Incremental Semantic Dependency Parsing

Marten van Schijndel William Schuler
Department of Linguistics
The Ohio State University

http://ling.osu.edu/∼vanschm

August 8, 2013

Introduction: Motivation

The person who officials say ___ stole millions . . .

Introduction: Motivation

The person who officials say ___ stole millions . . .

Goal: Incrementally obtain correct parse

- Filler-gap is hard for computers
  [Rimell et al., 2009, Nguyen et al., 2012]

Introduction: Motivation

The person who officials say ___ stole millions ...

Goal: Test human processing claims

- Filler-gap is hard for humans? [Chomsky and Miller, 1963]

Introduction: Motivation

The person who officials say ___ stole millions . . .

Goal: Test human processing claims

- Filler-gap is hard for humans [Gibson, 2000, Chen et al., 2005]

Introduction: Motivation

The person who officials say ___ stole millions . . .

Goal: Test human processing claims

- Filler-gap is hard for humans [Gibson, 2000, Chen et al., 2005]
- Embeddings speed processing [Pynte et al., 2008]
- Finishing embeddings = Fast
  [Wu et al., 2010, van Schijndel and Schuler, 2013]

Center embedding or filler-gap?

Overview

Contribution

Introduce an incremental semantic parser

- Fits reading times better than syntax parsing
- Replicate previous findings sans surface confounds

Overview

Contribution

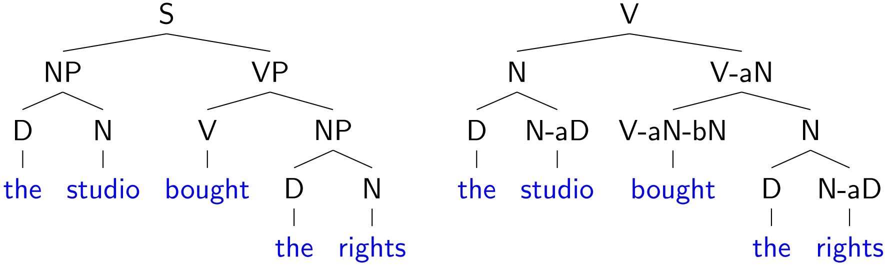
Introduce an incremental semantic parser

- Fits reading times better than syntax parsing
- Replicate previous findings sans surface confounds

1. Generalized Categorial Grammar
2. Incremental Semantic Parser
3. Eye-tracking evaluation
4. Results

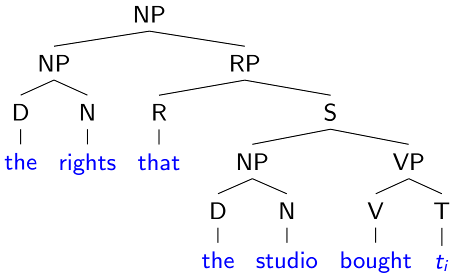
Generalized Categorial Grammar

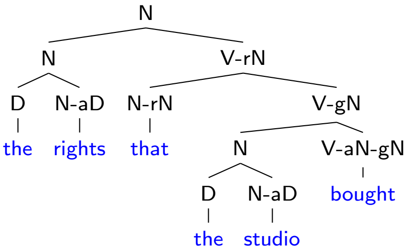
Reannotate WSJ [Nguyen et al., 2012]



Generalized Categorial Grammar

We can also keep the WSJ traces around.





Interpretation: Reannotation Rules

$$\frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d}{(f_{c\text{-}\mathbf{ad}}\; g\; h){:}\,c} \qquad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,(g\;k)\;h){:}\,c\psi} \qquad \frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\;g\,(h\;k)){:}\,c\psi} \qquad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,(g\;k)\,(h\;k)){:}\,c\psi}$$

$$\frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d}{(f_{c\text{-}\mathbf{bd}}\; g\; h){:}\,c} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\,(g\;k)\;h){:}\,c\psi} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\;g\,(h\;k)){:}\,c\psi} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\,(g\;k)\,(h\;k)){:}\,c\psi}$$

(Aa–h)

$$\frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c}{(f_{\mathrm{IM}}\; g\; h){:}c} \qquad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c}{\lambda_k\,(f_{\mathrm{IM}}\,(g\;k)\;h){:}c\psi} \qquad \frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c\psi}{\lambda_k\,(f_{\mathrm{IM}}\;g\,(h\;k)){:}c\psi} \qquad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c\psi}{\lambda_k\,(f_{\mathrm{IM}}\,(g\;k)\,(h\;k)){:}c\psi}$$

$$\frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d}{(f_{\mathrm{FM}}\; g\; h){:}c} \qquad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d}{\lambda_k\,(f_{\mathrm{FM}}\,(g\;k)\;h){:}c\psi} \qquad \frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{\mathrm{FM}}\;g\,(h\;k)){:}c\psi} \qquad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{\mathrm{FM}}\,(g\;k)\,(h\;k)){:}c\psi}$$

(Ma–h)

$$\frac{g{:}\,c\text{-}\mathbf{a}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,\{k\}\;g){:}\,c\text{-}\mathbf{g}d} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,\{k\}\;g){:}\,c\text{-}\mathbf{g}d} \qquad \frac{g{:}c}{\lambda_k\,(f_{\mathrm{IM}}\,\{k\}\;g){:}c\text{-}\mathbf{g}d}$$

(Ga–c)

$$\frac{g{:}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_i\,\exists_j\,(g\;i) \wedge (h\;i\;j){:}e} \qquad \frac{g{:}d\text{-}\mathbf{re} \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\,\exists_i\,(g\;k\;i) \wedge (h\;i\;j){:}\,c\text{-}\mathbf{re}} \qquad \frac{g{:}d\text{-}\mathbf{ie} \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\,\exists_i\,(g\;k\;i) \wedge (h\;i\;j){:}\,c\text{-}\mathbf{ie}}$$

