Supplemental Material (SOM-R)

TABLES

<table>
<tr><th>ATTEND
stimuli
LEFT | RIGHT</th><th>AV</th><th>A</th><th>V</th></tr>
<tr><td>A | V</td><td>12</td><td>12</td><td>12</td></tr>
<tr><td>V | A</td><td>12</td><td>12</td><td>12</td></tr>
<tr><td>A | A</td><td>12</td><td>12</td><td>0</td></tr>
<tr><td>V | V</td><td>12</td><td>0</td><td>12</td></tr>
</table>

Table S1: Number of trials per experimental condition.

FIGURES

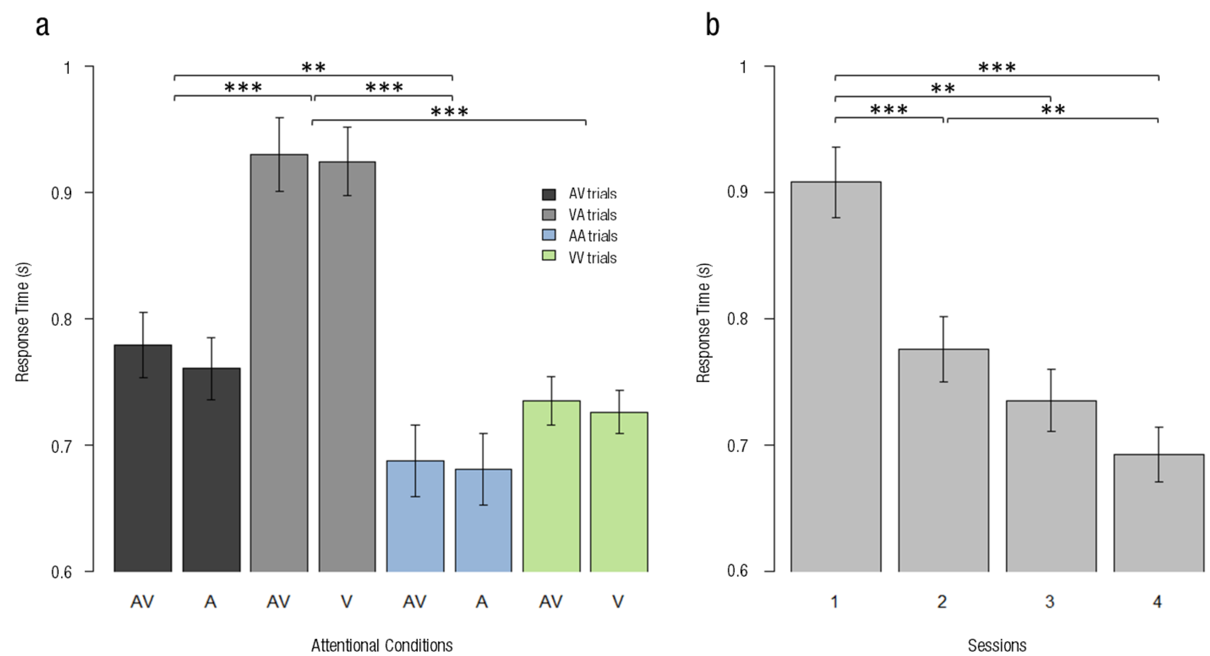

Fig. S1. Although instructions emphasized accuracy over speed, we report below the quantification of RT. **No effect of prior-entry on RT:** The law of prior-entry predicts faster responses when attention is oriented to the first presented stimulus. Hence, RTs were expected to be faster in unisensory trials when the stimulus in the attended modality was presented first as compared to when attention was split across sensory modalities. For audiovisual (AV) trials, RTs were expected to be faster when attention was oriented to the sensory modality of the first presented stimulus as compared to when attention was split across sensory modalities. RTs were separately averaged for visual, auditory, audiovisual audio-first and audiovisual visual-first trials, and for each attentional condition. RTs above 4 s were excluded from the analysis. A three-way ANOVA with factors of attention (3: auditory, visual, split), condition (4: visual, auditory, audiovisual audio-first, audiovisual visual-first), session (4) and a nested random effect modeling the between-subjects variability was performed. The main effect of attention was not significant ($F(1,18) = 1.27$, $p > .250$, $\eta^2_p = .066$), which may be due to task instructions emphasizing accuracy over speed, as well as the fact that prior-entry effects on RT have mostly been reported for single stimuli detection (Gibbon & Rutschmann, 1969; Sternberg & Knoll, 1973; Neumann, Esselmann, & Klotz, 1993; Shore, Spence, & Klein, Visual prior-entry, 2001). **RT in audiovisual TOJ is faster when the sound is presented first:** The main effect of conditions was significant ($F(3,18) = 26.01$, $p < .001$, $\eta^2_p = .591$). A