

**Interactions Between Auditory Elevation, Auditory Pitch and Visual  
Elevation during Multisensory Perception**

Yaseen Jamal<sup>1</sup>, Simon Lacey<sup>1</sup>, Lynne Nygaard<sup>2</sup> and K. Sathian<sup>1,2,3,4</sup>

<sup>1</sup>Department of Neurology, <sup>2</sup>Department of Psychology, <sup>3</sup>Department of Rehabilitation  
Medicine,

Emory University, Atlanta, GA 30322, USA

<sup>4</sup>Center for Visual and Neurocognitive Rehabilitation, Atlanta VAMC, Decatur, GA 30033,  
USA

**Supplementary Material**

## Supplementary results

### *Response times (RTs)*

For the initial analysis, in each discrimination task, data from the 8 stimulus types were sorted according to correspondence type (within-modal featural, cross-modal featural and cross-modal spatial) and trial type (congruent vs. incongruent) by averaging response times (RTs) from the two stimulus types containing the correspondence of interest. For example, to examine the within-modal featural correspondence between auditory pitch and auditory elevation (AP-AE), we averaged over stimulus types A and B of Table 1 in the main text for congruent trials and over stimulus types C and D (of Table 1 in the main text) for incongruent trials. We performed similar averaging for the cross-modal featural (auditory pitch – visual elevation, AP-VE) and cross-modal spatial (auditory elevation and visual elevation, AE-VE) correspondences: For the AP-VE correspondence, this meant averaging across stimulus types A and C for congruent trials and over stimulus types B and D for incongruent trials. For the AE-VE correspondence, averaging was performed across stimulus types A and D for congruent trials and over stimulus types B and C for incongruent trials.

### *RTs: global analysis*

A global, repeated-measures ANOVA (RM-ANOVA) with factors of attended feature (auditory pitch, auditory elevation, visual elevation), correspondence type (AP-AE, AP-VE and AE-VE), and trial type (congruent, incongruent) showed a significant main effect of attended feature ( $F_{2,44} = 33.9$ ,  $p < .001$ ,  $\eta^2 = .6$ ). *Post-hoc* t tests (Bonferroni-corrected  $\alpha = .017$ ) showed that RTs for discrimination of auditory pitch ( $749 \pm 41$ ms) and auditory elevation ( $762 \pm 45$ ms) were not significantly different ( $t_{22} = .33$ ,  $p = .7$ ), but both were significantly longer than RTs for discrimination of visual elevation ( $505 \pm 31$ ms:  $t_{23} = 6.9$ ,  $p < .001$  and  $t_{22} = 6.4$ ,  $p < .001$  respectively).

There was also a significant main effect of trial type ( $F_{1,22} = 60.5$ ,  $p < .001$ ,  $\eta^2 = .7$ ), congruent trial RTs ( $655 \pm 33$ ms) being faster than incongruent trial RTs ( $682 \pm 35$ ms). The main effect of correspondence type was not significant ( $F_{2,44} = 1.5$ ,  $p = .2$ ,  $\eta^2 = .06$ ).

There was a significant attended feature x trial type interaction ( $F_{2,44} = 17.9$ ,  $p < .001$ ,  $\eta^2 = .4$ ) in which RTs were faster for congruent than incongruent trials during discrimination of auditory elevation and pitch (Bonferroni-corrected  $\alpha = .0166$ :  $t_{22} = -6.5$ ,  $p < .001$  and  $t_{23} = -4.8$ ,  $p < .001$  respectively) but not during discrimination of visual elevation ( $t_{23} = 2.1$ ,  $p = .04$ ).

There was a significant correspondence type x trial type interaction ( $F_{2,44} = 9.2$ ,  $p < .001$ ,  $\eta^2 = .3$ ) in which congruent trial RTs were faster than incongruent RTs for each correspondence type (AP-AE  $t_{23} = 5.0$ ,  $p < .001$ ; AP-VE  $t_{23} = 3.8$ ,  $p = .001$ ; AE-VE  $t_{23} = 5.7$ ,  $p < .001$ ), the difference between congruent and incongruent trials being greatest for the AE-VE, and least for the AP-VE,

correspondences. Finally, there was a 3-way interaction between attended feature, correspondence type, and trial type ( $F_{4,88} = 29.4$ ,  $p < .001$ ,  $\eta^2 = .6$ ).