(Fa–c)

$$\frac{g{:}e \quad h{:}c\text{-}\mathbf{r}d}{\lambda_i\,\exists_j\,(g\;i) \wedge (h\;i\;j){:}e} \qquad \text{(R)}$$

Interpretation: Reannotation Rules

$$\frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d}{(f_{c\text{-}\mathbf{ad}}\; g\; h){:}\,c} \quad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,(g\; k)\; h){:}\,c\psi} \quad \frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\; g\,(h\; k)){:}\,c\psi} \quad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{ad}}\,(g\; k)\,(h\; k)){:}\,c\psi}$$

$$\frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d}{(f_{c\text{-}\mathbf{bd}}\; g\; h){:}\,c} \quad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\,(g\; k)\; h){:}\,c\psi} \quad \frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\; g\,(h\; k)){:}\,c\psi} \quad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d\psi}{\lambda_k\,(f_{c\text{-}\mathbf{bd}}\,(g\; k)\,(h\; k)){:}\,c\psi}$$

(Aa–h)

$$\frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c}{(f_{\mathrm{IM}}\; g\; h){:}c} \quad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c}{\lambda_k\,(f_{\mathrm{IM}}\,(g\; k)\; h){:}c\psi} \quad \frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c\psi}{\lambda_k\,(f_{\mathrm{IM}}\; g\,(h\; k)){:}c\psi} \quad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c\psi}{\lambda_k\,(f_{\mathrm{IM}}\,(g\; k)\,(h\; k)){:}c\psi}$$

$$\frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d}{(f_{\mathrm{FM}}\; g\; h){:}c} \quad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d}{\lambda_k\,(f_{\mathrm{FM}}\,(g\; k)\; h){:}c\psi} \quad \frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{\mathrm{FM}}\; g\,(h\; k)){:}c\psi} \quad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\,(f_{\mathrm{FM}}\,(g\; k)\,(h\; k)){:}c\psi}$$

(Ma–h)

$$\frac{g{:}\,c\text{-}\mathbf{a}d}{\lambda_k\,(f_{c\text{-ad}}\,\{k\}\; g){:}\,c\text{-}\mathbf{g}d} \quad \frac{g{:}\,c\text{-}\mathbf{b}d}{\lambda_k\,(f_{c\text{-ad}}\,\{k\}\; g){:}\,c\text{-}\mathbf{g}d} \quad \frac{g{:}c}{\lambda_k\,(f_{\mathrm{IM}}\,\{k\}\; g){:}c\text{-}\mathbf{g}d}$$

(Ga–c)

$$\frac{g{:}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_i\,\exists_j\,(g\; i) \wedge (h\; i\; j){:}e} \quad \frac{g{:}d\text{-}\mathbf{r}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\,\exists_i\,(g\; k\; i) \wedge (h\; i\; j){:}\,c\text{-}\mathbf{r}e} \quad \frac{g{:}d\text{-}\mathbf{i}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\,\exists_i\,(g\; k\; i) \wedge (h\; i\; j){:}\,c\text{-}\mathbf{i}e}$$

(Fa–c)

$$\frac{g{:}e \quad h{:}c\text{-}\mathbf{r}d}{\lambda_i\,\exists_j\,(g\; i) \wedge (h\; i\; j){:}e}$$

(R)

Interpretation: Reannotation Rules

$$\frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d}{(f_{c\text{-}\mathbf{ad}}\; g\; h){:}\,c} \qquad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\, (g\; k)\; h){:}\,c\psi} \qquad \frac{g{:}d \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\; g\, (h\; k)){:}\,c\psi} \qquad \frac{g{:}d\psi \quad h{:}\,c\text{-}\mathbf{a}d\psi}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\, (g\; k)\, (h\; k)){:}\,c\psi}$$

$$\frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d}{(f_{c\text{-}\mathbf{bd}}\; g\; h){:}\,c} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d}{\lambda_k\, (f_{c\text{-}\mathbf{bd}}\, (g\; k)\; h){:}\,c\psi} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d \quad h{:}d\psi}{\lambda_k\, (f_{c\text{-}\mathbf{bd}}\; g\, (h\; k)){:}\,c\psi} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d\psi \quad h{:}d\psi}{\lambda_k\, (f_{c\text{-}\mathbf{bd}}\, (g\; k)\, (h\; k)){:}\,c\psi}$$

(Aa–h)

$$\frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c}{(f_{\mathrm{IM}}\; g\; h){:}c} \qquad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c}{\lambda_k\, (f_{\mathrm{IM}}\, (g\; k)\; h){:}c\psi} \qquad \frac{g{:}\,u\text{-}\mathbf{a}d \quad h{:}c\psi}{\lambda_k\, (f_{\mathrm{IM}}\; g\, (h\; k)){:}c\psi} \qquad \frac{g{:}\,u\text{-}\mathbf{a}d\psi \quad h{:}c\psi}{\lambda_k\, (f_{\mathrm{IM}}\, (g\; k)\, (h\; k)){:}c\psi}$$

$$\frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d}{(f_{\mathrm{FM}}\; g\; h){:}c} \qquad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d}{\lambda_k\, (f_{\mathrm{FM}}\, (g\; k)\; h){:}c\psi} \qquad \frac{g{:}c \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\, (f_{\mathrm{FM}}\; g\, (h\; k)){:}c\psi} \qquad \frac{g{:}c\psi \quad h{:}\,u\text{-}\mathbf{a}d\psi}{\lambda_k\, (f_{\mathrm{FM}}\, (g\; k)\, (h\; k)){:}c\psi}$$

