± standard deviation across participants) of the approximately 30,000 trials. For 22% of the trials the judgments were ignored because there was no detectable saccade near the moment of the flash (including trials in which the participant blinked). In a further 3% of the trials the deviation of the direction of the saccade differed too much (>22.5°) from the expected saccade direction. Another 1% of the trials was discarded because the screen was touched more than 200 pixels (7°) from the actual location of the flash. 90% of these missed flashes had been presented at 4/3 of the last displacement of the white dot. This occurred as frequently when the tone was presented before as when it was presented after the flash. The last 4% of the total number of trials was discarded because we removed the whole condition...
for a participant if there were more than 5 ms of data missing in the critical period for judging the timing of the mislocalization (between 50 ms before and 50 ms after the saccade onset). This was so for the trials of 3 of the 4 participants when we presented a tone 202 ms before the flash in part 3.

Earlier research showed that saccadic reaction times depend on the radial direction (Dafoe et al., 2007). We found similar results (figure 4A). Presenting the tone before the flash means that on average it was presented before the saccade, because on average the flash was presented at the time of the saccade. We found no evidence that presenting a tone before the flash decreases the saccadic reaction time (illustrated for part 2 in figure 4B; see Table 1S of the supplementary material for the other parts) Also for various other saccade parameters we found no dependency on the condition (see Table 1S of the supplementary material).

Figure 4. Saccadic reaction times in part 2. A. Times for target jumps in different directions with standard errors across participants. Reaction times were largest for downward saccades. B. Reaction times for different moments of (or absence of) the tone. The axis on the right applies to both panels A and B.

Mislocalization
From the upper panel of figure 5 we can see that participants make systematic errors when targets are presented near the time of saccade onset (the left edge of the gray bar), and that the timing of these errors depends on when the tone was presented. The mislocalization curve is shifted to the right in the condition in which the tone was presented 62 ms before the flash (green) compared to when it was presented 42 ms after the flash (red). A rightward shift indicates that flashes presented at a certain time were perceived as if they were presented earlier. This is consistent with participants perceiving the flash at an earlier moment if it was preceded by a tone than if it was followed by a tone.

In order to see whether the differences between the conditions were really pure time shifts we calculated the residual variability with respect to that single mislocalization curve for each condition. These residuals were then smoothed with a (moving) Gaussian window ($\sigma = 5$ ms). The fact that the curves in the lower panels of figure 5 are flat and similar for both conditions shows that shifting the data in time captures most of the variability between the conditions.
Temporal information can influence spatial localization

Figure 5. Pattern of mislocalization in part 1. Data for targets flashed at 2/3 (left panels) and 4/3 (right panels) of the last displacement of the white dot. The gray bar indicates the average duration of the saccade (aligned at onset). Upper panels: green and red dots show one participants’ errors for individual flashes when the tone was presented 62 ms before the flash and 42 ms after the flash, respectively. The thick lines are smoothed averages of the dots of the same color. Curves for the other participants (thin lines) are shown without the raw data. The horizontal continuous line indicates the position of the flash (no mislocalization). The dashed line indicates the position of the saccade target. Lower panels: mislocalization that cannot be accounted for by shifting a single mislocalization curve in time (to obtain $\Delta(t)$). The curves show the difference between the mislocalization curves of the upper panel and shifted best fitting single average curves for each participant.

