Semantic Complexity and Linguistic Distributions

Jakub Szymanik

Institute for Logic, Language and Computation
University of Amsterdam

LEGO, 21 February 2014

Outline

Equivalent complexity thesis

Linguists and non-linguists alike agree in seeing human language as the clearest mirror we have of the activities of the human mind, and as a specially important of human culture, because it underpins most of the other components. Thus, if there is serious disagreement about whether language complexity is a universal constant or an evolving variable, that is surely a question which merits careful scrutiny. There cannot be many current topics of academic debate which have greater general human importance than this one. (Sampson, 2009)

How do we measure complexity?

Existing approaches depend on implementation/theory:

- Chomsky hierarchy
- Typological approach (McWhorther, 2001; Everett, 2008)
- Information-theoretic approach (Juola, 2009)

Outline

Inherent complexity

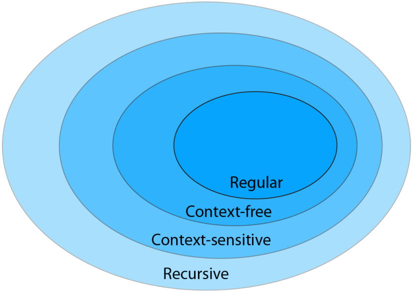
Inherent complexity

- Inherent complexity of the problem/concept

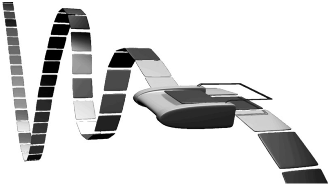
Inherent complexity

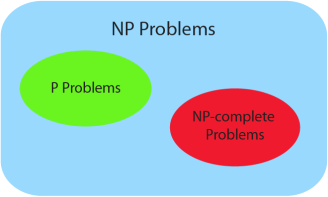
- Inherent complexity of the problem/concept
- and not the particular implementation.

E.g. in terms of Chomsky's Hierarchy



Or (in)tractability border





$$\exists x_1 \ldots \exists x_{k+1} \exists y_1 \ldots \exists x_{m+1} \Big[\bigwedge_{1 \le i < j \le k+1} x_i \neq x_j \ \wedge \bigwedge_{1 \le i < j \le m+1} y_i \neq y_j \\ \wedge \bigwedge_{1 \le i \le k+1} V(x_i) \ \wedge \bigwedge_{1 \le j \le m+1} T(y_j) \ \wedge \bigwedge_{\substack{1 \le i \le k+1 \\ 1 \le j \le m+1}} H(x_i, y_j) \Big].$$

Various semantic problems

- Inferential meaning
  $\hookrightarrow$ complexity of reasoning (satisfiability)
- Referential meaning
  $\hookrightarrow$ complexity of verification (model-checking)

They are closely related (Gottlob et al., 1999).

Outline

Intuition

- How complex are natural language arguments?
- It depends on the underlying natural logic (Moss, 2010; Muskens 2010).

Intuition

- How complex are natural language arguments?
- It depends on the underlying natural logic (Moss, 2010; Muskens 2010).

Example

$$\frac{\text{Every Italian loves pasta and football.} \quad \text{Camilo is Italian}}{\text{Camilo loves pasta}}$$

Intuition

- How complex are natural language arguments?
- It depends on the underlying natural logic (Moss, 2010; Muskens 2010).

Example

$$\frac{\text{Every Italian loves pasta and football.} \quad \text{Camilo is Italian}}{\text{Camilo loves pasta}}$$

$$\frac{\text{Everyone likes everyone who likes Pat} \quad \text{Pat likes every clarinetist}}{\text{Everyone likes everyone who likes everyone who likes every clarinetist}}$$