(Ma–h)

$$\frac{g{:}\,c\text{-}\mathbf{a}d}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\, \{k\}\; g){:}\,c\text{-}\mathbf{g}d} \qquad \frac{g{:}\,c\text{-}\mathbf{b}d}{\lambda_k\, (f_{c\text{-}\mathbf{ad}}\, \{k\}\; g){:}\,c\text{-}\mathbf{g}d} \qquad \frac{g{:}c}{\lambda_k\, (f_{\mathrm{IM}}\, \{k\}\; g){:}c\text{-}\mathbf{g}d}$$

(Ga–c)

$$\frac{g{:}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_i\, \exists_j\, (g\; i) \wedge (h\; i\; j){:}e} \qquad \frac{g{:}d\text{-}\mathbf{r}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\, \exists_i\, (g\; k\; i) \wedge (h\; i\; j){:}\,c\text{-}\mathbf{r}e} \qquad \frac{g{:}d\text{-}\mathbf{i}e \quad h{:}\,c\text{-}\mathbf{g}d}{\lambda_{k\,j}\, \exists_i\, (g\; k\; i) \wedge (h\; i\; j){:}\,c\text{-}\mathbf{i}e}$$

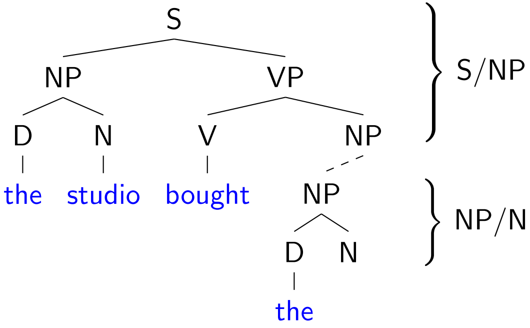
(Fa–c)

$$\frac{g{:}e \quad h{:}c\text{-}\mathbf{r}d}{\lambda_i\, \exists_j\, (g\; i) \wedge (h\; i\; j){:}e}$$

(R)

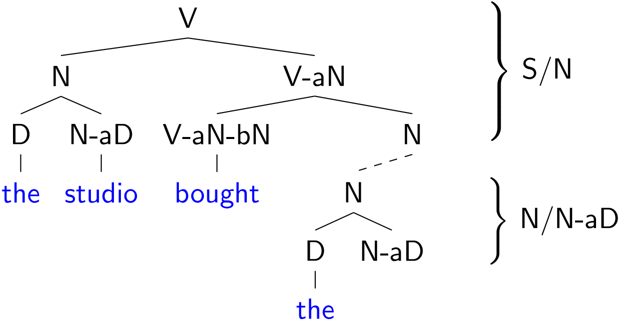
Interpretation

Connected Components

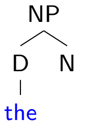


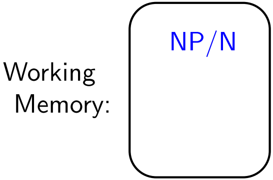
Interpretation

Reannotated Connected Components

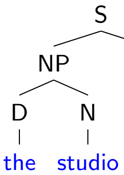


Connected Component Parsing

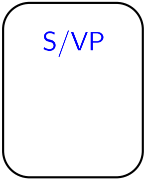




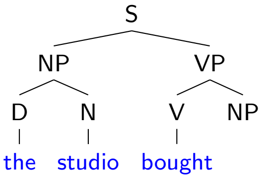
Connected Component Parsing

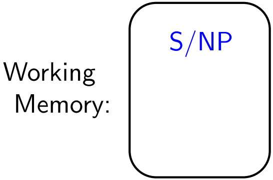


Working Memory:

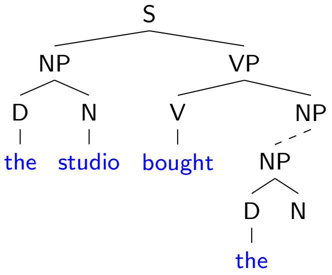


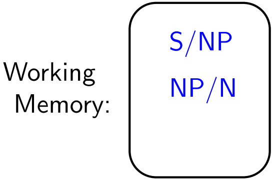
Connected Component Parsing



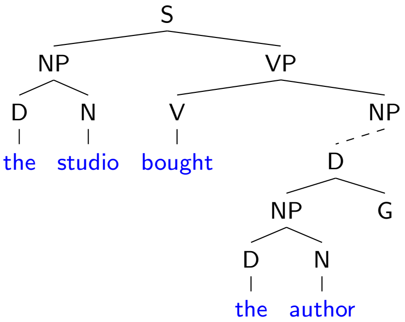


Connected Component Parsing





Connected Component Parsing

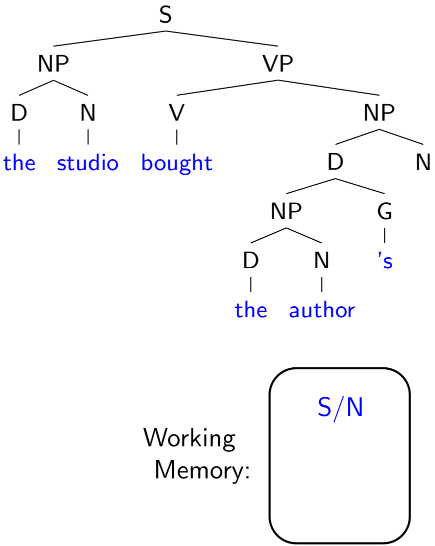


Working Memory:

