

# Does Language Influence How Visual Events Are Perceived?

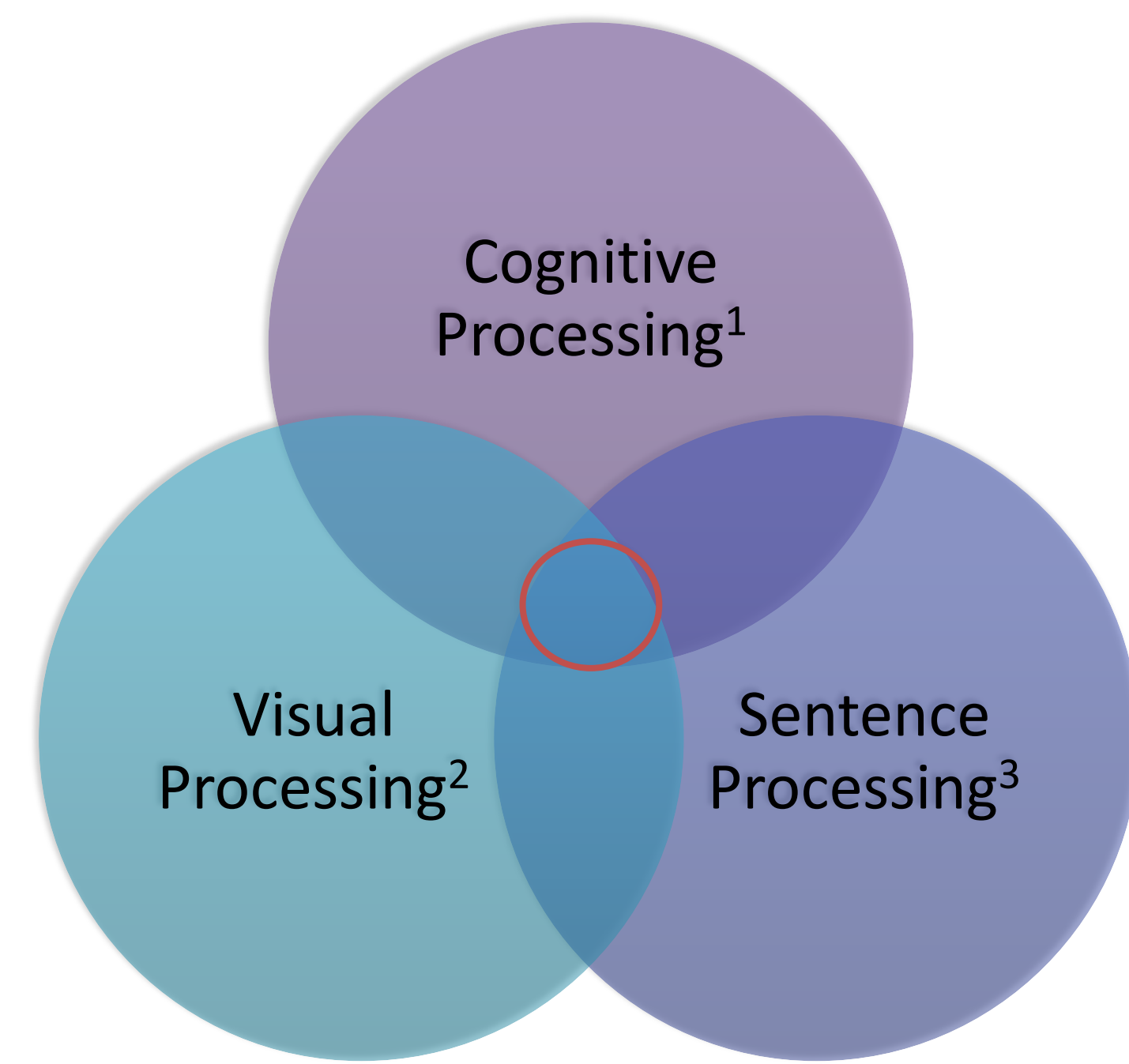
Gaurav Kharkwal<sup>1,2</sup>, Karin Stromswold<sup>1,2</sup>

kharkwal@rci.rutgers.edu karin@ruccs.rutgers.edu

1. Department of Psychology, Rutgers University – New Brunswick
2. Center for Cognitive Science, Rutgers University – New Brunswick



