(Association) measure for measure: Evaluating the corpus-based identification of lexical collocations

Stefan Evert
FAU Erlangen-Nürnberg

— joint work with Sabine Bartsch, Thomas Proisl, Peter Uhrig —
Lexical collocations

- Lexical collocation = “habitual” recurrent word combination (Firth 1957, Sinclair 1966)
  - complex phenomenon without clear definition
  - but supported by strong native-speaker intuitions

- Some examples
  - cow–milk, day–night (Firth 1957)
  - blonde hair, catchy tune, commit a crime, spill the beans
  - make/thwart an attempt, feeble/successful attempt, ...

- Important applications in lexicography, language teaching, digital humanities & social sciences, NLP, ...
Related phenomena

- Lexicalized multiword expressions (MWE)
  - kick the bucket, eat humble pie, spill the beans, ...
  - hand in, carry on, turn down, give a speech, ...
  - bucket seat, black hole, New York City, ...

- Psychological association
  - apple → red, orange, fruit, pie, pear, worm, ...
  - dolphin → fish, sea, fin, mammal, whale, blue, ...

- Co-occurrence in text corpora
  🔄 Overlap with lexical collocations & each other
Collocation identification

- Automatic identification of collocations and MWE based on large electronic corpora
- Main cue: co-occurrence frequency, quantified by various statistical association measures (AM)
- Other criteria derived from properties of MWE (Manning & Schütze 1999, 184)
  - non-compositionality (distributional semantics)
  - non-modifiability (syntactic flexibility, fixed ordering)
  - non-substitutability (substitution tests)
- AM are most appropriate cue for lexical collocations
Collocation identification

- Best techniques & parameters for collocation/MWE identification are still unclear
  - e.g. choice of more than 57 AMs (Pecina 2005)
- Need for empirical comparative evaluation
- Prior work: (subtype of) MWE with relatively clear definition → manual validation of MWE candidates

- Today: results of large-scale evaluation study on the automatic identification of English lexical collocations
  - special focus: linguistic usefulness of Web corpora
Research questions

- Which AM correlates best with lexical collocations?
- What is an appropriate window of co-occurrence?
- Which source corpora provide the best results?
- Does size matter? Or representativeness?
- Is automatic annotation with NLP tools useful?
- Are there interactions between these parameters?
- Are crawled Web corpora and n-gram databases a viable substitute for expensive reference corpora?
- long-term goal: better understanding of collocations
Gold standard

- Standard evaluation methodology (Evert & Krenn 2001, 2005)
  - obtain ranked list of candidates from automatic system
  - manual validation by expert → TPs, FPs
  - compute & compare quantitative evaluation scores such as precision (P), recall (R) and F-score (F) for n-best lists

- Difficult to apply in this study
  - not suitable for comprehensive large-scale evaluation
  - lack of clear annotation criteria → pro-corpus bias
Gold standard

- Alternative: use existing gold standard of “known” lexical collocations ➔ collocation dictionary
  - but see critical remarks by Evert (2004)

- BBI: The BBI Combinatory Dictionary of English (Benson, Benson & Ilson 1986)
  - based on lexicographic native-speaker intuitions
  - pre-corpus era ➔ no bias towards specific method/corpus

- OCD2: Oxford Collocations Dictionary for students of English, 2nd ed. (McIntosh, Francis & Poole 2009)
  - corpus-based, much more comprehensive
The Bartsch224 gold standard

- Set of 203 node words selected by Sabine Bartsch
  - original set contained approx. 224 node words
  - some obscure nodes with few collocates omitted
- Manually extracted all lexical words (nouns, verbs, adjectives, adverbs) from corresponding BBI entries
  - set of 2,845 node-collocate pairs
  - lemmatized, reduced to two-word collocations
- Automatic extraction from XML version of OCD2
  - also from other entries (node word listed as collocate)
  - set of 18,545 node-collocate pairs
Gold standard example