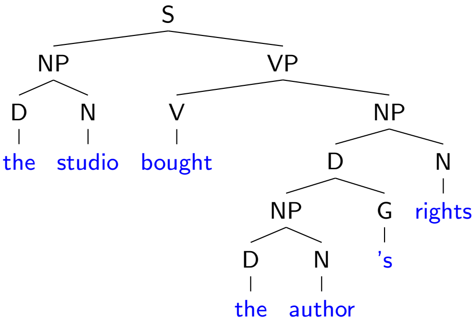
S/NP

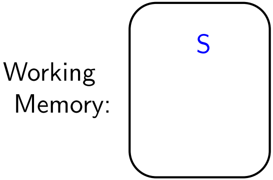
D/G

Connected Component Parsing

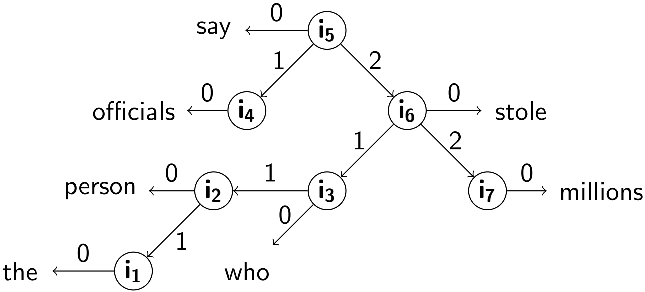


Connected Component Parsing





Interpretation: Referent States



Interpretation: Fa/La

First or Last element of a CC

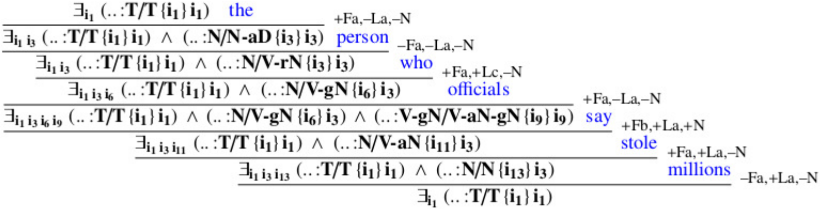
$$\frac{\exists_{i^1 j^1 .. i^\ell j^\ell} \ \dots \ \wedge \ (g^\ell{:}c/d\,\{j^\ell\}\ i^\ell) \quad x_t}{\exists_{i^1 j^1 .. i^\ell} \ \dots \ \wedge \ ((g^\ell f){:}c\ i^\ell)} \ x_t \mapsto_M f{:}d \tag{–Fa}$$

$$\frac{\exists_{i^1 j^1 .. i^\ell j^\ell} \ \dots \ \wedge \ (g^\ell{:}c/d\,\{j^\ell\}\ i^\ell) \quad x_t}{\exists_{i^1 j^1 .. i^\ell j^\ell i^{\ell+1}} \ \dots \ \wedge \ (g^\ell{:}c/d\,\{j^\ell\}\ i^\ell) \ \wedge \ (f{:}e\ i^{\ell+1})} \ x_t \mapsto_M f{:}e \tag{+Fa}$$

$$\frac{\exists_{i^1 j^1 .. i^{\ell-1} j^{\ell-1} i^\ell} \ \dots \ \wedge \ (g^\ell{:}d\ i^\ell)}{\exists_{i^1 j^1 .. i^\ell j^\ell} \ \dots \ \wedge \ ((f g^\ell){:}c/e\,\{j^\ell\}\ i^\ell)} \begin{cases} g{:}d\ \ h{:}e \Rightarrow (f\ g\ h){:}c \\ g{:}d\ \ h{:}e \Rightarrow \lambda_k(f\ (g\ k)\ h){:}c \\ g{:}d\ \ h{:}e \Rightarrow \lambda_k(f\ g\ (h\ k)){:}c \\ g{:}d\ \ h{:}e \Rightarrow \lambda_k(f\ (g\ k)\ (h\ k)){:}c \end{cases} \tag{–La}$$

$$\frac{\exists_{i^1 j^1 .. i^{\ell-1} j^{\ell-1} i^\ell} \ \dots \ \wedge \ (g^{\ell-1}{:}a/c\,\{j^{\ell-1}\}\ i^{\ell-1}) \ \wedge \ (g^\ell{:}d\ i^\ell)}{\exists_{i^1 j^1 .. i^{\ell-1} j^{\ell-1}} \ \dots \ \wedge \ (g^{\ell-1} \circ (f\ g^\ell){:}a/e\,\{j^{\ell-1}\}\ i^{\ell-1})} \begin{cases} g{:}d\ \ h{:}e \Rightarrow (f\ g\ h){:}c \\ g{:}d\ \ h{:}e \Rightarrow \lambda_k(f\ (g\ k)\ h){:}c \\ g{:}d\ \ h{:}e \Rightarrow \lambda_k(f\ g\ (h\ k)){:}c \\ g{:}d\ \ h{:}e \Rightarrow \lambda_k(f\ (g\ k)\ (h\ k)){:}c \end{cases} \tag{+La}$$

Interpretation



Interpretation: Fb/Lb

$$\psi \in \{\text{-r,-i}\} \times \text{C}$$

$$\frac{\exists_{i^1 j^1 .. \; i^n j^n .. \; i^\ell j^\ell} \ldots \wedge (g^n{:}y/z\psi \, \{j^n\} \, i^n) \wedge \ldots \wedge (g^\ell{:}c/d \, \{j^\ell\} \, i^\ell) \quad x_t}{\exists_{i^1 j^1 .. \; i^n j^n .. \; i^\ell} \ldots \wedge (g^n{:}y/z\psi \, \{j^n\} \, i^n) \wedge \ldots \wedge ((g^\ell (f'\{j^n\} f)){:}c \, i^\ell)}$$