NL fragments

Fragment	Coverage	FO Operators and Relations
COP(¬)	Copula ("is a"), nouns ("man"), intransitive verbs ("runs"), "every", "some" names ("Joe"), adjectives ("thin") (+"not"))	$\{\forall, \exists, (\neg)\}$ $\cup$ $\{P_i^1 \mid i \in \mathbb{N}\}$
COP(¬)+TV	COP(¬) +transitive verbs ("loves")	$\{\forall, \exists, (\neg)\}$ $\cup \{P_i^1, P_j^2 \mid i, j \in \mathbb{N}\}$
COP(¬)+DTV	COP(¬) +ditransitive verbs ("gives")	$\{\forall, \exists, (\neg)\}$ $\cup \{P_i^1, P_j^3 \mid i, j \in \mathbb{N}\}$
COP(¬)+TV+DTV	COP(¬)+TV + ditransitive verbs	$\{\forall, \exists, (\neg)\}$ $\cup \{P_i^1, P_j^2, P_k^3 \mid i, j, k \in \mathbb{N}\}$
COP(¬)+Rel	COP(¬)+relative pronouns ("who", "that", "which") "and", intersective adjectives (+"or")	$\{\forall, \exists, \wedge, (\neg, \vee)\}$ $\cup$ $\{P_i^1 \mid i \in \mathbb{N}\}$
COP(¬)+Rel+TV	COP(¬)+Rel +transitive verbs	$\{\forall, \exists, \wedge, (\neg, \vee)\}$ $\cup \{P_i^1, P_j^2 \mid i, j \in \mathbb{N}\}$
COP(¬)+Rel+DTV	COP(¬)+Rel +ditransitive verbs	$\{\forall, \exists, \wedge, (\neg, \vee)\}$ $\cup \{P_i^1, P_j^3 \mid i, j \in \mathbb{N}\}$
COP(¬)+Rel+TV+DTV	COP(¬)+Rel+TV +ditransitive verbs	$\{\forall, \exists, \wedge, (\neg, \vee)\}$ $\cup \{P_i^1, P_j^2, P_k^3 \mid i, j, k \in \mathbb{N}\}$

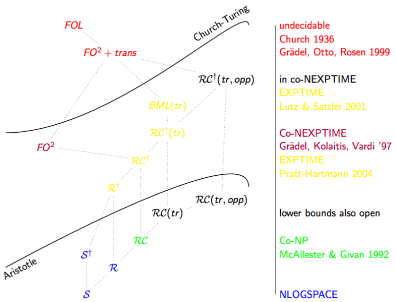
(Pratt-Hartmann & Third 2010; Thorne, 2010)

Examples of fragments

Fragment	Example	Fo
COP	Every politician cheats	$\forall x(\mathit{Politician}(x) \rightarrow \mathit{Cheat}(x))$
COP$^{\neg}$	Some philosopher is not trustworthy	$\exists x(\mathit{Philosopher}(x) \wedge \neg \mathit{Trusted}(x))$
COP$^{\neg}$+TV	John does not love Luke	$\neg \mathit{Loves}(\mathsf{John}, \mathsf{Luke})$
COP+TV +DTV	John gives a book to Jane Some man likes every candy	$\exists x \mathit{Book}(x) \wedge$ $\mathit{Gives}(\mathsf{John}, x, \mathsf{Jane})$ $\exists x(\mathit{Man}(x) \wedge$ $\forall y \mathit{Candy}(x) \rightarrow \mathit{Likes}(x, y))$
COP +Rel	Every idiot who is a philosopher cheats	$\forall x(\mathit{Idiot}(x) \wedge \mathit{Philosopher}(x)$ $\rightarrow \mathit{Cheat}(x))$
COP$^{\neg}$ +Rel	Some man who does not cheat is trustworthy	$\forall x(\mathit{Man}(x) \wedge \neg \mathit{Cheat}(x)$ $\rightarrow \mathit{Trusted}(x))$
⋮	⋮	⋮

Complexity results

- Fragments that contain either negation or relatives are tractable.
- Having both makes for intractable semantic complexity.



(Pratt-Hartmann 2010; Thorne, 2010; Larry Moss, 2010)

Outline

Quantifiers

1. All poets have low self-esteem.
2. Some dean danced nude on the table.
3. At least 3 grad students prepared presentations.
4. An even number of the students saw a ghost.
5. Most of the students think they are smart.
6. Less than half of the students received good marks.
7. Many of the soldiers have not eaten for several days.
8. A few of the conservatives hate each other.