post-hoc Bonferroni-corrected paired t-test revealed that RTs in auditory (Fig. S1a, A) and visual (Fig. S1a, V) unisensory trials were not significantly different ($t(18) = 1.44$, $p > .250$, $d = .20$). In audiovisual trials (Fig. S1a, AV), audio-first (black) RTs were significantly faster than visual-first (gray) RTs (769 +/- 12 ms and 927 +/- 14 ms, respectively; $t(18) = 7.30$, $p < .001$, $d = .62$) indicating that, irrespective of attentional instructions and spatial side, a sound presented first yielded a faster response by ~158 ms on average as compared to a flash presented first (Fig. S1b). This is congruent with reports of asymmetry between audition and vision in temporal tasks: participants typically detect more accurately asynchrony when audition is leading vision (van Eijk, Kohlrausch, Juola, & van de Par, Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type, 2008) and visual leads stimuli need larger temporal integration/encoding window than audio lead stimuli (van Wassenhove, Grant, & Poeppel, 2007). Still, audio-first trials in audiovisual conditions were slower than the auditory unisensory trials (684 +/- 14 ms, $t(18) = -3.63$, $p = .011$, $d = .34$) and not significantly different from the visual unisensory trials (730 +/- 9ms, $t(18) = -1.39$, $p > .250$, $d = .19$). **Overall RT decreases over sessions and is thus decorrelated with the stability of PSS and prior-entry over time:** interestingly, the session factor is significant ($F(3,18) = 20.18$, $p < .001$, $\eta^2_p = .528$). A second post-hoc Bonferroni-corrected paired t-test (Fig. S1a) showed that RTs were significantly different between session 1 and 2 ($t(18) = -4.66$, $p = .001$, $d = .51$), 1 and 3 ($t(18) = -4.36$, $p = .002$, $d = .69$), 1 and 4 ($t(18) = -5.73$, $p < .001$, $d = .91$), and 2 and 4 ($t(18) = -3.45$, $p = .017$, $d = .36$). Overall, the mean RT decreased over time (908 +/- 14ms, 776 +/- 13ms, 735 +/-12ms, 692 +/- 11ms, respectively for each session). RT was thus largely de-correlated from PSS, stable over months, consistently with previous studies showing that PSS and RT behave differently (Jaskowsi, 1999).