We investigated the third order interaction by conducting RM-ANOVAs, with factors of correspondence type and trial type, for each attended feature separately.

For the visual elevation condition, congruent trial RTs were significantly faster than incongruent trial RTs ( $F_{1,23} = 4.6$ ,  $p = .04$ ) but there was no effect of correspondence type and no significant interaction (Supplementary Figure S1). Note that the main effect for trial type is formally the same as that tested in the attended feature x trial type interaction above that failed Bonferroni correction; although significant here, the absolute difference between congruent and incongruent RTs was quite small (503ms vs 506ms).

In the auditory pitch condition, congruent trial RTs ( $734 \pm 39$ ms) were significantly faster than incongruent trial RTs ( $764 \pm 42$ ms:  $F_{1,23} = 22.8$ ,  $p < .001$ ,  $\eta^2 = .5$ ) and correspondence interacted significantly with trial type ( $F_{2,46} = 9.1$ ,  $p < .001$ ,  $\eta^2 = .3$ ). Congruent trial RTs were faster for the AP-AE ( $t_{23} = 4.6$ ,  $p < .001$ ) and AP-VE ( $t_{23} = 4.4$ ,  $p < .001$ ) correspondences, but not the AE-VE, correspondence (Supplementary Figure S2).

For the auditory elevation condition, congruent trial RTs ( $737 \pm 44$ ms) were again significantly faster than incongruent trial RTs overall ( $788 \pm 46$ ms:  $F_{1,22} = 42.4$ ,  $p < .001$ ,  $\eta^2 = .6$ ). Correspondence type and trial type again interacted significantly ( $F_{2,44} = 25.3$ ,  $p < .001$ ,  $\eta^2 = .5$ ), such that congruent trial RTs were faster for the AE-VE ( $t_{22} = 6.0$ ,  $p < .001$ ) and AP-AE ( $t_{22} = 3.6$ ,  $p < .001$ ) correspondences, but not AP-VE, correspondences (Supplementary Figure S3).

#### *Congruency magnitudes: global analysis*

An RM-ANOVA of congruency magnitudes showed a main effect of attended feature ( $F_{2,44} = 25.2$ ,  $p < .001$ ,  $\eta^2 = .5$ ) and correspondence type ( $F_{2,44} = 13.2$ ,  $p < .001$ ,  $\eta^2 = .4$ ) and an interaction between the two ( $F_{4,88} = 39.5$ ,  $p < .001$ ,  $\eta^2 = .6$ ). *Post hoc* t tests showed that in the auditory elevation condition, the congruency magnitudes for all three correspondence types were significantly different from one another (AP-AE  $1.8 \pm .5$ , AP-VE  $-.2 \pm .4$ , AE-VE  $8.7 \pm 1.2$ : all  $t_{22} > -2.8$ , all  $p < \text{Bonferroni-corrected } .017$ ). In the auditory pitch condition, the AE-VE congruency magnitude ( $.5 \pm .4$ ) was significantly smaller than both the AP-AE ( $2.9 \pm .5$ :  $t_{23} = -4.9$ ,  $p < .001$ ) and AP-VE ( $2.4 \pm .5$ :  $t_{23} = 2.8$ ,  $p = .009$ ) congruency magnitudes, but the AP-AE and AP-VE congruency magnitudes did not differ ( $t_{23} = -1.1$ ,  $p = .3$ ). There was no significant difference between the congruency magnitudes for any correspondence type in the visual elevation condition.

#### *Accuracy compared with RTs*

There was a main effect of the attended feature on accuracy ( $F_{2,44} = 14.3$ ,  $p < .001$ ) – just as there was for response times (RTs). Visual elevation (VE) accuracy (mean  $\pm$  sem:  $97.4 \pm .7\%$ ) was significantly higher than both auditory pitch (AP:  $93.0 \pm 1.5\%$ ;  $t_{23} = -3.3$ ,  $p = .003$ ) and auditory elevation (AE:  $89.5 \pm 1.5\%$ ;  $t_{22} = -5.4$ ,  $p < .001$ ) accuracy. Mirroring this, VE RTs were significantly faster than both AE and AP RTs. But AP and AE accuracy did not differ ( $t_{22} = -2.1$ ,  $p = .052$ ; similarly, AP and AE RTs were not different either.

