

*Quite diverse: A corpus-driven analysis  
of quite + ADJ in twenty varieties of English*

Natalia Levshina  
F.R.S. – FNRS  
Université catholique de Louvain

# Thinking globally...

- Traditionally, variational linguistics focuses on geographic and social distribution of formal variants (cf. eWAVE: Kortmann & Lunkenheimer 2013), whereas the semasiological differences between geographic and social variants of the same form remain understudied.

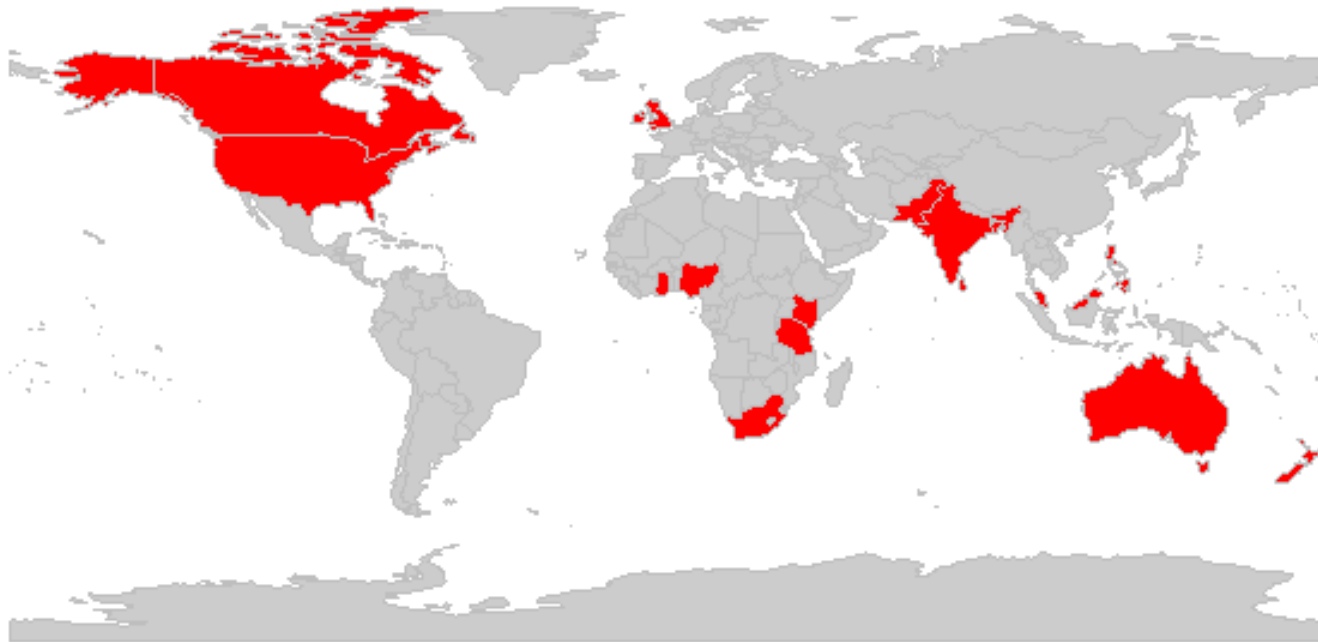
# Thinking globally...

- Traditionally, variational linguistics focuses on geographic and social distribution of formal variants (cf. eWAVE: Kortmann & Lunkenheimer 2013), whereas the semasiological differences between geographic and social variants of the same form remain understudied.
- This situation can be explained by more labour-intensive and subjective procedures involved in semantic comparisons and a lack of reliable *tertia comparationis*.