RUTGERS



Two experiments investigated how the linguistic and visual processing systems interact by asking subjects to “match” captions with visual events. Results suggest that the two systems work fairly independently and a “race” between the two determined how fast subjects performed the task.

## Experiment 1- Actives Only

**Task:** Does the caption match the video?

**Subjects:** 21 college-aged adults.

**Independent Variables:**

- Verb:
  - Chase } “Behind” ○ ⇒ □
  - Follow }
  - Flee } “Front” ○ ⇒ □
  - Lead }
- Eyes (acentric dots):
  - vs. ○
- Caption Veracity
  - True
  - False (NPs interchanged)

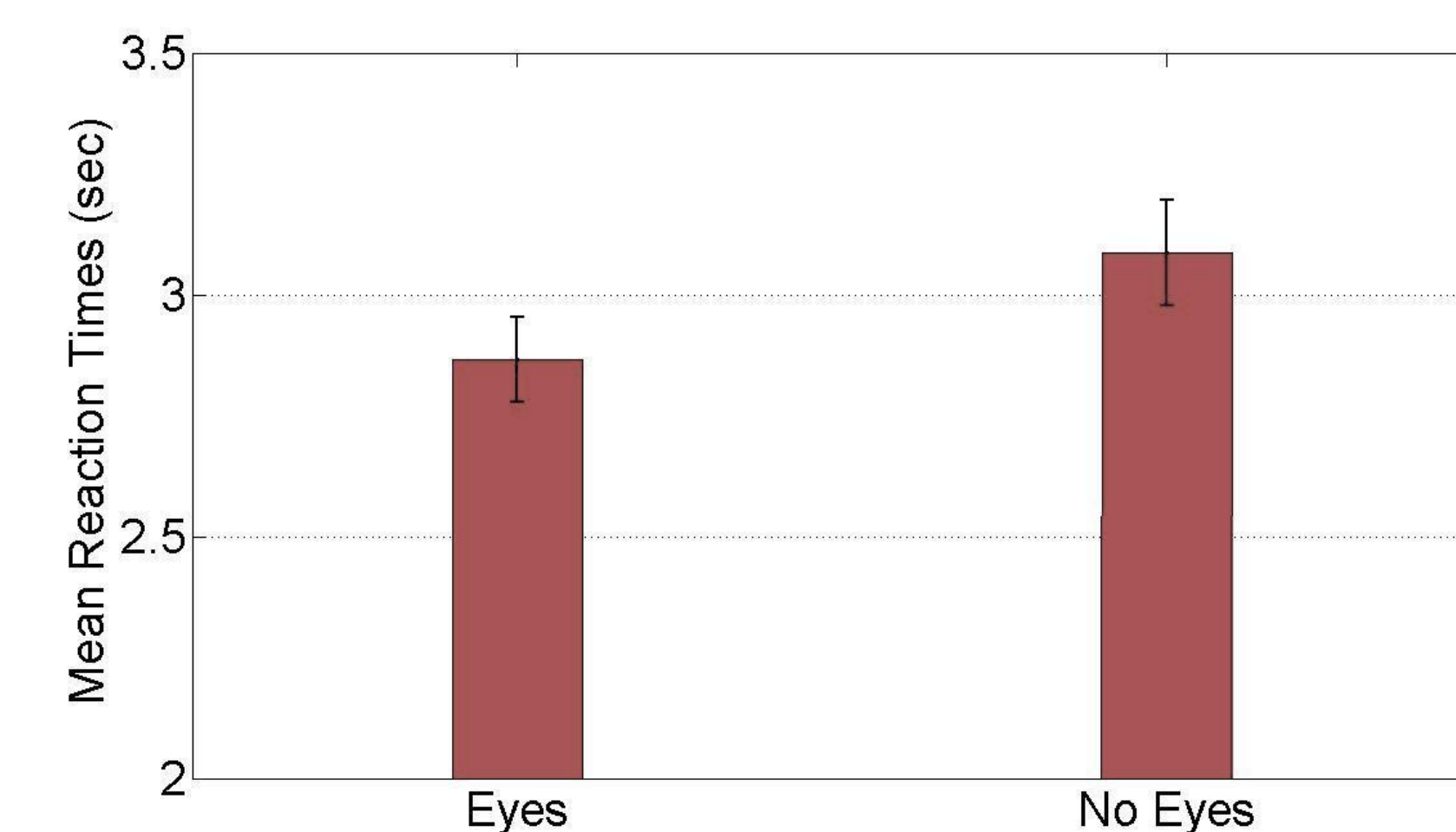
**Dependent Variables:**

- Reaction Time, [Error Rates]

## Results

- Mean accuracy: 94.47%
- RTs for correct trials only were analyzed.

- Main effect of Eyes ( $p < .01$ )



**Eyes** (Error bars = 95% CI.)

## Discussion

- Only Visual Parameter, Eyes, affected performance.
  - Eyes may indicate direction of motion and/or indicate which object is in front.
- *Hypothesis:* linguistic parsing is not necessary to complete the task.

## Acknowledgements

This work was supported by the National Science Foundation (BCS-0446850, BCS-0124095, & IGERT DGE-0549115). We thank Jacob Feldman, Sabrina Angelini, Tina Hou-Imerman, Nora Isacoff, Choonkyu Lee, Heather Yaden, and Nikhita Karki for their suggestions and advice.

## Experiment 2- Actives and Passives

**Task:** Same as Exp. 1.

**Subjects:** 20 college-aged adults.

**Independent Variables:**

- Voice:
  - Active (e.g. *The circle is chasing the square.*)
  - Passive (e.g. *The square is chased by the circle.*)
- Verb:
  - Trail } “Behind” ○ ⇒ □
  - Follow }
  - Guide } “Front” ○ ⇒ □
  - Lead }
- Eyes:
  - same as Exp. 1
- Caption Veracity:
  - same as Exp. 1