There was a main effect of trial type on accuracy ( $F_{1,44} = 35.3$ ,  $p < .001$ ) – as there was for RTs – in which accuracy was significantly higher on congruent ( $94.7 \pm .7\%$ ) compared to incongruent ( $91.9 \pm 1.1\%$ ) trials; reflecting this, congruent RTs were significantly faster.

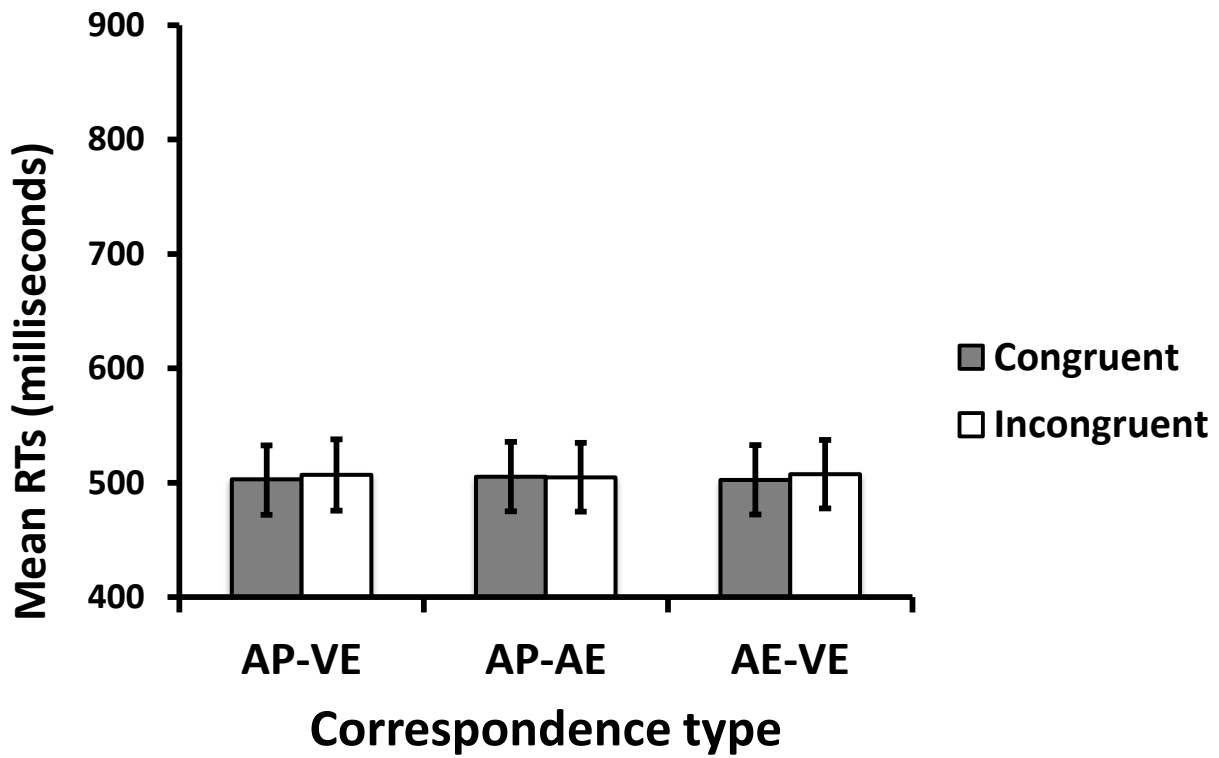
There was a significant interaction between attended feature and trial type for the accuracy data ( $F_{2,44} = 17.9$ ,  $p < .001$ ) – again, this was so for RTs as well. Accuracy was higher for congruent compared to incongruent trials for both AP ( $94.1 \pm 1.1\%$  vs  $91.9 \pm 1.7\%$ ;  $t_{23} = 3.1$ ,  $p = .005$ ) and AE ( $92.4 \pm 1.1\%$  vs  $86.6 \pm 2.0$ ;  $t_{22} = 6.0$ ,  $p < .001$ ); consistent with both AP and AE congruent trial RTs being faster compared to incongruent trials. However, accuracy for VE congruent and incongruent trials were not significantly different ( $97.4 \pm .6\%$  vs  $97.2 \pm .7\%$ ;  $t_{23} = .6$ ,  $p = .5$ ). Although VE RTs were significantly faster for congruent trials, the absolute difference was only 3ms, compared to 30-50ms for AP and AE. Thus, the essential nature of the interaction between attended feature and trial type was similar for both accuracy and RT data: larger congruent/incongruent differences for AP and AE compared to VE.