# Thinking globally...

- Traditionally, variational linguistics focuses on geographic and social distribution of formal variants (cf. eWAVE: Kortmann & Lunkenheimer 2013), whereas the semasiological differences between geographic and social variants of the same form remain understudied.
- This situation can be explained by more labour-intensive and subjective procedures involved in semantic comparisons and a lack of reliable *tertia comparationis*.
- We can use large-scale corpus data and distributional cues (collexemes) to pinpoint subtle semantic differences between multiple geographic variants of a construction
  - a case study of *quite* + ADJ in 20 varieties of English

# GloWbE varieties



1.9 billion words, 20 countries  
web-based corpus by Mark Davies (2013)  
<http://corpus.byu.edu/glowbe/>

# Outline

1. Theoretical background: variation of *quite*
2. Top-down approach:
  - GloWbE corpus
  - TripAdvisor reviews
3. Bottom-up approach: a clustering model of geographic variants of *quite* + ADJ
4. Conclusions and a warning

# *quite* + ADJ: main functions

- Maximizer: *All art is quite useless.* (O. Wilde)  
*≈ absolutely, completely, entirely*

(based on Quirk et al. 1985  
and Paradis 1997)

# *quite* + ADJ: main functions

- Maximizer: *All art is quite useless.* (O. Wilde)  
*≈ absolutely, completely, entirely*
- Booster: *Quite DELICIOUS! Hunger stricken upon arrival, we waddled happy full bellies back to the hotel...* (TripAdvisor)  
*≈ very, really*

(based on Quirk et al. 1985  
and Paradis 1997)



# *quite* + ADJ: main functions

- Maximizer: *All art is quite useless.* (O. Wilde)  
*≈ absolutely, completely, entirely*
- Booster: *Quite DELICIOUS! Hunger stricken upon arrival, we waddled happy full bellies back to the hotel...* (TripAdvisor)  
*≈ very, really*
- Moderator: *The paella was quite good although I have had better.* (TripAdvisor)  
*≈ fairly, pretty, rather*

(based on Quirk et al. 1985  
and Paradis 1997)

# Some geographic variation

- Maximizer: *All art is quite useless.* (O. Wilde)  
*≈ absolutely, completely, entirely*
- Booster: *Quite DELICIOUS! Hunger stricken upon arrival, we waddled happy full bellies back to the hotel...* (TripAdvisor)  
*≈ very, really*
- Moderator: *The paella was quite good although I have had better.* (TripAdvisor)  
*≈ fairly, pretty, rather*



# Some geographic variation

- Maximizer: *All art is quite useless.* (O. Wilde)

*≈ absolutely, completely, entirely*

- Booster: *Quite DELICIOUS! Hunger stricken upon arrival, we waddled happy full bellies back to the hotel...* (TripAdvisor)

*≈ very, really*



- Moderator: *The paella was quite good although I have had better.* (TripAdvisor)

*≈ fairly, pretty, rather*

# Some geographic variation

- Maximizer: *All art is quite useless.* (O. Wilde)

Exception:

*The whole experience was quite amazing.*



- Boost: *waddled happy full bellies back to the hotel...* (TripAdvisor)

*≈ very, really*



- Moderator: *The paella was quite good although I have had better.* (TripAdvisor)

*≈ fairly, pretty, rather*

# Research questions

- Do these differences really exist? And what about other geographic varieties of English?

# Research questions

- Do these differences really exist? And what about other geographic varieties of English?



top-down approach

# Research questions

- Do these differences really exist? And what about other geographic varieties of English?



top-down approach

- Are there any other differences?

# Research questions

- Do these differences really exist? And what about other geographic varieties of English?



top-down approach

- Are there any other differences?