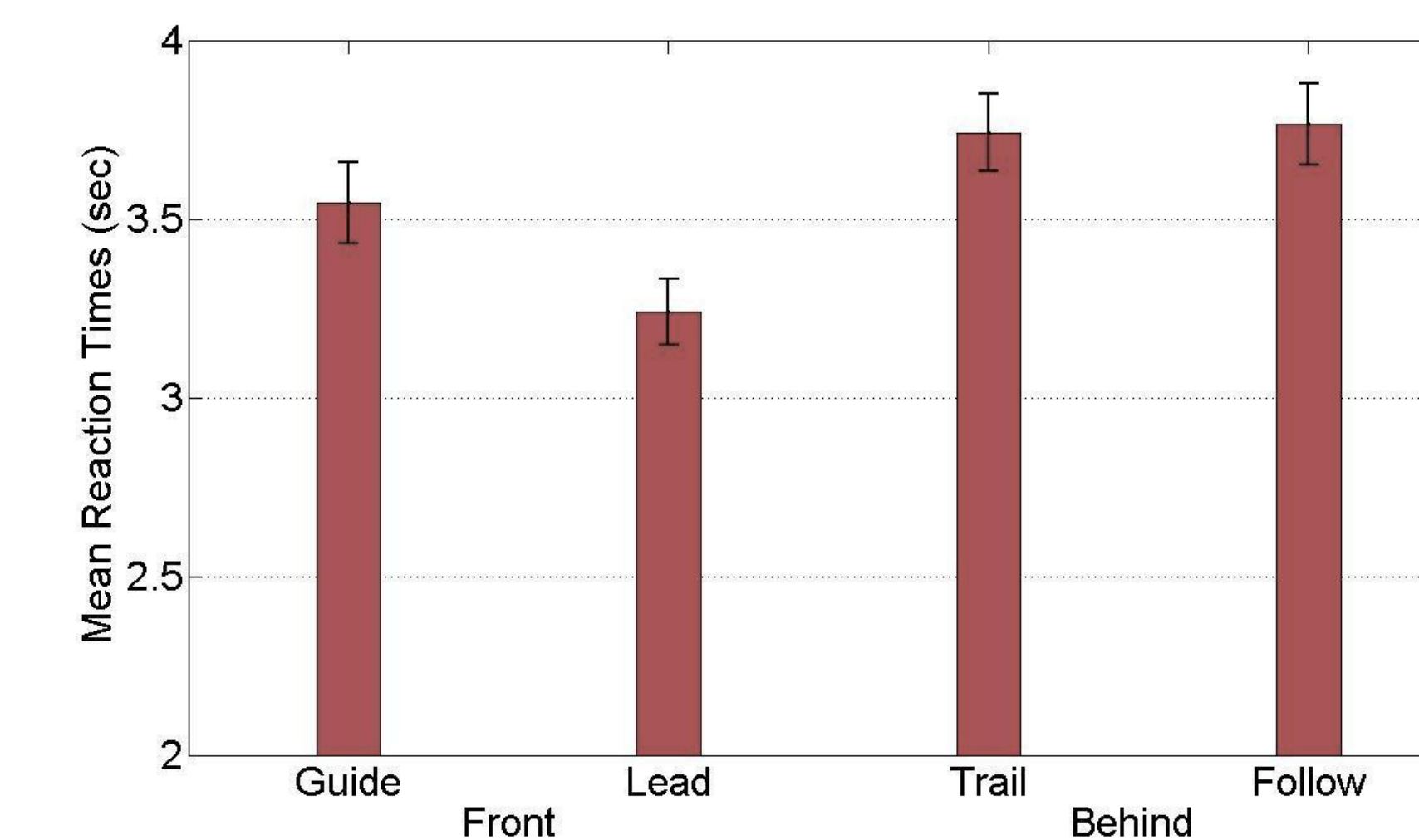
**Dependent Variables:**

- Reaction Time, [Error Rates]

## Results

- Mean accuracy: 95.55%
- RTs for correct trials only were analyzed

- Main effect of Verb ( $p < .001$ )