Node: measure (noun or verb)

cubic, dry, liquid, metric, tape, certain, good, make, take,

measure I  n.  1. a cubic; dry; liquid; metric ~ 2. a tape ~ 3. in a certain ~ (in large ~) 4. (misc.) for good ~ (‘as smt. extra’); made to ~ (‘custom-made’); to take smb.’s ~ (‘to evaluate smb.’) (see also measures)

measure II  v.  1. (d; tr.) to ~ against (to ~ one’s accomplishments against smb. else’s) 2. (P; intr.) the room ~s twenty feet by ten
Gold standard example

Node: measure (noun or verb)

- cubic, dry, liquid, metric, tape, certain, good, make, take, carry, coercive, compulsory, draconian, drastic, harsh, stern, stringent, emergency, extreme, radical, preventive, prophylactic, safety, security, stopgap, temporary

measures n. 1. to carry out, take ~ 2. coercive; compulsory; draconian; drastic, harsh, stern, stringent; emergency; extreme, radical; preventive, prophylactic; safety, security; stopgap, temporary ~ 3. ~ to + inf. (we took ~ to insure their safety) 4. ~ against (to take ~ against smuggling)
Parameters: association measure

- log-likelihood ($G^2$, Dunning 1993)
- Mutual Information ($MI$, Church & Hanks 1990)
- t-score ($t$, Church et al. 1991)
- $MI^2$ (Daille 1994)
- chi-squared ($X^2$) with Yates's correction
- co-occurrence frequency ($f$)
- $\Delta P$ (Gries 2013, a directional measure)
- Dice coefficient ($Dice$, SketchEngine)

(selection based on recommendations of Evert 2008)
Parameters: co-occurrence context

- syntactic co-occurrence: dependency relations
  - direct dependency (all types, both directions)
- surface co-occurrence: L1 / R1
- surface co-occurrence: L2 / R2
- surface co-occurrence: L3 / R3
- surface co-occurrence: L5 / R5
- surface co-occurrence: L10 / R10
- textual co-occurrence: sentence
### Parameters: corpus & annotation

<table>
<thead>
<tr>
<th>corpus</th>
<th>annotation</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>British National Corpus (BNC, Aston &amp; Burnard 1996)</td>
<td>C&amp;C, Stanford</td>
<td>0.1 G</td>
</tr>
<tr>
<td>Darmstadt English Movie Subtitle Corpus (DESC)</td>
<td>C&amp;C, Stanford</td>
<td>0.1 G</td>
</tr>
<tr>
<td>Gigaword newspaper corpus (2nd ed.)</td>
<td>C&amp;C, Stanford</td>
<td>2.0 G</td>
</tr>
<tr>
<td>English wikipedia of 2009 (Wackypedia)</td>
<td>C&amp;C, Malt, Stfd</td>
<td>1.0 G</td>
</tr>
<tr>
<td>Subcorpus WP500 (500 words per article)</td>
<td>C&amp;C, Malt, Stfd</td>
<td>0.2 G</td>
</tr>
<tr>
<td>Web corpus ukWaC (Baroni et al. 2009)</td>
<td>C&amp;C, Malt</td>
<td>2.0 G</td>
</tr>
<tr>
<td>Web corpus WebBase (Han et al. 2013)</td>
<td>C&amp;C</td>
<td>3.0 G</td>
</tr>
<tr>
<td>Web corpus UKCOW 2012 (Schäfer et al. 2012)</td>
<td>C&amp;C</td>
<td>4.0 G</td>
</tr>
<tr>
<td>Web corpus ENCCOW 2014</td>
<td>C&amp;C, Malt</td>
<td>10.0 G</td>
</tr>
<tr>
<td>All Web corpora + Wackypedia (JOINT)</td>
<td>C&amp;C</td>
<td>16.0 G</td>
</tr>
</tbody>
</table>
## Parameters: corpus & annotation

<table>
<thead>
<tr>
<th>corpus</th>
<th>annotation</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Web 1T 5-Grams (Web1T5, Brants &amp; Franz 2006)</td>
<td></td>
<td>1000 G</td>
</tr>
<tr>
<td>Google Books N-Grams 2012 (BooksEN, Lin et al. 2012)*</td>
<td>parsed</td>
<td>500G</td>
</tr>
<tr>
<td>Google Books N-Grams 2012 GB (BooksGB)*</td>
<td>parsed</td>
<td>50 G</td>
</tr>
</tbody>
</table>