There was also a significant interaction between correspondence type and trial type for the accuracy data ( $F_{2,44} = 16.0$ ,  $p < .001$ ). This was also the case for RTs, but here the nature of the interaction was slightly different. For RTs, congruent RTs were faster than incongruent RTs for all correspondences, the absolute difference being largest for the AE-VE correspondence and smallest for the AP-VE correspondence. The gradient across correspondence types was similar for accuracy, however, accuracy was significantly higher on congruent vs. incongruent trials for the AE-VE ( $96.0 \pm .8\%$  vs.  $90.5 \pm 1.2\%$ ;  $t_{23} = 6.3$ ,  $p < .001$ ) and AP-AE ( $94.7 \pm .7\%$  vs  $91.8 \pm 1.2\%$ ;  $t_{23} = 3.4$ ,  $p = .002$ ) correspondences but not for the AP-VE correspondence ( $93.3 \pm .9\%$  vs  $93.4 \pm 1.0\%$ ;  $t_{23} = -.4$ ,  $p = .7$ ).

The main effect of correspondence type and its interaction with the attended feature were non-significant in the accuracy data, which was also the case in the RT analyses. There was a third order interaction of attended feature, correspondence type, and trial type, which was also true for RTs, but this was not analyzed further for accuracy data.

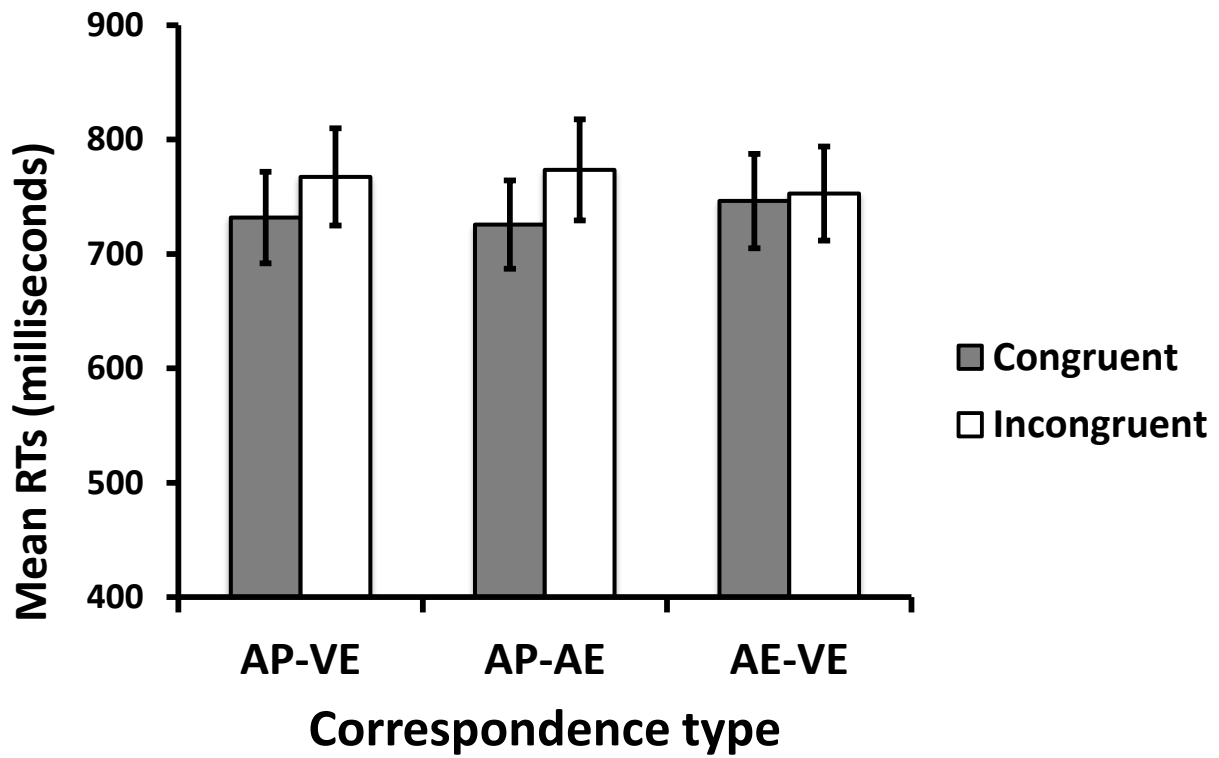
The close coupling between higher accuracy rates and faster RTs makes a speed/accuracy trade-off unlikely.