$$x_t \mapsto_M \lambda_k (f' \{k\} \, f){:}d \quad (\text{–Fb})$$

$$\frac{\exists_{i^1 j^1 .. \; i^n j^n .. \; i^\ell j^\ell} \ldots \wedge (g^n{:}y/z\psi \, \{j^n\} \, i^n) \wedge \ldots \wedge (g^\ell{:}c/d \, \{j^\ell\} \, i^\ell) \quad x_t}{\exists_{i^1 j^1 .. \; i^n j^n .. \; i^\ell j^\ell i^{\ell+1}} \ldots \wedge (g^n{:}y/z\psi \, \{j^n\} \, i^n) \wedge \ldots \wedge (g^\ell{:}c/d \, \{j^\ell\} \, i^\ell) \wedge ((f'\{j^n\} f){:}e \, i^{\ell+1})}$$

$$x_t \mapsto_M \lambda_k (f' \{k\} \, f){:}e \quad (+\text{Fb})$$

$$\frac{\exists_{i^1 j^1 .. \; i^n j^n .. \; i^{\ell-1} j^{\ell-1} i^\ell} \ldots \wedge (g^n{:}y/z\psi \, \{j^n\} \, i^n) \wedge \ldots \wedge (g^\ell{:}d \, i^\ell)}{\exists_{i^1 j^1 .. \; i^n j^n .. \; i^\ell j^\ell} \ldots \wedge (g^n{:}y/z\psi \, \{j^n\} \, i^n) \wedge \ldots \wedge ((fg^\ell) \circ (f'\{j^n\}){:}c\psi/e \, \{j^\ell\} \, i^\ell)}$$

$$g{:}d \;\; h{:}e \Rightarrow \lambda_k (fg \, (f' \{k\} \, h)){:}c\psi \quad (\text{–Lb})$$

$$\frac{\exists_{i^1 j^1 .. \; i^n j^n .. \; i^{\ell-1} j^{\ell-1} i^\ell} \ldots \wedge (g^n{:}y/z\psi \, \{j^n\} \, i^n) \wedge \ldots \wedge (g^{\ell-1}{:}a/c\psi \, \{j^{\ell-1}\} \, i^{\ell-1}) \wedge (g^\ell{:}d \, i^\ell)}{\exists_{i^1 j^1 .. \; i^n j^n .. \; i^{\ell-1} j^{\ell-1}} \ldots \wedge (g^n{:}y/z\psi \, \{j^n\} \, i^n) \wedge \ldots \wedge (g^{\ell-1} \circ (fg^\ell) \circ (f'\{j^n\}){:}a/e \, \{j^{\ell-1}\} \, i^{\ell-1})}$$

$$g{:}d \;\; h{:}e \Rightarrow \lambda_k (fg \, (f' \{k\} \, h)){:}c\psi \quad (+\text{Lb})$$

Interpretation: Lc/N

$$\frac{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1} i^{\ell}} \;\ldots\; \wedge\; (g^{\ell}{:}d\; i^{\ell})}{\exists_{i^1 j^1 .. \, i^{\ell} j^{\ell}} \;\ldots\; \wedge\; ((fg^{\ell}) \circ (\lambda_{h\,k\,i}\,(h\,k)){:}a/e\psi\;\{j^{\ell}\}\; i^{\ell})} \quad g{:}d \;\; h{:}e\psi \Rightarrow (fg\; h){:}c \qquad (-\text{Lc})$$

$$\frac{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1} i^{\ell}} \;\ldots\; \wedge\; (g^{\ell-1}{:}a/c\;\{j^{\ell-1}\}\; i^{\ell-1}) \;\wedge\; (g^{\ell}{:}d\; i^{\ell})}{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1}} \;\ldots\; \wedge\; (g^{\ell-1} \circ (fg^{\ell}) \circ (\lambda_{h\,k\,i}\,(h\,k)){:}a/e\psi\;\{j^{\ell-1}\}\; i^{\ell-1})} \quad g{:}d \;\; h{:}e\psi \Rightarrow (fg\; h){:}c \qquad (+\text{Lc})$$

$$\frac{\exists_{i^1 j^1 .. \, i^{\ell} j^{\ell}} \;\ldots\; \wedge\; (g^{\ell-1}{:}c/d\psi\;\{j^{\ell-1}\}\; i^{\ell-1}) \;\wedge\; (g^{\ell}{:}d\psi/e\;\{j^{\ell}\}\; i^{\ell})}{\exists_{i^1 j^1 .. \, i^{\ell-1} j^{\ell-1}} \;\ldots\; \wedge\; (g^{\ell-1} \circ (\lambda_{h\,i} \exists_j (h\,j)) \circ g^{\ell}{:}c/e\;\{j^{\ell-1}\}\; i^{\ell-1})} \qquad (+\text{N})$$

All of these rules may be made probabilistic

Interpretation

$$\exists_{i_1}\ (..:\mathbf{T/T}\ \{\mathbf{i_1}\}\ \mathbf{i_1}) \quad \text{the} \qquad {+}\text{Fa},{-}\text{La},{-}\text{N}$$

$$\exists_{i_1\, i_3}\ (..:\mathbf{T/T}\ \{\mathbf{i_1}\}\ \mathbf{i_1}) \wedge (..:\mathbf{N/N\text{-}aD}\ \{\mathbf{i_3}\}\ \mathbf{i_3}) \quad \text{person} \qquad {-}\text{Fa},{-}\text{La},{-}\text{N}$$

$$\exists_{i_1\, i_3}\ (..:\mathbf{T/T}\ \{\mathbf{i_1}\}\ \mathbf{i_1}) \wedge (..:\mathbf{N/V\text{-}rN}\ \{\mathbf{i_3}\}\ \mathbf{i_3}) \quad \text{who} \qquad {+}\text{Fa},{+}\text{Lc},{-}\text{N}$$

$$\exists_{i_1\, i_3\, i_6}\ (..:\mathbf{T/T}\ \{\mathbf{i_1}\}\ \mathbf{i_1}) \wedge (..:\mathbf{N/V\text{-}gN}\ \{\mathbf{i_6}\}\ \mathbf{i_3}) \quad \text{officials} \qquad {+}\text{Fa},{-}\text{La},{-}\text{N}$$