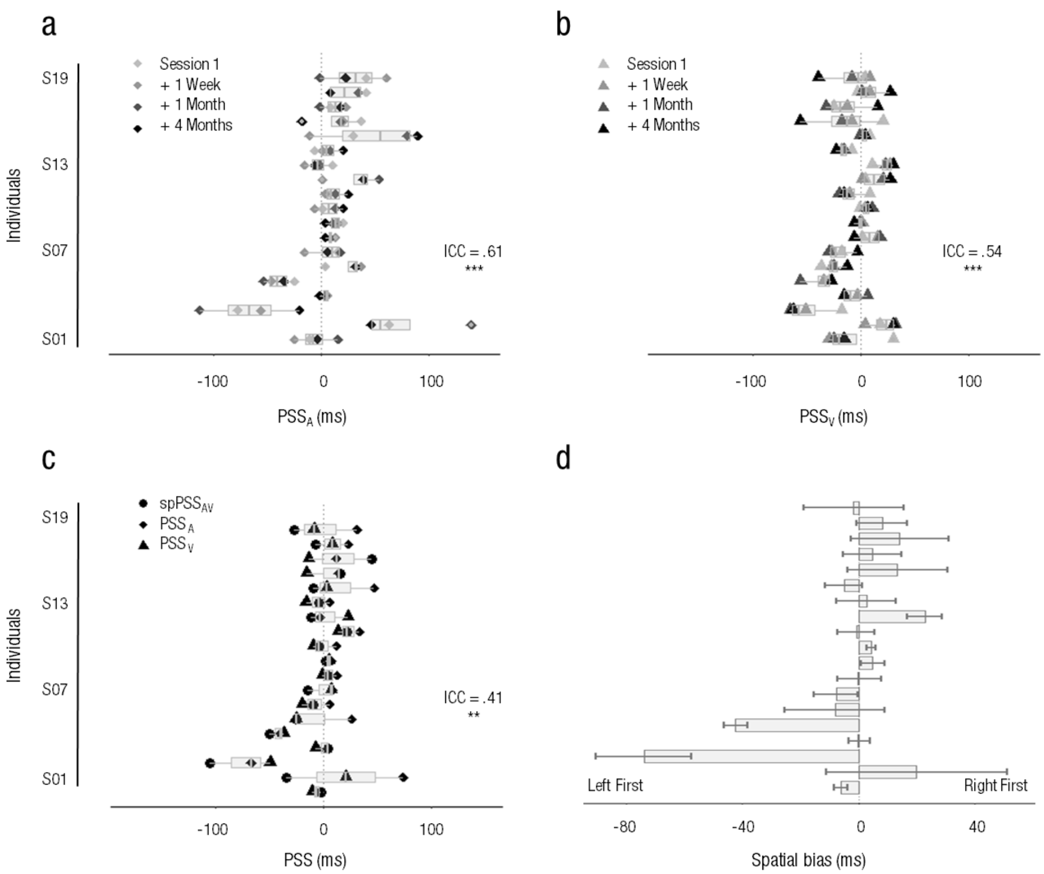

Fig. S2. PSS: a time or a space bias? We tested the stability of unisensory PSSs collected in unisensory auditory (AA, PSS_A) and visual (VV, PSS_V) trials using ICC. PSS_A (Fig. S2a) showed a significant ICC of .61 ($CI_{95\%}$ = [.39, .80], $F(18,38) = 7.19$, $p < .001$). PSS_V (Fig. S2b) showed a significant ICC of .54 ($CI_{95\%}$ = [.32, .75], $F(18,38) = 5.68$, $p < .001$). Hence, both unisensory PSSs were stable over time for all individuals. PSSs were sorted as previously, namely according to audiovisual PSS (PSS_{AV}). As PSS_A and PSS_V were computed on the basis of the '*right stimulus first*' responses, they could also be interpreted as a spatial bias, leaving unclear whether what was being quantified was effectively a temporal or a spatial order threshold in the unisensory conditions of the task. However, in the audiovisual conditions, sensory modality and spatial position were orthogonal dimensions so that the spatial bias could be computed by disregarding the sensory modality of the spatialized stimuli. Thus, we computed the spatial bias in the split-attention conditions ($spPSS_{AV}$) using the audiovisual trials and preserving the '*right stimulus first*' responses irrespective of the sensory modality. $spPSS_{AV}$ was significantly stable (ICC = .53, $CI_{95\%}$ = [.31, .75], $F(18,38) = 5.53$, $p < .001$) suggesting that individuals do have a spatial bias irrespective of the sensory modality. We then performed an ICC comparing all three measures (PSS_A, PSS_V and $spPSS_{AV}$) combined across all four experimental sessions (Fig. S2c). The resulting ICC of .41 was significant ($CI_{95\%}$ = [.13, .68], $F(18,38) = 3.10$, $p = .002$) indicating a reliable consistency among the three PSSs (Fig. S2c) supporting the notion of an individual's spatial bias. Combining all measures, we found that 11 out of 19 participants perceived more often the left stimulus first (Fig. S2d) consistent with a reported perceptual bias towards the left visual field (Matthews & Welch, 2015) providing support to the possibility of a spatialized bias or mental time line for the estimation of temporal ordinality (e.g. Bonato *et al.*, 2012).