## Attended feature: visual elevation



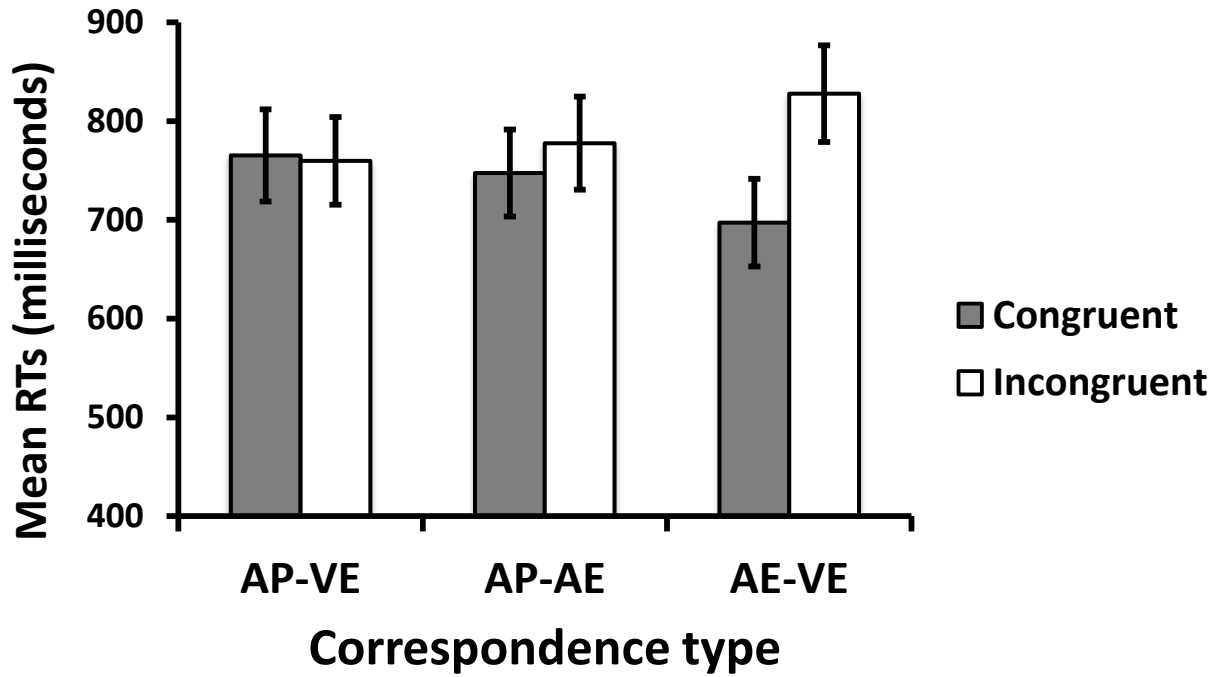
**Figure S1:** Mean RTs for each correspondence type when the attended feature was visual elevation (error bars = sem).

## Attended feature: auditory pitch



**Figure S2:** Mean RTs for each correspondence type when the attended feature was auditory pitch (error bars = sem).

## Attended feature: auditory elevation



**Figure S3:** Mean RTs for each correspondence type when the attended feature was auditory elevation (error bars = sem).