Effects of integration in eye tracking

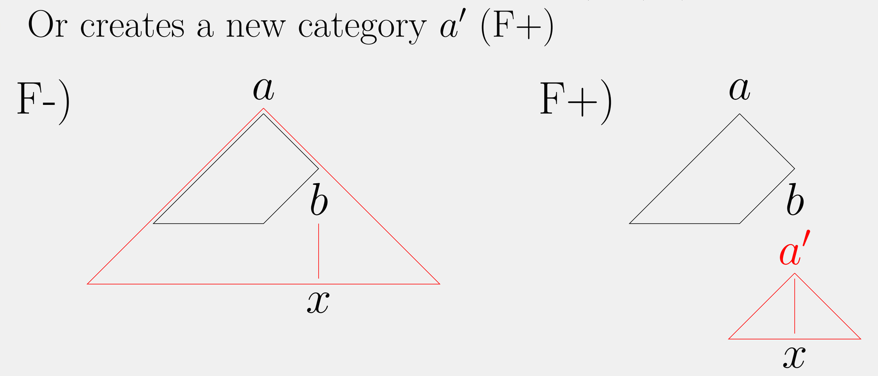
William Schuler and Marten van Schijndel
Linguistics, The Ohio State University
The Ohio State University

Introduction

This work investigates an apparent discrepancy between previous eye-tracking experiments some of which show a significant positive integration effect on first-pass fixation durations, and some of which show no positive effect for integration cost (and instead a significant negative effect). One reason for this discrepancy may be assumptions of serial vs parallel processing. Findings of positive integration calculate integration cost over the best parse, while findings of negative integration cost usually calculate integration cost weighted proportionally to the number of hypothesized parallel parses undergoing integration. Is positive integration cost a casualty of parallel processing models?

Integration

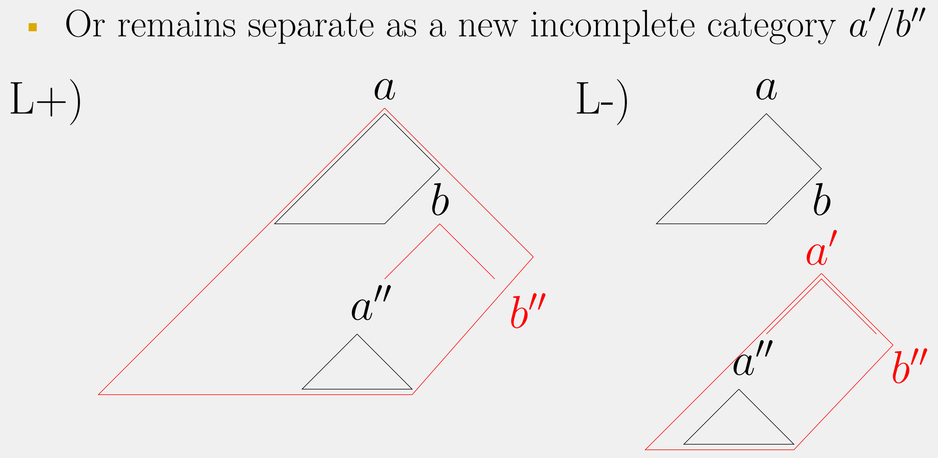
- Top-down models of processing assume readers predict upcoming observations to some degree
- Categorial grammars reflect this
  - Incomplete categories (trapezoids) with outstanding requirements
  - Complete categories (triangles) used to fulfill requirements
- Each new observation (x)...
  - Completes an incomplete category a/b (F-)
  - Or creates a new category a' (F+)



$$\frac{\text{NP/N} \quad \text{wind}}{\text{NP}} \; \text{N} \rightarrow \text{wind} \qquad \text{(F-)}$$

$$\frac{\text{T/S} \quad \text{the}}{\text{T/S} \quad \text{D}} \; \text{S} \stackrel{+}{\rightarrow} \text{D} \ldots ; \text{D} \rightarrow \text{the} \qquad \text{(F+)}$$