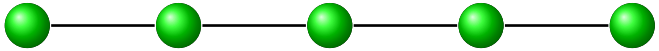
Simple quantifiers

Quantifier Q	Model Class	S. C.	Class
some	$\{\mathcal{I} \mid A^{\mathcal{I}} \cap B^{\mathcal{I}} \neq \emptyset\}$	$\mathbf{AC}^0$	Aristotelian (*ari*)
all	$\{\mathcal{I} \mid A^{\mathcal{I}} \subseteq B^{\mathcal{I}}\}$	$\mathbf{AC}^0$	
the	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) = 1\}$	$\mathbf{AC}^0$	Counting (*cnt*)
$> k$	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) > k\}$	$\mathbf{AC}^0$	
$< k$	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) < k\}$	$\mathbf{AC}^0$	
k	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) = k\}$	$\mathbf{AC}^0$	
most	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) > \#(A^{\mathcal{I}} \setminus B^{\mathcal{I}})\}$	$\mathbf{L}$	Proportional (*pro*)
few	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) < \#(A^{\mathcal{I}} \setminus B^{\mathcal{I}})\}$	$\mathbf{L}$	
$> p/k$	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) > p \cdot (\#(A)/k)\}$	$\mathbf{L}$	
$< p/k$	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) < p \cdot (\#(A)/k)\}$	$\mathbf{L}$	
p/k	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) = p \cdot (\#(A)/k)\}$	$\mathbf{L}$	
$> k\%$	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) > k \cdot (\#(A)/100)\}$	$\mathbf{L}$	
$< k\%$	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) < k \cdot (\#(A)/100)\}$	$\mathbf{L}$	
$k\%$	$\{\mathcal{I} \mid \#(A^{\mathcal{I}} \cap B^{\mathcal{I}}) = k \cdot (\#(A)/100)\}$	$\mathbf{L}$	

(In)tractable Reciprocal Constructions

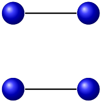
(In)tractable Reciprocal Constructions

Five pitchers sat alongside each other.

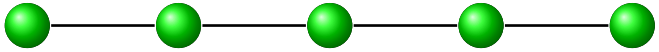


(In)tractable Reciprocal Constructions

Some Pirates were staring at each other.

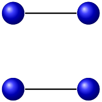


Five pitchers sat alongside each other.

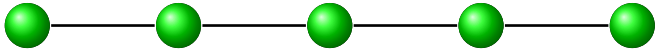


(In)tractable Reciprocal Constructions

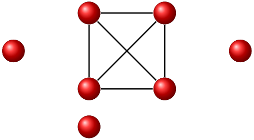
Some Pirates were staring at each other.



Five pitchers sat alongside each other.

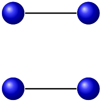


Most PMs referred to each other.

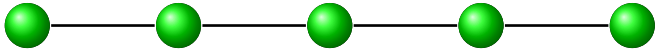


(In)tractable Reciprocal Constructions

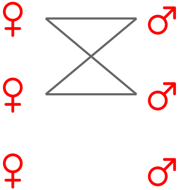
Some Pirates were staring at each other.



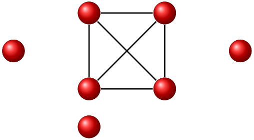
Five pitchers sat alongside each other.



Most girls and most boys hate each other



Most PMs referred to each other.



(Gierasimczuk & Szymanik, 2009; Szymanik, 2010)

Outline

Principle of least effort in communication

1. Speakers tend to use “simple" messages.

Principle of least effort in communication

1. Speakers tend to use “simple" messages.
2. Therefore, semantic complexity should correlate with linguistic frequency.
3. We would expect power law distributions (Zipf law).

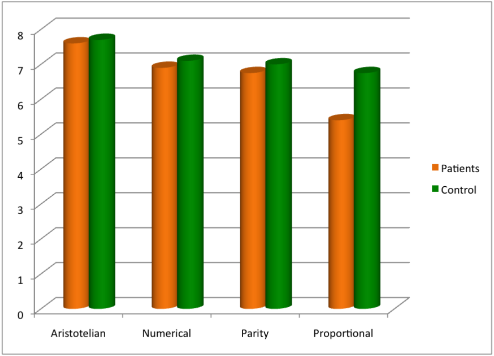
Intermezzo: semantic complexity and processing load

Verification times, WM involvement, comprehension, cognitive load, etc.
All can be predicted by semantic complexity.

Intermezzo: semantic complexity and processing load

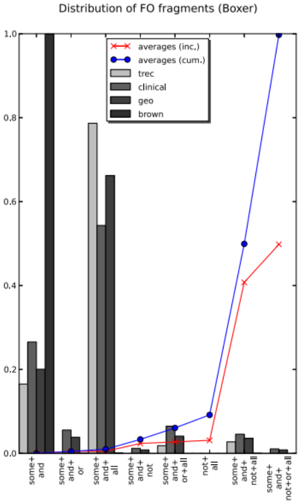
Verification times, WM involvement, comprehension, cognitive load, etc.
All can be predicted by semantic complexity.