$$\exists_{i_1\, i_3\, i_6\, i_9}\ (..:\mathbf{T/T}\ \{\mathbf{i_1}\}\ \mathbf{i_1}) \wedge (..:\mathbf{N/V\text{-}gN}\ \{\mathbf{i_6}\}\ \mathbf{i_3}) \wedge (..:\mathbf{V\text{-}gN/V\text{-}aN\text{-}gN}\ \{\mathbf{i_9}\}\ \mathbf{i_9}) \quad \text{say} \qquad {+}\text{Fb},{+}\text{La},{+}\text{N}$$

$$\exists_{i_1\, i_3\, i_{11}}\ (..:\mathbf{T/T}\ \{\mathbf{i_1}\}\ \mathbf{i_1}) \wedge (..:\mathbf{N/V\text{-}aN}\ \{\mathbf{i_{11}}\}\ \mathbf{i_3}) \quad \text{stole} \qquad {+}\text{Fa},{+}\text{La},{-}\text{N}$$

$$\exists_{i_1\, i_3\, i_{13}}\ (..:\mathbf{T/T}\ \{\mathbf{i_1}\}\ \mathbf{i_1}) \wedge (..:\mathbf{N/N}\ \{\mathbf{i_{13}}\}\ \mathbf{i_3}) \quad \text{millions} \qquad {-}\text{Fa},{+}\text{La},{-}\text{N}$$

$$\exists_{i_1}\ (..:\mathbf{T/T}\ \{\mathbf{i_1}\}\ \mathbf{i_1})$$

Evaluation: Syntactic vs Semantic

Syntactic Parser [van Schijndel et al., 2013]

- Only Fa/La
- Trained on WSJ 02-21
- Split-merged ×5 [Petrov et al., 2006]

Semantic Parser

- Trained on Reannotated WSJ 02-21
- Split-merged ×3

Evaluation: Syntactic vs Semantic

Test Corpus: Dundee

- Log-transformed go-past durations
- Omit:
  - first and last of each line (wrap-up)
  - $<$ 5 times in WSJ (accuracy) [Fossum and Levy, 2012]
  - saccade length $>$ 4 (track loss) [Demberg and Keller, 2008]

Eye Tracking

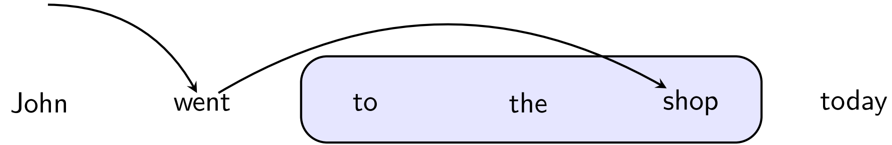
Go-past durations:

Cumulative factors are summed over the go-past region

Non-cumulative factors are based on the initial word in a region (shop)

Eye Tracking

Go-past durations:



X = Go-past region

Cumulative factors are summed over the go-past region

Non-cumulative factors are based on the initial word in a region (shop)

Evaluation: Syntactic vs Semantic

Fitting a linear mixed effects model (*lmer* in R)

Fixed Effects

- Word length
- Sentence position
- Prev, Next word fixated?
- Unigram and bigram probs
- Surprisal [Hale, 2001]
- Region length
- Cum. surprisal
- Cum. entropy reduction [Hale, 2003]
- Joint interactions
- Spillover predictors

By-subject random slopes

- Region length
- Prev word fixated?
- Cumulative surprisal

Subject and Item random intercepts

Evaluation: Syntactic vs Semantic

Model	log-likelihood	AIC
syntactic	-64175	128619
semantic	-64169	128609

Goodness-of-fits

Relative likelihood: 0.0009 (n = 151,331)

Evaluation: Semantic Factors

Factor	coeff	std. err.	t-score	p-value
F+L– (encoding)	0.014	0.005	2.665	0.02
F–L+ (integration)	-0.021	0.005	-4.109	0.001
F–L+|N+	-0.021	0.005	-4.109	?
F–L–	–	–	–	.50
F+L+	–	–	–	–

Significance of residualized factors on reading time.

Positive t-score: inhibition

Negative t-score: facilitation

Evaluation: Semantic Factors

Corpus: Reannotated WSJ

- Remove sentences with modifier embeddings [Pynte et al., 2008]

For example:
The CEO sold [[the shares] of the company]

Evaluation: Semantic Factors

Model	coeff	std err	t-score
Canonical	-0.040	0.010	-4.05
Other	-0.017	0.004	-4.20

Significance of residualized factors on reading time.
Positive t-score: inhibition
Negative t-score: facilitation

To achieve convergence, residualization was used

Conclusion

Results

- Described incremental semantic dependency parser
- General metrics are not hurt by semantic calculation
- Semantic metrics predict reading times better than syntactic
- Replicated negative integration cost without FG confound
- Failed to find support for maintenance cost

Fin

Thanks to Elliot Schumacher (and viewers like you)!
Questions?