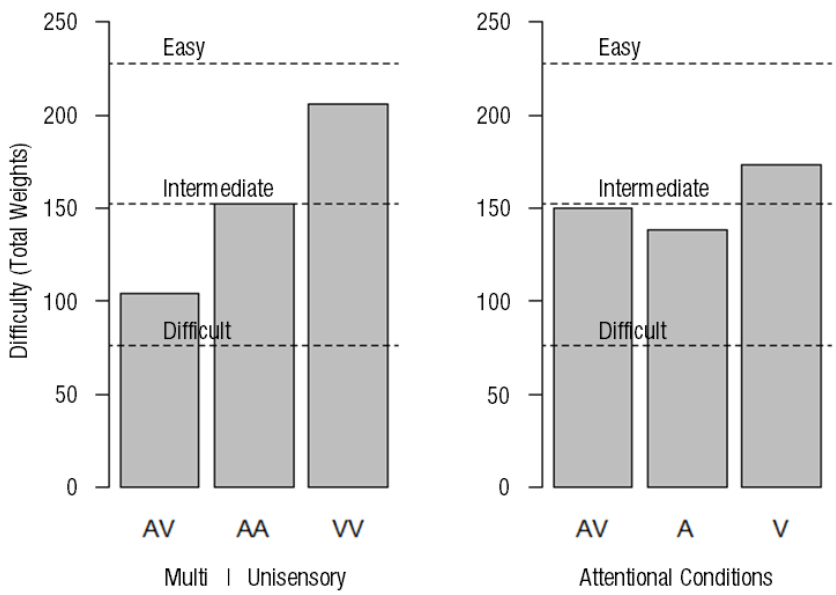

Fig. S3. Participants' introspection on task difficulty. At the end of each experimental session, participants were asked to rank their felt easiness with which they accomplished the TOJ task according to the sensory modality (AV: audiovisual; AA: two sounds; VV: two flashes) or to the attentional condition (AV: split-attention; A: auditory attention; V: visual attention). Each participant's response was weighted according to the answer: 3 for the easiest and 1 for the most difficult one. The left panel reports the sum of weights as function of the sensory modality of the stimuli pair and the right panel as a function of attentional condition. The dashed lines correspond to the maximal weight for each response. Overall, the audiovisual TOJ was ranked as most difficult and the visual TOJ as easiest. Attentional conditions were more comparably ranked.

References

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: evidence for a mental time line. *Neuroscience & Biobehavioral Reviews*, 36(10), 2257-2273.

Gibbon, J., & Rutschmann, R. (1969). Temporal order judgment and reaction time. *Science, 165(3891)*, 413-415.

Jaskowsi, P. (1999). *Cognitive Contributions to the Perception of Spatial and Temporal Events*. Elsevier.

Matthews, N., & Welch, L. (2015). Left visual field attentional advantage in judging simultaneity and temporal order. *Journal of Vision, 15(2)*, 1-13.

Neumann, O., Esselmann, U., & Klotz, W. (1993). Differential effects of visual-spatial attention on response latency and temporal-order judgment. *Psychological Research, 56(1)*, 26-34.

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior-entry. *American Psychological Society, 12(3)*, 205-212.

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: fundamental issues and a general model. *Attention & Performance*, 630-681.

van Eijk, R. L., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics, 70(6)*, 955-968.

van Wassenhove, V. v., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45(3)*, 598-607.