In figure 6 we show the mislocalization curves (before any time shifts) for each condition of the other parts (see Table 1). We noticed that some participants showed peculiar patterns of mislocalization. For example the participant in part 2 whose data are shown in figure 6A shows an overall systematic error that is independent of the timing for flashes presented at 4/3 of the last displacement of the white dot. Dassonville et al. (1992) attributed similar results to misjudged retinal eccentricity (evidence for misjudged eccentricity can be found in studies such as Müßeler, van der Heijden, Mahmud, Deubel, & Ertsey, 1999). Importantly, in our study participants that showed unusual patterns of mislocalization showed a similar temporal shift as the others, so such participants’ behavior need not be considered separately. One other aspect that is worth mentioning is that the mislocalization was similar for the eight radial directions that we used (not shown), which is in line with earlier reports (Honda, 1991). There is no overall bias towards mislocalizing targets in a specific direction in space, such as the location of the loudspeakers (also not shown).
**Temporal shift**

The data in figure 6 are typical examples, showing only one participant per part. The best fitting shift values for each participant in each condition are shown in figure 7. How the data of the different participants and parts are aligned is explained in the data analysis section of the methods (and in the supplementary material). The temporal shift was about 20 ms when the tone was presented 62 ms before the flash and about 5 ms when the tone was presented 42 ms after the flash. The dashed line in figure 7 shows the model prediction with the parameters from the two temporal judgment experiments. This prediction overestimates the effect of the tone.

![Image of mislocalization curves](image_url)

*Figure 6. Mislocalization curves in part 2 (A), part 3 (B), part 4 (C) and part 5 (D). Data of a different participant are presented for each part. See figure 5 for other details.*

The solid line in figure 7 shows a fit of our model (equation 1) to the data (solid line in figure 7). The following parameters fit our data best: \( c \cdot w = 0.28 \), \( b = -51 \text{ ms} \) and \( \sigma = 65 \text{ ms} \). Based on the value of \( c \) from the simultaneity judgments of 0.82, the weight \( w \) given to the tone is about 0.34. A \( \chi^2 \)-test (Press, Flannery, Teukolsky, & Vetterling, 1992) indicates, that the data do not deviate
Temporal information can influence spatial localization significantly from this model ($\chi^2(54) = 62.22, p > .05$). A similar test indicates that we can reject a model whereby the tone has no effect ($\chi^2(57) = 303.78, p < .001$; i.e. the data do deviate significantly from zero). The data is also significantly different from the model prediction with the parameters from the two temporal judgment experiments ($\chi^2(54) = 149.10, p < 0.001$).

**Figure 7.** Temporal shift of the mislocalization curve as a function of the interval between tone onset and flash. Each symbol represents the best shift for one participant in one condition of a part of the experiment (considering both flashes at 2/3 and 4/3 of the last displacement of the white dot). Different symbols represent different parts. The dashed purple line represents the model of equation 1 with parameter values that were predicted from the two temporal judgment experiments (reproduced from figure 3C). The purple solid line represents the best fit of equation 1 to the data.

**Discussion and conclusion**

In this study we examined whether the presence of an irrelevant tone near the time of a flash influences the location at which the flash is perceived. We presented the tone at different moments with respect to the flash and found a temporal shift of the mislocalization as a result of doing so. This temporal shift was largest when the tone was presented 40-100 ms before the flash, and negligible when the tone was presented more than 150 ms after the flash or more than 200 ms before the flash. When it had an effect, presenting a tone influenced the perceived location of the flash in the way that presenting the flash nearer to the time of the tone would have done. We interpret this as evidence that the tone changes the perceived time of the flash. Note that this happened although participants were not instructed to pay attention to the tone, so we can consider the tone to be task irrelevant. In this respect our study differs from a study by Binda et al. (2007) where tones were presented at different positions and the position of the tone was relevant for the task.