- The newly completed category (a'')...
  - Serves as the first piece of the outer requirement b (L+)
  - Or remains separate as a new incomplete category a'/b'' (L-)

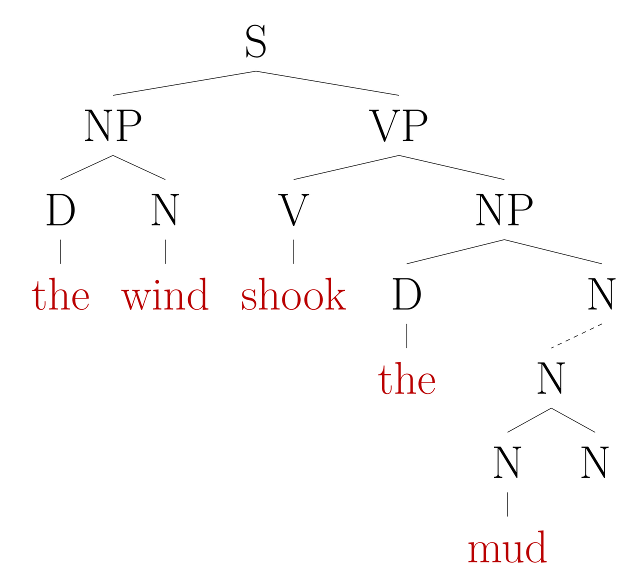


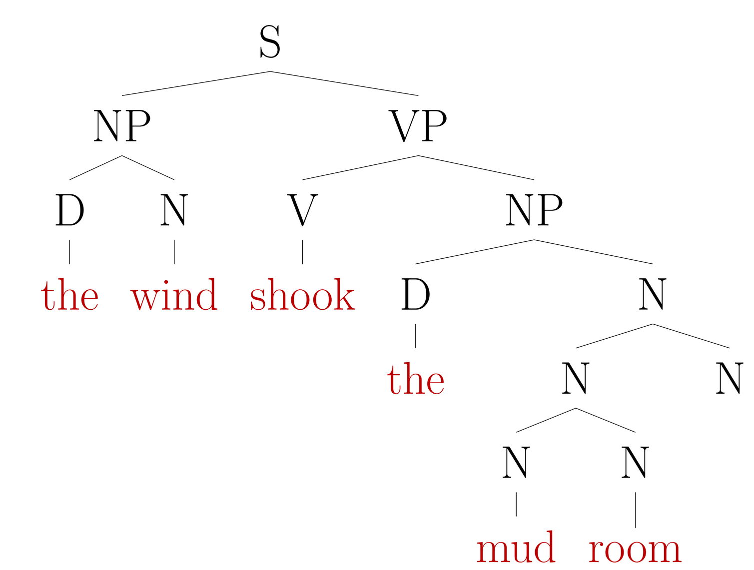
$$\frac{\text{S/VP} \quad \text{V}}{\text{S/NP}} \; \text{VP} \rightarrow \text{V NP} \qquad \text{(L+)}$$

$$\frac{\text{T/S} \quad \text{D}}{\text{T/S} \quad \text{NP/N}} \; \text{S} \stackrel{+}{\rightarrow} \text{NP} \ldots ; \text{NP} \rightarrow \text{D N} \qquad \text{(L-)}$$

- *Integration*: an outstanding dependency is non-trivially satisfied (occurs in operation L+)

Example





Integration in parsing 'the wind shook the mud room door.'