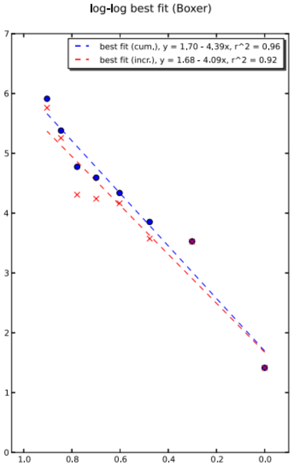
Example



(Zajenkowski et al., 2010)

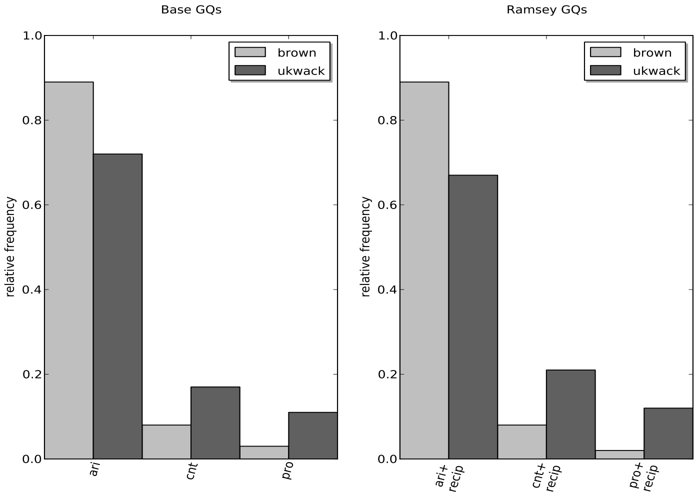
Fragments’ distribution and power law regression





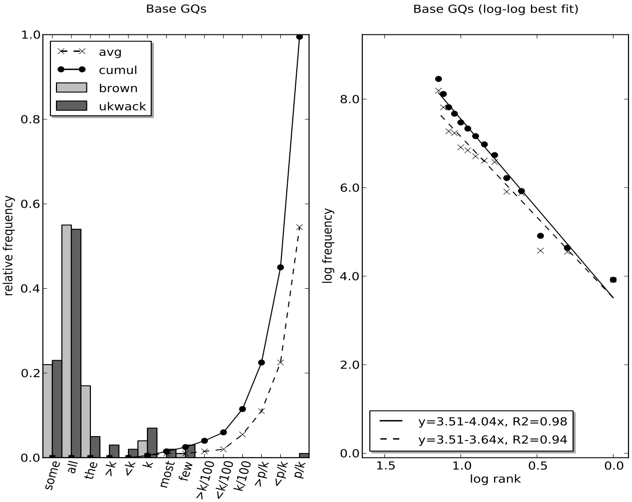
(Thorne, 2012)

Quantifier distribution by classes

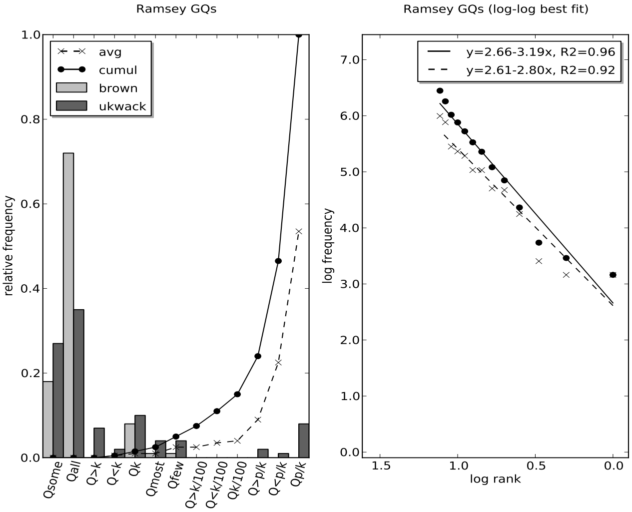


(Thorne & Szymanik, 2014)

Base quantifier distribution and power law regression



Ramsey quantifier distribution and power law regression



Summary

- Computationally easier expressions occur exponentially more frequent.
- Semantic complexity can quantify linguistic simplicity.
- Additional support for the cognitive studies.
- Semantic complexity is an empirically fruitful notion.
- Next step, apply it to equivalent complexity thesis.