We presented the flash at two different locations: at 2/3 and 4/3 of the distance between the last displacements of the white dot. Having different locations revealed that the mislocalization was mainly in the direction of the saccade target. Some other studies have reported such compression of the locations of flashed targets (Ross et al., 1997), while others did not find such
compression (Dassonville et al., 1992). This compression is still not fully understood. The compression is not found when the experiment is performed in the dark (Dassonville et al., 1992), the critical issue being whether there are visual references after the saccade (Lappe et al., 2000; Morrone, Ma-Wyatt et al., 2005), but even studies performed in a dimly lit room do not always show this compression (Brenner et al., 2005). The compression has been related to saccadic speed, because it is negligible for saccades that are smaller than 5° (Ostendorf et al., 2007), or to changing receptive field size and location (Hamker, Zirnsak, Calow, & Lappe, 2008). Compression has also been related to temporal uncertainty (Brenner, van Beers, Rotman, & Smeets, 2006). Our results support the latter interpretation by showing that there is indeed uncertainty about the time of the flash. We here show that the compression does not only occur for isolated saccades made repeatedly to the same position (e.g. Lappe et al., 2000; Morrone, Ross et al., 2005), but also during continuous scanning behavior, with frequent saccades in unpredictable directions.

Our main results are nicely summarized by equation 1. When we compare the estimates of \( w \), \( b \) and \( \sigma \) from the two temporal judgment experiments described in the model section \((w \cdot c = 0.43, b = -31\text{ ms}, \sigma = 82\text{ ms})\) with the results of the fit to the main experiment \((w \cdot c = 0.28, b = -51\text{ ms}, \sigma = 65\text{ ms})\) we find that the values are similar but not identical. We are not concerned by these differences because the variability of the temporal judgment tasks across participants is large (thin lines in figure 3). Moreover, it has been shown that the precise values of parameters for temporal precision depend strongly on the methods used (Vatakis, Navarra, Soto-Faraco, & Spence, 2008), and our two temporal judgment tasks are not completely comparable to the main experiment. For instance, in the main experiment we did not instruct the participants to pay attention to the tone, whereas in the simultaneity judgment task (which was therefore performed later) we obviously did. Similarly, in the main experiment the flashes were presented at unpredictable locations on the screen. Whereas in the experiment in which a sequence of three flashes was presented the flashes were always presented at the centre of the screen. Moreover, in the main experiment the participant was making saccades to follow the randomly moving dot, whereas in the three-flash experiment the participant was probably fixating the screen centre.

To conclude, we show that introducing an irrelevant tone at different moments with respect to the time of a flash shifts the pattern of mislocalization of flashes presented near the time of saccades. We show that a model based on weighted averaging of the judged times of the flash and the tone, and which considers the probability of the two being perceived as arising from one event, provides a good description of the data. We therefore conclude that temporal information is taken into consideration when combining spatial information from different sensory streams for localization.
Supplementary material
S1. Saccade parameters
The main experiment consisted of five parts, with three or four participants for each part. In order to determine whether the saccades were influenced by the tones, we determined the median saccade duration, peak velocity, latency and amplitude for each participant, part and condition. The values in table 1S are averages of these median values across participants (mean ± standard error). None of the parameters showed any dependency on the timing of the tone. We identified saccades by the tangential velocity exceeding 35°/s. Note that saccade latency differences between conditions within a part are smaller than between the same condition in different parts and much smaller than those between directions (shown for part 2 in figure 4 of the main article).

Table 1S. Saccade parameters for each part and condition.
S2. Temporal shift

To determine whether the pattern of localization errors is shifted in time between the different conditions (which only differed in the timing of the tone), we determined the temporal shifts for each of the conditions that would produce the smallest deviations around a single mislocalization curve for each session and flash location. We did this by minimizing the median squared difference between all (shifted) data points (considering both flashes at 2/3 and 4/3 of the last displacement of the white dot) and a single mislocalization curve (a curve through all the shifted data points that was created by smoothing the data by averaging the values with weights defined by a moving Gaussian window). We used the median rather than the mean because it is less sensitive to outliers. For each participant in each part:

\[ SS_{op}(\Delta) = \frac{1}{2} \sum_{f=1}^{2} \sum_{i=1}^{n} \mu_{k}^{i} \left( \left( M_{fop}^{(\Delta)}(tf) - x_{kiop}(tf + \Delta'_{op}) \right)^{2} \right) \]  

