# **Academic Collocation List: Development Chart**

PICAE

### •Pearson International Corpus of Academic English

•Quantitative analysis of the written academic component of PICAE

### Complete reference list

•List of all content words that occur more than 12 times in the corpus

**CRL** 

## •Reference list of important academic words

- •Frequency: ≥5 per million; distribution: ≥5 different texts
- •GSL words were removed

List of Academic Words

- •List to exclude frequent function words that express purely grammatical meaning
- •to exclude be, do, have

Stop List

### Initial collocation list

- •c. 130,000 collocations: academic word + co-occurring words in the span of ±3
- •Frequency: at least ≥1 per million; distribution: ≥5 different texts; MI score: ≥3; t-score: ≥2

**IACL** 

- Tagging of initial collocation list
- •Cleaning based on quantitative parameters: normed frequency per million: ≥1; dispersion: ≥.2; MI score: ≥3; t-score: ≥4

Manual Vetting

POS-filtering

Expert Judgment •Expert panel to review the academic collocation list and qualitatively assess the pedagogical value of each collocation

#### Academic Collocation List

•Completion September 2011

**ACL**