Extras 1: Probabilistic Formulae

$$\mathsf{P}_{\phi_\ell}(\text{`}-\text{'}\ r^F \mid \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \mathsf{E}_{\gamma^*_\ell}(c \overset{0}{\rightarrow} d\ ...) \cdot \textstyle\sum_x \mathsf{P}_\gamma(d \rightarrow x) \cdot [\![ r^F = \langle i, \text{`id'}, j\rangle]\!] \quad \text{(1a)}$$

$$\mathsf{P}_{\phi_\ell}(\text{`}+\text{'}\ r^F \mid \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \mathsf{E}_{\gamma^*_\ell}(c \overset{+}{\rightarrow} d\ ...) \cdot \textstyle\sum_x \mathsf{P}_\gamma(d \rightarrow x) \cdot [\![ r^F = \langle \text{`}-\text{'}, \text{`}-\text{'}, \text{`}-\text{'}\rangle]\!] \quad \text{(1b)}$$

$$\mathsf{P}_{\lambda_\ell}(\text{`}+\text{'} \mid \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \textstyle\sum_{c',e} \mathsf{E}_{\gamma^*_\ell}(c \overset{0}{\rightarrow} c'\ ...) \cdot \mathsf{P}_{\gamma_{B,\ell}}(c' \rightarrow d\ \ e) \quad \text{(2a)}$$

$$\mathsf{P}_{\lambda_\ell}(\text{`}-\text{'} \mid \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \textstyle\sum_{c',e} \mathsf{E}_{\gamma^*_\ell}(c \overset{+}{\rightarrow} c'\ ...) \cdot \mathsf{P}_{\gamma_{A,\ell}}(c' \rightarrow d\ \ e) \quad \text{(2b)}$$

$$\mathsf{P}_{\nu_\ell}(\text{`}+\text{'} \mid \langle i,c\rangle\ \langle j,d\rangle\ \langle j',d'\rangle\ \langle k,e\rangle) \overset{\text{def}}{=} [\![c,d,d' \in C \times \{\textbf{-g}\} \times C \ \wedge \ e \notin C \times \{\textbf{-g}\} \times C]\!] \quad \text{(3a)}$$

$$\mathsf{P}_{\nu_\ell}(\text{`}-\text{'} \mid \langle i,c\rangle\ \langle j,d\rangle\ \langle j',d'\rangle\ \langle k,e\rangle) \overset{\text{def}}{=} [\![c,d,d' \notin C \times \{\textbf{-g}\} \times C \ \vee \ e \in C \times \{\textbf{-g}\} \times C]\!] \quad \text{(3b)}$$

Extras 1: Probabilistic Formulae

$$\mathsf{P}_{\alpha_\ell}(\langle i',c'\rangle\ r^A \mid l\ \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \begin{cases}\text{if } l = \text{`+'} : \sum_e \mathsf{E}_{\gamma^*_\ell}(c \overset{+}{\rightarrow} c'\ \ldots) \cdot \mathsf{P}_{\gamma_{A,\ell}}(c' \rightarrow d\ e)\\ \text{if } l = \text{`--'}\ : [\![c' = d]\!]\end{cases}$$

$$\cdot \begin{cases}\text{if } l = \text{`+'} \vee [d \mathinner{..} \Rightarrow c'] \in \text{Ae–h,Me–h} : [\![i' = j]\!]\\ \text{if } l = \text{`--'} \wedge [d \mathinner{..} \Rightarrow c'] \in \text{Aa–d,Ma–d}\ : [\![i' = \mathbf{i}_{Z+1}]\!]\end{cases}$$

$$\cdot \begin{cases}\text{if } l = \text{`+'} : [\![r^A = \langle i', \text{`id'}, j\rangle]\!]\\ \text{if } l = \text{`--'}\ : [\![r^A = \langle \text{`--'}, \text{`--'}, \text{`--'}\rangle]\!]\end{cases} \quad (1)$$

$$\mathsf{P}_{\beta_{s,\ell}}(\langle k,e\rangle\ r^B \mid \langle i,c\rangle\ \langle j,d\rangle) \overset{\text{def}}{\propto} \mathsf{P}_{\gamma_{s,\ell}}(c \rightarrow d\ e) \cdot \begin{cases}\text{if } [d\ e \Rightarrow c] \in \text{Aa–d,Ma–d} : [\![k = i]\!]\\ \text{if } [d\ e \Rightarrow c] \in \text{Ae–h,Me–h}\ : [\![k = \mathbf{i}_{Z+1}]\!]\end{cases}$$

$$\cdot \begin{cases}\text{if } [d\ e \Rightarrow c] \in \text{Aa–d,Me–h} : [\![r^B = \langle k, V(e), j\rangle]\!]\\ \text{if } [d\ e \Rightarrow c] \in \text{Ae–h,Ma–d} : [\![r^B = \langle j, V(d), k\rangle]\!]\\ \text{if } [d\ e \Rightarrow c] \in \text{Fa–c}\quad : [\![r^B = \langle \text{`--'}, \text{`--'}, \text{`--'}\rangle]\!]\end{cases} \quad (2)$$

$$\mathsf{P}_{\kappa_\ell}(r^K \mid \langle i,c\rangle\ \langle i',c'\rangle\ \langle j,d\rangle\ \langle k,e\rangle) \overset{\text{def}}{=}$$

$$\begin{cases}\text{if } c \in C \times \{\textbf{-g}\} \times C\ \wedge\ \exists_{d'}\ d'\ e \Rightarrow c'\ \wedge\ [d \Rightarrow d'] \in \text{Ga–b} : [\![r^K = \langle i', V(d), i\rangle]\!]\\ \text{if } c \in C \times \{\textbf{-g}\} \times C\ \wedge\ \exists_{e'}\ d\ e' \Rightarrow c'\ \wedge\ [e \Rightarrow e'] \in \text{Ga–b} : [\![r^K = \langle i', V(e), i\rangle]\!]\\ \text{if } c \in C \times \{\textbf{-g}\} \times C\ \wedge\ \exists_{d'}\ d'\ e \Rightarrow c'\ \wedge\ [d \Rightarrow d'] \in \text{Gc} : [\![r^K = \langle i, 1, i'\rangle]\!]\\ \text{if } c \in C \times \{\textbf{-g}\} \times C\ \wedge\ \exists_{e'}\ d\ e' \Rightarrow c'\ \wedge\ [e \Rightarrow e'] \in \text{Gc} : [\![r^K = \langle i, 1, i'\rangle]\!]\\ \text{otherwise} : [\![r^K = \langle \text{`--'}, \text{`--'}, \text{`--'}\rangle]\!]\end{cases} \quad (3)$$