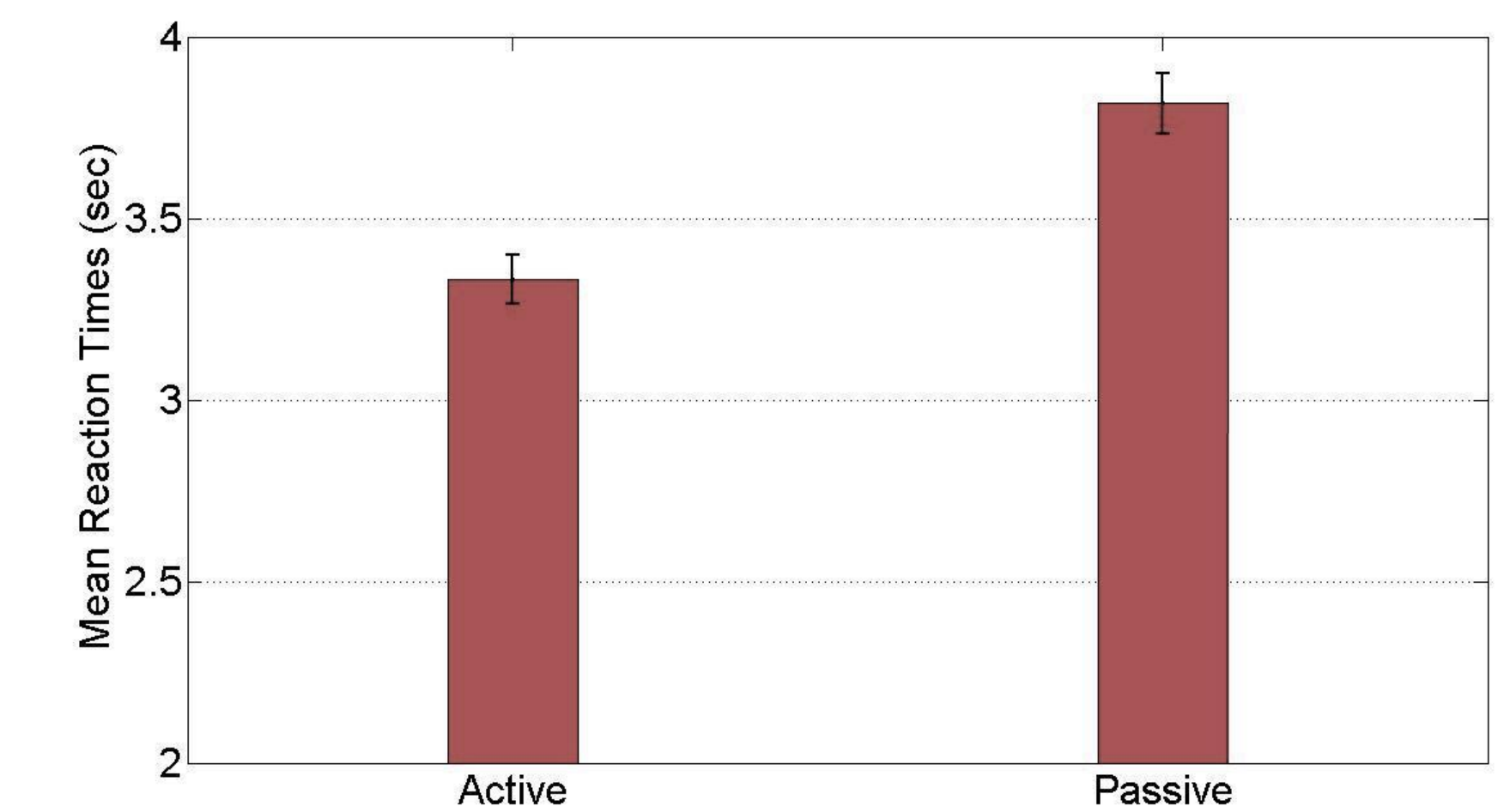


**Verb** (Error bars = 95% CI)

- Main effect of Verb “Perspective” ( $p < .001$ ):
  - “Behind” verbs ( $M = 3.74$  sec)
  - “Front” verbs ( $M = 3.38$  sec)