Integration Cost

Integration can predict reading time delays

- Termed a *positive* cost because it predicts delays
- Weighted by dependency length [Gibson, 2000]
- Definition of 'length' matters
  - English: Linear distance?
  - Korean: Structural distance? [Kwon et al., 2010]
  - Baumann (2012) adds structural predictors
    - Only *structural* distance improves model fit
- Studies assume serial processing
  - Is there an integration or not?
- Often based on constructed stimuli

Integration can predict *shorter* reading times [Demberg and Keller, 2008, Wu et al., 2010] [van Schijndel et al., 2013]

- Studies use additional factors
  - PCFG surprisal
  - Fixation histories
- These studies typically assume parallel processing
  - What proportion of hypothesized parses undergo integration?
- These studies all use large eye-tracking corpora
- Found on multiple corpora...
  - So difference likely not just due to data
  - Difference possibly due to additional factors
  - Difference possibly due to assumption of parallelism

Models

Model A

- Purpose: Is parallelism to blame?
- Same set of predictors as Baumann (2012)
  - Fixed: Word length
  - Fixed: Sentence position
  - Fixed: Unigram frequency
  - Fixed: Bigram frequency
  - Fixed: Joint interactions
  - Random: Subject/Item intercepts
  - Test: Parallel integration cost
- Result: Numerically positive integration cost ($.19\text{ms} \pm 1.2\text{ms}$, $p = .88$)
- Not quite a replication
  - Uses parallel (not 1-best) integration cost

Model B

- Purpose: Add by-subject random slopes
- Same set of predictors as Model A, plus:
  - Random: Word length
  - Random: Sentence position
  - Random: Unigram frequency
  - Test: Parallel integration cost
- Random: Bigram frequency wouldn't converge
- Result: Numerically negative integration cost ($-.20\text{ms} \pm 1.8\text{ms}$, $p = .91$)

Model C

- Purpose: Account for parafoveal processing
- Same set of predictors as Model B, plus:
  - Fixed: Was prev. fixation on prev. word?
  - Random: Was prev. fixation on prev. word?
  - Test: Parallel integration cost
- Result: Significant negative integration cost ($-3.4\text{ms} \pm 1.6\text{ms}$, $p < .05$)

Model D

- Purpose: Better account for frequency effects
- Same set of predictors as Model C, plus:
  - Fixed: PCFG surprisal
  - Random: PCFG surprisal
  - Test: Parallel integration cost
- Result: Strongly significant negative integration cost ($-6.3\text{ms} \pm 1.9\text{ms}$, $p < .001$)

Results and Conclusions

Results

- Loss of positive integration cost is not due to parallel processing model
- Negative integration cost arises from accounting for
  - PCFG surprisal
  - Parafoveal processing
  - Random slopes
- Predictors are independently motivated
- Predictors significantly increase model fit

Conclusion

- Results cast some doubt on existence of broad positive integration cost on reading times
- Highlights need to eliminate possible confounds in constructed stimuli (ala Bartek et al. 2011)
  - Perhaps confounds related to grammar rule probabilities

References

[Bartek et al., 2011] Bartek, B., Lewis, R. L., Vasishth, S., and Smith, M. (2011). In search of on-line locality effects in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 37(5):1178–1198.

[Baumann, 2012] Baumann, P. (2012). The role of hierarchical structure in syntactic dependency integration. In *Twenty-Fifth Annual CUNY Conference on Human Sentence Processing*, page 91.

[Demberg and Keller, 2008] Demberg, V. and Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, 109(2):193–210.

[Gibson, 2000] Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In *Image, language, brain: Papers from the first mind articulation project symposium*, pages 95–126, Cambridge, MA. MIT Press.

[Kennedy et al., 2003] Kennedy, A., Pynte, J., and Hill, R. (2003). The Dundee corpus. In *Proceedings of the 12th European conference on eye movement*.

[Kwon et al., 2010] Kwon, N., Lee, Y., Gordon, P. C., Kluender, R., and Polinsky, M. (2010). Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of pre-nominal relative clauses in korean. *Language*, 86(3):561.

[van Schijndel et al., 2013] van Schijndel, M., Nguyen, L., and Schuler, W. (2013). An analysis of memory-based processing costs using incremental deep syntactic dependency parsing. In *Proceedings of CMCL 2013*. Association for Computational Linguistics.

[Wu et al., 2010] Wu, S., Bachrach, A., Cardenas, C., and Schuler, W. (2010). Complexity metrics in an incremental right-corner parser. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL'10)*, pages 1189–1198.