Extras 1: Probabilistic Formulae

$$\mathsf{P}_{\sigma}(q_t^{1..N} \mid q_{t-1}^{1..N} x_{t-1}) \stackrel{\text{def}}{=} \mathsf{P}_{\phi_\ell}(\text{`}-\text{'} \mid b_{t-1}^{\ell}\ x_{t-1}) \cdot \mathsf{P}_{\sigma'_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N}\ a_{t-1}^{\ell})$$

$$+ \mathsf{P}_{\phi_\ell}(\text{`}+\text{'} \mid b_{t-1}^{\ell}\ x_{t-1}) \cdot \mathsf{P}_{\sigma'_{\ell+1}}(q_t^{1..N} \mid q_{t-1}^{1..N}\ x_{t-1});\quad \ell \stackrel{\text{def}}{=} \max\{\ell' \mid q_{t-1}^{\ell'} \neq \text{`}-\text{'}\} \tag{4}$$

$$\mathsf{P}_{\sigma'_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a') \stackrel{\text{def}}{=} \mathsf{P}_{\lambda_\ell}(\text{`}+\text{'} \mid b_{t-1}^{\ell-1}\ a') \cdot [\![a = a_{t-1}^{\ell-1}]\!] \cdot \mathsf{P}_{\beta_{\mathrm{B},\ell-1}}(b \mid b_{t-1}^{\ell-1}\ a') \cdot \mathsf{P}_{\sigma''_{\ell-1}}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b\, a')$$

$$+ \mathsf{P}_{\lambda_\ell}(\text{`}-\text{'} \mid b_{t-1}^{\ell-1}\ a') \cdot \mathsf{P}_{\alpha_\ell}(a \mid b_{t-1}^{\ell-1}\ a') \cdot \mathsf{P}_{\beta_{\mathrm{A},\ell}}(b \mid a_t^{\ell}\ a') \cdot \mathsf{P}_{\sigma''_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b\, a') \tag{5}$$

$$\mathsf{P}_{\sigma''_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b\, a') \stackrel{\text{def}}{=} \mathsf{P}_{\nu_{\ell-1}}(\text{`}+\text{'} \mid a_{t-1}^{\ell-1}\ b_{t-1}^{\ell-1}\ a_{t-1}^{\ell}\ b) \cdot \mathsf{P}_{\kappa_{\ell-1}}(r^K \mid b_{t-1}^{n}\ b_{t-1}^{\ell}\ a'\ b) \cdot \mathsf{P}_{\sigma'''_{\ell-1}}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b)$$

$$+ \mathsf{P}_{\nu_{\ell-1}}(\text{`}-\text{'} \mid a_{t-1}^{\ell-1}\ b_{t-1}^{\ell-1}\ a_{t-1}^{\ell}\ b) \cdot \mathsf{P}_{\kappa_{\ell-1}}(r^K \mid b_{t-1}^{n}\ b_{t-1}^{\ell}\ a'\ b) \cdot \mathsf{P}_{\sigma'''_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b) \tag{6}$$

$$\mathsf{P}_{\sigma'''_\ell}(q_t^{1..N} \mid q_{t-1}^{1..N} a\, b) \stackrel{\text{def}}{=} [\![q_t^{1..\ell-1} = q_{t-1}^{1..\ell-1}]\!] \cdot [\![a_t^{\ell} = a]\!] \cdot [\![b_t^{\ell} = b]\!] \cdot [\![q_t^{\ell+1..N} = \text{`}-\text{'}]\!] \tag{7}$$

Bibliography I

Chen, E., Gibson, E., and Wolf, F. (2005).
Online syntactic storage costs in sentence comprehension.
Journal of Memory and Language, 52(1):144–169.

Chomsky, N. and Miller, G. A. (1963).
Introduction to the formal analysis of natural languages.
In *Handbook of Mathematical Psychology*, pages 269–321. Wiley, New York, NY.

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

Bibliography II

Fossum, V. and Levy, R. (2012).
Sequential vs. hierarchical syntactic models of human incremental sentence processing.
In *Proceedings of CMCL 2012*. Association for Computational Linguistics.

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.

Hale, J. (2001).
A probabilistic earley parser as a psycholinguistic model.
In *Proceedings of the second meeting of the North American chapter of the Association for Computational Linguistics*, pages 159–166, Pittsburgh, PA.

Bibliography III

Hale, J. (2003).
Grammar, Uncertainty and Sentence Processing.
PhD thesis, Cognitive Science, The Johns Hopkins University.

Nguyen, L., van Schijndel, M., and Schuler, W. (2012).
Accurate unbounded dependency recovery using generalized categorial grammars.
In *Proceedings of the 24th International Conference on Computational Linguistics (COLING '12)*, Mumbai, India.

Petrov, S., Barrett, L., Thibaux, R., and Klein, D. (2006).
Learning accurate, compact, and interpretable tree annotation.
In *Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics (COLING/ACL'06)*.

Bibliography IV

Pynte, J., New, B., and Kennedy, A. (2008).
On-line contextual influences during reading normal text: A multiple-regression analysis.
Vision research, 48(21):2172–2183.

Rimell, L., Clark, S., and Steedman, M. (2009).
Unbounded dependency recovery for parser evaluation.
In *Proceedings of EMNLP 2009*, volume 2, pages 813–821.

van Schijndel, M., Exley, A., and Schuler, W. (2013).
A model of language processing as hierarchic sequential prediction.
Topics in Cognitive Science.

van Schijndel, M. and Schuler, W. (2013).
An analysis of frequency- and recency-based processing costs.
In *Proceedings of NAACL-HLT 2013*. Association for Computational Linguistics.

Bibliography V

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.