* Google Books data sets only include n-gram counts from contemporary books published in 1980 and later (evaluation on full 20th century yields very similar results)
Evaluation methodology

<table>
<thead>
<tr>
<th>node</th>
<th>collocate</th>
<th>$G^2$</th>
<th>TP?</th>
</tr>
</thead>
<tbody>
<tr>
<td>measure</td>
<td>introduce</td>
<td>762.60</td>
<td>—</td>
</tr>
<tr>
<td>measure</td>
<td>take</td>
<td>600.13</td>
<td>++</td>
</tr>
<tr>
<td>measure</td>
<td>measure</td>
<td>510.33</td>
<td>—</td>
</tr>
<tr>
<td>measure</td>
<td>preventive</td>
<td>496.79</td>
<td>++</td>
</tr>
<tr>
<td>measure</td>
<td>success</td>
<td>475.13</td>
<td>—</td>
</tr>
<tr>
<td>measure</td>
<td>use</td>
<td>450.28</td>
<td>—</td>
</tr>
<tr>
<td>measure</td>
<td>concentration</td>
<td>428.35</td>
<td>—</td>
</tr>
<tr>
<td>measure</td>
<td>safety</td>
<td>413.83</td>
<td>++</td>
</tr>
<tr>
<td>measure</td>
<td>adopt</td>
<td>397.25</td>
<td>—</td>
</tr>
<tr>
<td>measure</td>
<td>distance</td>
<td>337.12</td>
<td>—</td>
</tr>
</tbody>
</table>

most strongly associated co-occurrences for node measure (n = 10)

$P = 3 \text{ TP} / 10 \text{ cand.} = 30\%$

$R = 3 / 26 \text{ TP} = 11.5\%$

TP = true positive

(according to BBI dictionary)
Evaluation methodology

- Precision / recall of n-best lists for each node
  - strategy used by Uhrig & Proisl (2012)
  - task: determine most salient collocates for given node
  - results averaged over all 203 nodes

- Precision vs. recall for list of all candidate pairs
  - strategy used by Bartsch & Evert (2013), following Evert & Krenn (2001, 2005)
  - task: determine most salient collocational pairs
    - we present results for this approach
**Evaluation: global ranking**

<table>
<thead>
<tr>
<th>node</th>
<th>collocate</th>
<th>$G^2$</th>
<th>TP?</th>
</tr>
</thead>
<tbody>
<tr>
<td>minister</td>
<td>prime</td>
<td>111653.34</td>
<td>++</td>
</tr>
<tr>
<td>prime</td>
<td>minister</td>
<td>103587.58</td>
<td>—</td>
</tr>
<tr>
<td>authority</td>
<td>local</td>
<td>64395.65</td>
<td>++</td>
</tr>
<tr>
<td>take</td>
<td>place</td>
<td>43787.49</td>
<td>—</td>
</tr>
<tr>
<td>place</td>
<td>take</td>
<td>42551.75</td>
<td>++</td>
</tr>
<tr>
<td>set</td>
<td>up</td>
<td>37871.00</td>
<td>—</td>
</tr>
<tr>
<td>state</td>
<td>secretary</td>
<td>37588.70</td>
<td>—</td>
</tr>
<tr>
<td>door</td>
<td>open</td>
<td>37287.35</td>
<td>++</td>
</tr>
<tr>
<td>open</td>
<td>door</td>
<td>37193.15</td>
<td>—</td>
</tr>
<tr>
<td>head</td>
<td>shake</td>
<td>32301.90</td>
<td>++</td>
</tr>
</tbody>
</table>

Most strongly associated node-collocate pairs (n = 10)

\[ P = 5 \text{ TP} / 10 \text{ cand.} = 50\% \]

\[ R = 5 / 2845 \text{ TP} = 0.2\% \]