Outline

Generalized Quantifiers

Definition

A quantifier Q is a way of associating with each set M a function from pairs of subsets of M into $\{0, 1\}$ (False, True).

Example

$$\text{every}_M[A, B] = 1 \text{ iff } A \subseteq B$$

Generalized Quantifiers

Definition

A quantifier Q is a way of associating with each set M a function from pairs of subsets of M into $\{0, 1\}$ (False, True).

Example

$$\text{every}_M[A, B] = 1 \text{ iff } A \subseteq B$$

$$\text{even}_M[A, B] = 1 \text{ iff } \text{card}(A \cap B) \text{ is even}$$

Generalized Quantifiers

Definition

A quantifier Q is a way of associating with each set M a function from pairs of subsets of M into $\{0, 1\}$ (False, True).

Example

$$\text{every}_M[A, B] = 1 \text{ iff } A \subseteq B$$

$$\text{even}_M[A, B] = 1 \text{ iff } \text{card}(A \cap B) \text{ is even}$$

$$\text{most}_M[A, B] = 1 \text{ iff } \text{card}(A \cap B) > \text{card}(A - B)$$

Space of GQs

- If $\text{card}(M) = n$, then there are $2^{2^{2n}}$ GQs.
- For $n = 2$ it gives 65,536 possibilities.

Space of GQs

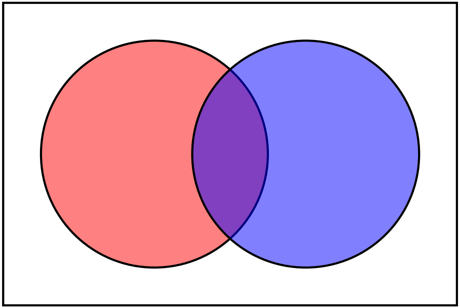
- If $\text{card}(M) = n$, then there are $2^{2^{2n}}$ GQs.
- For $n = 2$ it gives 65,536 possibilities.

Question

Which of those correspond to simple determiners?

Isomorphism closure

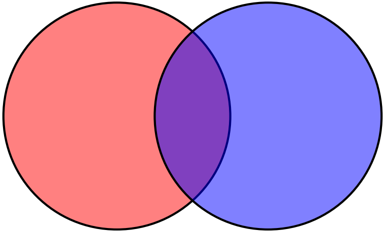
(ISOM) If $(M, A, B) \cong (M', A', B')$, then $\mathrm{Q}_{\mathbf{M}}(A, B) \Leftrightarrow \mathrm{Q}_{\mathbf{M'}}(A', B')$



Topic neutrality

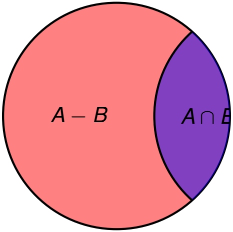
Extensionality

(EXT) If $M \subseteq M'$, then $\mathsf{Q}_{\mathbf{M}}(A, B) \Leftrightarrow \mathsf{Q}_{\mathbf{M'}}(A, B)$



Conservativity

(CONS) $Q_{\mathbf{M}}(A, B) \Leftrightarrow Q_{\mathbf{M}}(A, A \cap B)$



Semantic complexity as universale

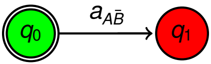
- Some expressions may be even too hard to appear in NL.
  - E.g, some collective quantifiers can be crazy complex!
- Complexity as a test of methodological plausibility of linguistic theories.

(Ristad, 1993; Mostowski & Szymanik, 2012; Kontinen & Szymanik, 2014)

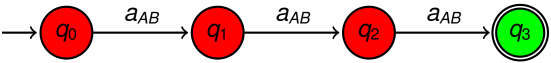
Thanks for your attention

Quantifiers and Chomsky's Hierarchy

All As are B.

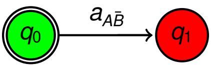


More than 2 As are B.

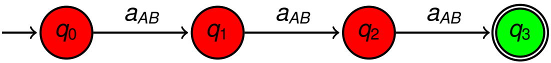


Quantifiers and Chomsky's Hierarchy

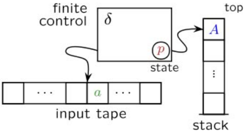
All As are B.



More than 2 As are B.

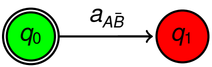


Most As are B.

