Semantic Complexity of Quantifiers and their Distribution in Corpora

Jakub Szymanik\textsuperscript{1}  Camilo Thorne\textsuperscript{2}

\textsuperscript{2}IBM CAS Trento - Trento RISE  
Povo di Trento, Italy  
c.thorne.email@trentorise.eu

\textsuperscript{1}Institute of Logic, Language and Computation - UvA  
Amsterdam, The Netherlands  
J.K.Szymanik@uva.nl

IWCS 2015, April 15–17, London
Generalized Quantifiers

some men are happy

all humans are mammals

50 MPs voted against the war in Irak

fewer than 100 violins are Stradivari

more than 2/3 of planets are lifeless

33% of Peru’s population lives in Lima

some children like each other

more than 2/3 of female MPs sit next to each other
Generalized Quantifiers

\[\text{some men are happy}\]
\[\text{all humans are mammals}\]
\[50 \text{ MPs voted against the war in Irak}\]
\[\text{fewer than 100 violins are Stradivari}\]
\[\text{more than } \frac{2}{3} \text{ of planets are lifeless}\]
\[33\% \text{ of Peru’s population lives in Lima}\]
\[\text{some children like each other}\]
\[\text{more than } \frac{2}{3} \text{ of female MPs sit next to each other}\]
| Base Quantifiers $Q$                        | Complexity |
|--------------------------------------------|------------|
| Aristotelian ($ari$)                       | $AC^0$     |
| counting ($cnt$)                            | $AC^0$     |
| proportional ($pro$)                        | $P$        |

| Ramseyfied Quantifiers $R_Q$                | Complexity |
|--------------------------------------------|------------|
| Ramsey+Aristotelian ($ari$+$recip$)        | $AC^0$     |
| Ramsey+counting ($cnt$+$recip$)            | $AC^0$     |
| Ramsey+proportional ($pro$+$recip$)        | NP-complete |
Quantifier Distribution

Corpus distribution is biased towards tractable quantifiers.
A fine-grained analysis yields power law distributions.
1. Corpus distribution is biased towards tractable quantifiers

2. A fine-grained analysis yields power law distributions