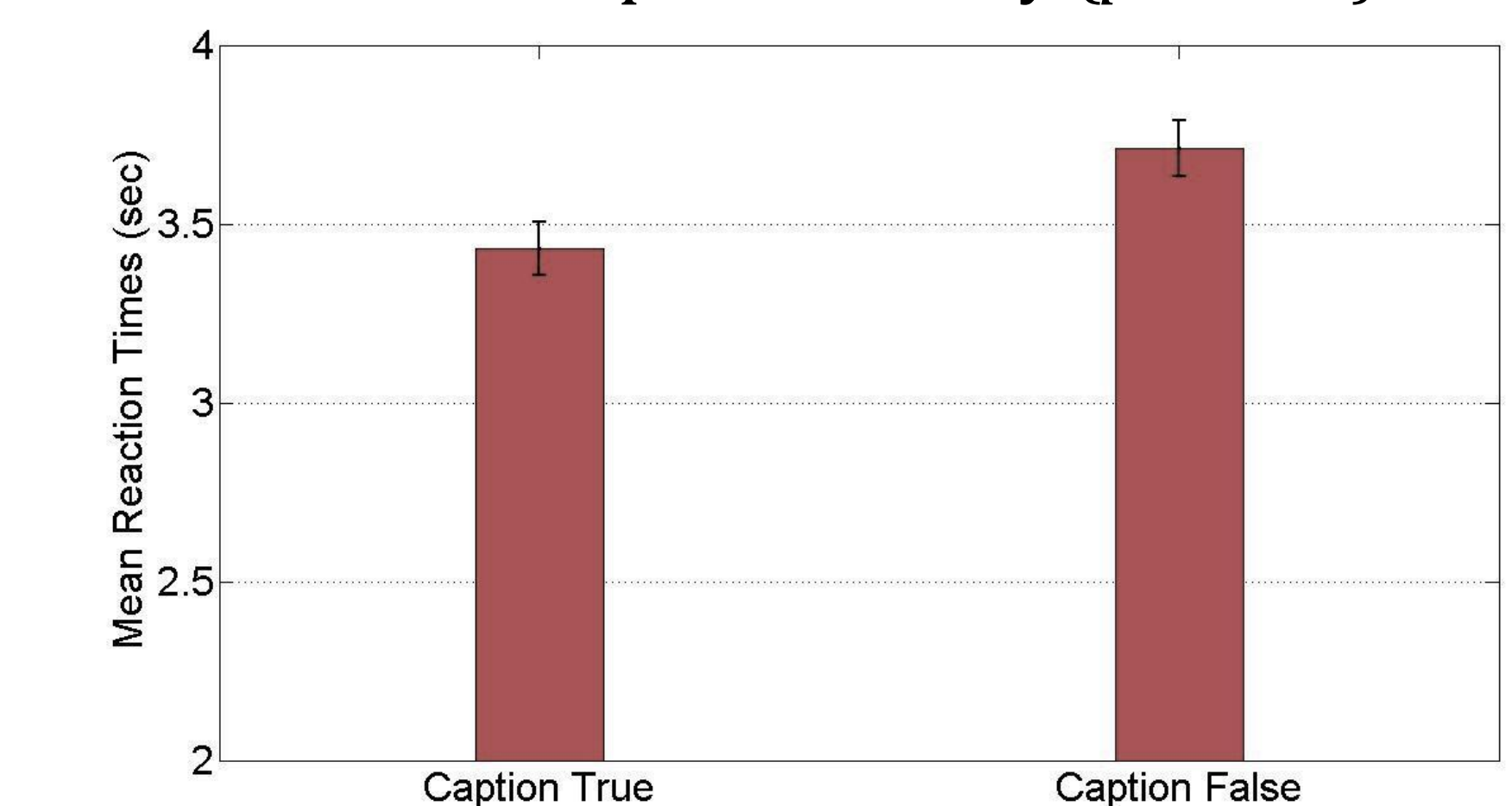
## Results (contd.)

- Main effect of Voice ( $p < .001$ )



**Voice** (Error bars = 95% CI)

- Main effect of Caption Veracity ( $p < .001$ )



**Caption Veracity** (Error bars = 95% CI)

## Discussion

- Visual Parameter (Eyes) did not affect performance.
- All Linguistic Parameters affected performance.

## Conclusion

- The linguistic and visual systems are fairly independent.
- *Hypothesis:* The two systems run in parallel (a “race”), and the slower influences the RTs.
- *Exp. 1:* The linguistic sub-task is easy -> only visual parameter played a role.
- *Exp. 2:* The linguistic sub-task is hard -> only linguistic parameters played a role.

## References

1. Zwaan, R. A. (2004). The immersed experienter: towards an embodied theory of language comprehension.
2. Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy.
3. Stromswold, K., Eisenband, J., Norland, E., Ratzan, J. (2002). Tracking the acquisition and processing of English passives: Using acoustic cues to disambiguate actives and passives.