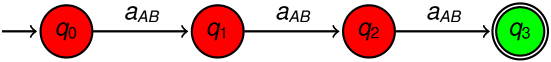


Quantifiers and Chomsky's Hierarchy

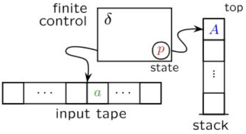
All As are B.



More than 2 As are B.



Most As are B.



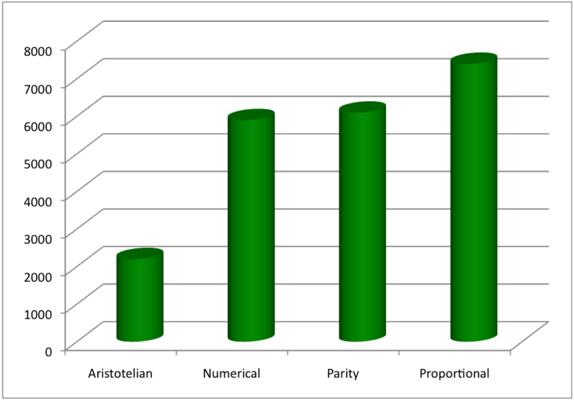
van Benthem, Essays in logical semantics, 1986

Mostowski, Computational semantics for monadic quantifiers, 1998

A simple study

More than half of the cars are yellow.

Verification times can be predicted by complexity



Szymanik & Zajenkowski, Comprehension of simple quantifiers. Empirical evaluation of a computational model, Cognitive Science, 2010

Neurobehavioral prediction wrt working memory is satisfied

Differences in brain activity.

- Only proportional quantifiers activate working-memory capacity: recruit right dorsolateral prefrontal cortex.

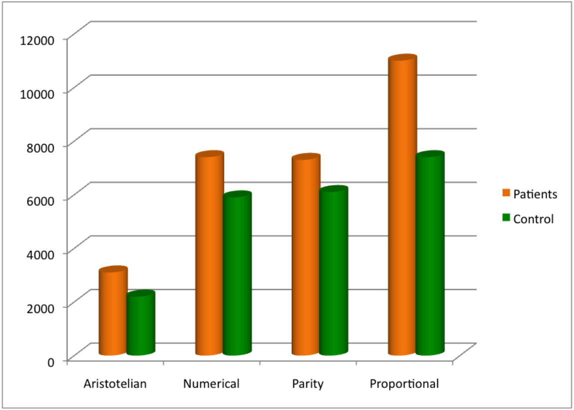
McMillan et al., Neural basis for generalized quantifiers comprehension, Neuropsychologia, 2005

Szymanik, A Note on some neuroimaging study of natural language quantifiers comprehension, Neuropsychologia, 2007

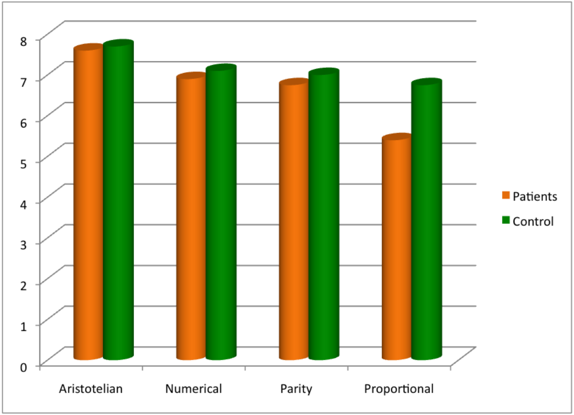
Experiment with schizophrenic patients

- Compare performance of:
  - Healthy subjects.
  - Patients with schizophrenia.
    - Known WM deficits.

Patients are generally slower



Patients are only less accurate with proportional quantifiers



Zajenkowski et al., A computational approach to quantifiers as an explanation for some language impairments in schizophrenia, Journal of Communication Disorders, 2011.

Comprehension and verification are influenced by complexity

1. Draw and verify:
   - All/Most of the dots are directly connected to each other.

Comprehension and verification are influenced by complexity

1. Draw and verify:
   - All/Most of the dots are directly connected to each other.
2. In line with complexity:
   - Fewer strong pictures for ‘most’
   - Better performance on complete graphs for ’All’-condition

Bott et al., Interpreting Tractable versus Intractable Reciprocal Sentences, Proceedings of the International Conference on Computational Semantics, 2011.

Schlotterbeck & Bott, Easy solutions for a hard problem? The computational complexity of reciprocals with quantificational antecedents, Proc. of the Logic & Cognition Workshop at ESSLLI 2012.