TP = true positive

(according to BBI dictionary)
Evaluation: precision vs. recall | BBI

British National Corpus [100M] | L5/R5 span | gold: Bartsch224–BBI

coverage: 96.4%

n = 1193
R = 10.0%
P = 23.9%

Berlin, 9 Feb 2016
Evaluation: precision vs. recall | BBI

British National Corpus [100M] | L5/R5 span | gold: Bartsch224–BBI

coverage: 96.4%

Average Precision
up to 50% recall

AP50 = 17.4%
AP50 = 18.8%

baseline = 0.3%
Evaluation: precision vs. recall | BBI

British National Corpus [100M] | L5/R5 span | gold: Bartsch224–BBI

coverage: 96.4%

precision (%) vs. recall (%)

baseline = 0.3%

MI² = \log_2 \frac{O^2}{E}

heuristic measure (Daille 1994)
Evaluation: precision vs. recall | BBI

British National Corpus [100M] | L5/R5 span | gold: Bartsch224–BBI

Coverage: 96.4%

ΔP_{2|1} = \frac{O_{11}}{R_1} - \frac{O_{21}}{R_2}

Asymmetric ΔP
(Allan 1980; Gries 2013)
Factor: context size | BBI

British National Corpus [100M] | sentence window | gold: Bartsch224–BBI

coverage: 98.0%

precision (%) vs recall (%)
Factor: context size | BBI

British National Corpus [100M] | syntactic C&C | gold: Bartsch224–BBI

coverage: 91.6%

precision (%) vs recall (%)

baseline = 0.7%
Factor: different corpora | L2 / R2 | BBI

British National Corpus [100M] | L2/R2 span | gold: Bartsch224–BBI

coverage: 93.4%

precision (%) vs. recall (%) graph

baseline = 0.57%
Factor: different corpora | L2 / R2 | BBI

Joint Web corpus [16G] | L2/R2 span | gold: Bartsch224–BBI

coverage: 99.4%

precision (%) vs recall (%)

baseline = 0.1%
Result overview | BBI

Coverage of Bartsch224–BBI gold standard | L2/R2

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNC</td>
<td></td>
</tr>
<tr>
<td>DESC</td>
<td></td>
</tr>
<tr>
<td>Gigaword</td>
<td></td>
</tr>
<tr>
<td>WP500</td>
<td></td>
</tr>
<tr>
<td>Wiki</td>
<td></td>
</tr>
<tr>
<td>UKWAC</td>
<td></td>
</tr>
<tr>
<td>WEBBASE</td>
<td></td>
</tr>
<tr>
<td>UKCOW</td>
<td></td>
</tr>
<tr>
<td>ENCow</td>
<td></td>
</tr>
<tr>
<td>JOINT</td>
<td></td>
</tr>
<tr>
<td>WEB1T5</td>
<td></td>
</tr>
<tr>
<td>BooksGBmod</td>
<td></td>
</tr>
<tr>
<td>BooksENmod</td>
<td></td>
</tr>
</tbody>
</table>

Berlin, 9 Feb 2016
Result overview | BBI

AP 50 | L2/R2 span | gold: Bartsch224–BBI

Berlin, 9 Feb 2016
Result overview | BBI

AP 50 | syntactic dependency | gold: Bartsch224–BBI

0 10 20 30 40
AP50 (%)

C&C / Google
Malt
Stanford

BNC  DESC  Gigaword  WP500  Wiki  UKWAC  WEBBASE  UKCOW  ENCOW  JOINT  BooksGBmod  BooksENmod

Berlin, 9 Feb 2016
Result overview | BBI

AP 50 \( f \geq 1 \) gold: Bartsch224–BBI

- **BNC**
- **Wiki**
- **ENCOW**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
</tr>
</tbody>
</table>
Result overview | BBI

AP 50 | L2/R2 span | f >= 1 | gold: Bartsch224–BBI

AP50 (%)

BNC    DESC    Gigaword    WP500    Wiki    UKWAC    WEBBASE    UKCOW    ENCOW    JOINT    WEB1T5    BooksGBmod    BooksENmod

Berlin, 9 Feb 2016
Result overview | BBI

AP 50 | British National Corpus | f >= 1 | gold: Bartsch224–BBI

|------------|----------|---------|---------|---------|---------|---------|-----------|----------|

Berlin, 9 Feb 2016
Conclusions 1

- Chi-squared / MI2 seem to be best-suited AMs for identification of (BBI) lexical collocations
- Web corpora are a good replacement for BNC, but they need to be much larger & reasonably clean
- Corpus size matters, but very slow “learning curve”
- Small context windows & syntactic context optimal
- NLP annotation helps, but only with the right parser
- Few interactions: always same ranking of AMs
- But how good and reliable is our “gold standard”?
**rose** *noun*