where \( x_{kiop}(tf + \Delta'_{op}) \) is the indicated position on trial \( k \) attributed to a time that is shifted by \( \Delta'_{op} \) and \( M_{fop}^{(\Delta)}(tf) \) is the value of the smooth mislocalization curve that takes into account all the temporal shifts in the part. The best fitting temporal shift was determined simultaneously for both flash locations (2/3 and 4/3 of the last displacement of the white dot), but obviously with a different mislocalization curve for each flash location. Adding the same time to all values of \( \Delta \) does not influence the value of \( SS_{op}(\Delta) \). The -62 ms conditions was present in all parts so we considered it as a baseline (\( \Delta(t=-62)=0 \)), but this choice is arbitrarily and does not influence the final values. In Table 2S we show the values of the temporal shift relative to the -62 ms condition for each condition, participant and part (\( \Delta'_{op} \)). In the next section we describe how we align the temporal shifts to the no tone condition.
Table 2S. The temporal shifts as determined by the best fit of the data points to a mislocalization curve (see equation 1) for each part and participant. These values are then aligned across participants and parts (see equations 2 to 4) to obtain the values shown in figure 7 (the change in sign is because having to shift the data points in a certain direction implies that the tone shifted them in the opposite direction).

<table>
<thead>
<tr>
<th>Part.Participant</th>
<th>No tone</th>
<th>-202 ms</th>
<th>-152 ms</th>
<th>-98 ms</th>
<th>-62 ms</th>
<th>-46 ms</th>
<th>-26 ms</th>
<th>-4 ms</th>
<th>42 ms</th>
<th>143 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>0</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>0</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>0</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>0</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>18</td>
<td>0</td>
<td>12</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>25</td>
<td>0</td>
<td>32</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>16</td>
<td>0</td>
<td>-5</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>26</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>-1</td>
<td>5</td>
<td>0</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>-1</td>
<td>-7</td>
<td>0</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>0</td>
<td></td>
<td></td>
<td>1</td>
<td>-1</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>0</td>
<td></td>
<td></td>
<td>11</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>0</td>
<td>-11</td>
<td>0</td>
<td>-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2</td>
<td>0</td>
<td>3</td>
<td>23</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.3</td>
<td>0</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
S3. Aligning the temporal shifts

The method for finding temporal shifts that is described in section S2 only yields differences between conditions within a part. In order to align the different parts, and to relate all the values to the condition with no tone (while not all parts included a no-tone condition), we combined the above-mentioned differences in three steps. We first aligned the shifts of the different participants within each part by minimizing the total between-participant variability across conditions. This was achieved by subtracting the same time from all of each participant’s temporal shifts, so that rather than the value for the -62 condition being zero, the average value across all conditions (for every participant in each part) was zero:

\[ \Delta''_{lop} = \Delta'_{lop} - \frac{1}{n} \sum_{i=1}^{n} \Delta'_{lop} \]  
\[ \text{(s2)} \]

We then aligned the temporal shifts between parts (without shifting any of the relative positions within each part) on the basis of the overall average values across participants of the common conditions of each part with part 2.

\[ \Delta'''_{lop} = \Delta''_{lop} - \frac{1}{m \cdot n'} \sum'_{i=1}^{m} \sum_{o=1}^{n'} \Delta''_{lop} \]  
\[ \text{(s3)} \]

Finally, we subtracted the average value of the no-tone condition \((i = 1 \text{ and } p = 2)\) from all temporal shifts so that the average value for the no-tone condition is zero.

\[ \Delta_{op} (t) = \Delta'''_{lop} - \frac{1}{m} \sum_{o=1}^{n} \Delta'''_{lop} \]  
\[ \text{(s4)} \]

The values of \( \Delta_{op} (t) \) are shown in figure 7 of the main article for each part \((p; \text{each part is represented by a different symbol}), \text{condition } (t; \text{the position on the ordinate}) \text{ and participant } (o). \)
Temporal information can influence spatial localization