**ADJ.** deep pink, red, yellow, etc. | early, late | climbing, rambler, rambling, shrub | long-stemmed | scented | fresh *She put fresh roses in the vases.* | wild

**ROSE + NOUN** garden | bed | bush | petals | bowl

**PHRASES** a bed of roses, the scent/smell of a rose > Note at FLOWER (for verbs)
Replication on OCD2 gold standard
Replication on OCD2 gold standard

Joint Web corpus [16G] | L2/R2 span | gold: Bartsch224–OCD2

coverage: 99.6%

baseline = 0.66%
Result overview | OCD

Coverage of Bartsch224–OCD2 gold standard | L2/R2

Coverage (%)

BNC  DESC  Gigaword  WP500  Wiki  UKWAC  WEBBASE  UKCOW  ENCow  JOINT  WEB1T5  BooksGBmod  BooksENmod

Berlin, 9 Feb 2016
Result overview | OCD

AP 50 | best result | gold: Bartsch224–OCD2

AP50 (%)

BNC  DESC  Gigaword  WP500  Wiki  UKWAC  WEBSSE  UKCOW  ENCow  JOINT  WEB1T5  BooksGBmod  BooksENmod

Berlin, 9 Feb 2016
Result overview | OCD

![Graph showing AP 50 for syntactic dependency with gold data from Bartsch224–OCD2](image_url)
Result overview | OCD

AP 50 | f >= 1 | gold: Bartsch224–OCD2

AP50 (%)

BNC
Wiki
ENCOW

Berlin, 9 Feb 2016
## Result overview | OCD

### AP 50 I L2/R2 span I f >= 1 I gold: Bartsch224–OCD2

<table>
<thead>
<tr>
<th>Dataset</th>
<th>G²</th>
<th>MI</th>
<th>t</th>
<th>Mi²</th>
<th>X²</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DESC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gigaword</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiki</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UKWAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEBBASE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UKCOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENCow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JOINT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEB1T5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BooksGBmod</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BooksENmod</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Berlin, 9 Feb 2016
But what about frequency thresholds?

- Evaluation experiments based on *all* co-occurrences found in corpora, even if not recurrent \( (f = 1) \)
- It is common in collocation research to apply a frequency threshold (esp. with MI measure)
  - prior work often reports substantially better results than without the frequency threshold
- Approach 1: fixed threshold, e.g. \( f \geq 5 \)
- Approach 2: relative threshold, e.g. \( f \geq 10 / G \) words
### Frequency thresholds

#### Coverage of Bartsch224–BBI gold standard | L2/R2

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Coverage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNC</td>
<td></td>
</tr>
<tr>
<td>DESC</td>
<td></td>
</tr>
<tr>
<td>Gigaword</td>
<td></td>
</tr>
<tr>
<td>WP500</td>
<td></td>
</tr>
<tr>
<td>Wiki</td>
<td></td>
</tr>
<tr>
<td>UKWAC</td>
<td></td>
</tr>
<tr>
<td>WEBBASE</td>
<td></td>
</tr>
<tr>
<td>UKCOW</td>
<td></td>
</tr>
<tr>
<td>ENcow</td>
<td></td>
</tr>
<tr>
<td>JOINT</td>
<td></td>
</tr>
<tr>
<td>WEB1T5</td>
<td></td>
</tr>
<tr>
<td>BooksGBmod</td>
<td></td>
</tr>
<tr>
<td>BooksENmod</td>
<td></td>
</tr>
</tbody>
</table>

#### Coverage (%)

- **f ≥ 1**
- **f ≥ 5**
- **f ≥ 10/G**

---

*Berlin, 9 Feb 2016*
Frequency thresholds

AP 50 | L2/R2 span | gold: Bartsch224–BBI

- f ≥ 1
- f ≥ 5
- f ≥ 10/G

Berlin, 9 Feb 2016
Frequency thresholds

AP 50 | L2/R2 span | gold: Bartsch224–OCD2

AP50 (%)

0 10 20 30 40 50 60 70

BNC DESC Gigaword WP500 Wiki UKWAC WEBBASE UKCOW ENCOW JOINT WEB1T5 BooksGBmod BooksENmod

f ≥ 1
f ≥ 5
f ≥ 10/G

Berlin, 9 Feb 2016
Frequency effect?

For high thresholds, coverage falls below 50%.
Frequency effect?

AP 50 | syntactic dependency | gold: Bartsch224–BBI

- $f \geq 1$
- $f \geq 3$
- $f \geq 5$
- $f \geq 10$
- $f \geq 20$
- $f \geq 50$
- $f \geq 100$
- $f \geq 500$
- $f \geq 1000$

AP50 (%)

BNC
Wiki
ENCOW
Frequency effect?

AP 50 | ENCow (syntactic) | gold: Bartsch224−BBI

![Graph showing the frequency effect with different AP50 (% values for different frequency thresholds)]

- $f \geq 1$
- $f \geq 3$
- $f \geq 5$
- $f \geq 10$
- $f \geq 20$
- $f \geq 50$
- $f \geq 100$
- $f \geq 500$
- $f \geq 1000$
- $f \geq 5000$
- $f \geq 10000$
Conclusions 2

- Suitable choice of AM depends on gold standard, i.e. the precise type of collocations to be extracted
  - chi-squared or MI² for BBI collocations
  - log-likelihood (or t-score) for OCD2 collocations
- Plausible bias of OCD2 towards BNC & certain AMs
- Good news: no interaction with other parameters
  - best AMs are the same for (nearly) all parameter settings
  - consistent effect of window size, source corpus & parser
- Frequency thresholds have virtually no effect
  - AMs successful at weeding out low-frequency noise
Questions?

THANK YOU!
Appendix: Per-node evaluation

by node I British National Corpus [100M] I L5/R5 span (f >= 5) I gold: Bartsch224–BB
Appendix: Per-node evaluation

by node | British National Corpus [100M] | L5/R5 span (f >= 5) | gold: Bartsch224–BB

recall (%)

n–best list
Appendix: Manual validation
discrepancies between BBI and OCD2

<table>
<thead>
<tr>
<th>Bartsch 224, ENCOW, Malt dependencies, X2 'argue' (1000 candidates)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1535292</td>
<td>argue convincingly</td>
</tr>
<tr>
<td>1536640</td>
<td>argue forcefully</td>
</tr>
<tr>
<td>1539055</td>
<td>argue plausibly</td>
</tr>
<tr>
<td>1534717</td>
<td>argue case</td>
</tr>
<tr>
<td>1539037</td>
<td>argue plaintiff</td>
</tr>
<tr>
<td>1540747</td>
<td>argue strongly</td>
</tr>
<tr>
<td>1535589</td>
<td>argue defendant</td>
</tr>
<tr>
<td>1534170</td>
<td>argue author</td>
</tr>
<tr>
<td>1538056</td>
<td>argue passionately</td>
</tr>
<tr>
<td>1536059</td>
<td>argue economist</td>
</tr>
<tr>
<td>1540725</td>
<td>argue strenuously</td>
</tr>
<tr>
<td>989811</td>
<td>argue also</td>
</tr>
</tbody>
</table>

variable ordering: A..B 74.5% / B..A 25.5%  non-contiguous: 79.2% adjacent

Although no one has argued that amendments other than the Tenth and Fourteenth have impliedly amended the AC, one might plausibly argue that each has necessarily done so.

It cannot plausibly be argued, in my opinion, that the Human Rights Act erodes the sovereignty of Parliament or amounts to a usurpation of power by the judges.

Rather than viewing them as epistemically-morally-politically pernicious forms of hasty generalization by contrast, say, with Gadamerian pre-judgements or putatively more benign practices of categorization, Fricker argues plausibly for a "neutral" sense of stereotype which catches their frequent reliability as part of a "hearer's rational resources" in making credibility judgements.

But, it can be plausibly argued that we are getting there.

Given any statement, we can argue plausibly that it is about Maine.


References


References

- Gries, Stefan Th. (2013). 50-something years of work on collocations: What is or should be next ... . International Journal of Corpus Linguistics, 18(1), 137–165.
References