bottom-up approach



# Outline

1. Theoretical background: variation of *quite*

2. Top-down approach:

- GloWbE corpus

- TripAdvisor reviews

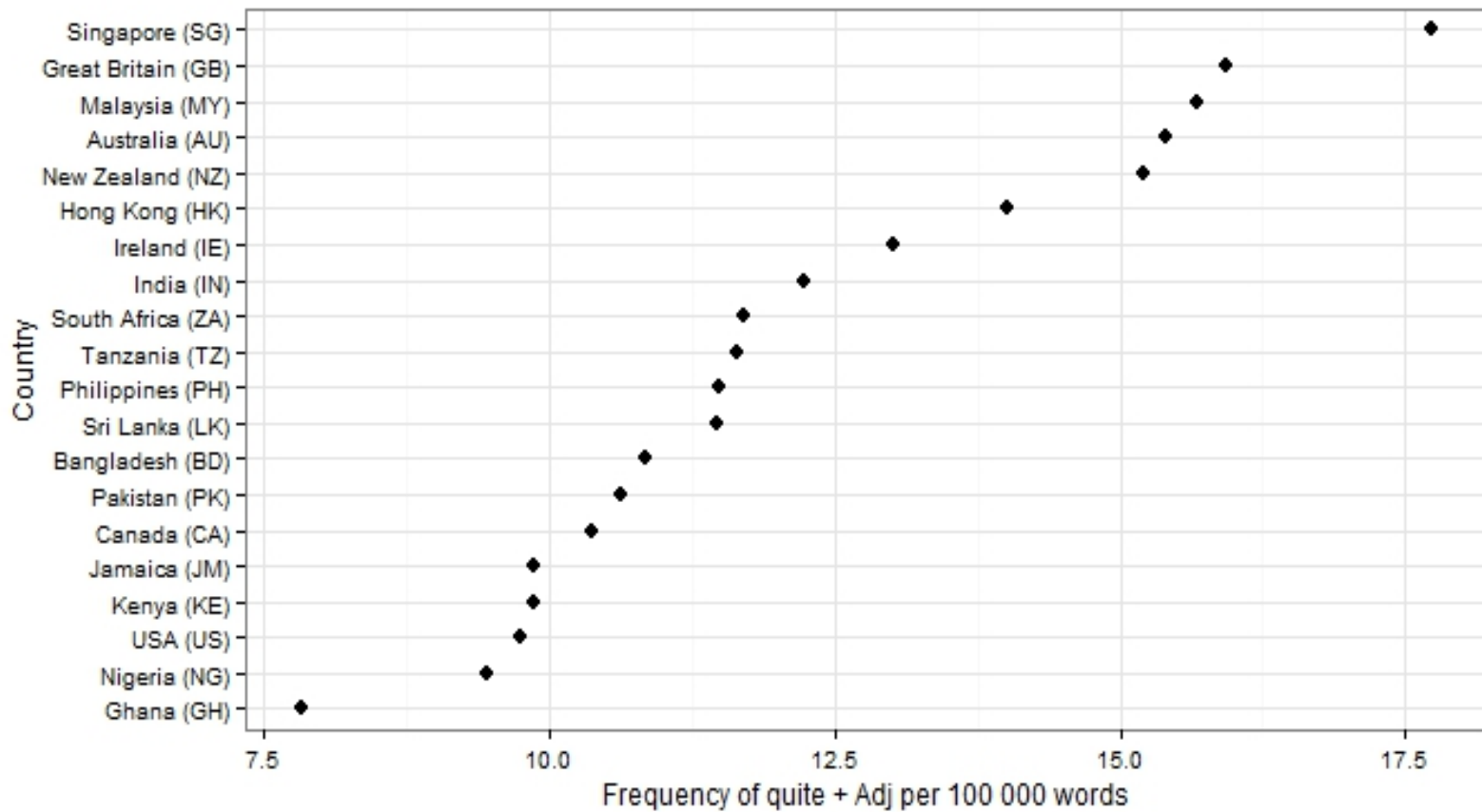
3. Bottom-up approach: a clustering model of geographic variants of *quite* + ADJ

4. Conclusions and a warning

# *quite* + ADJ GloWbE

- 237,951 instances of adverb *quite* immediately followed by an adjective in 20 varieties
- 6,096 adjectives

# Frequencies of *quite* + ADJ



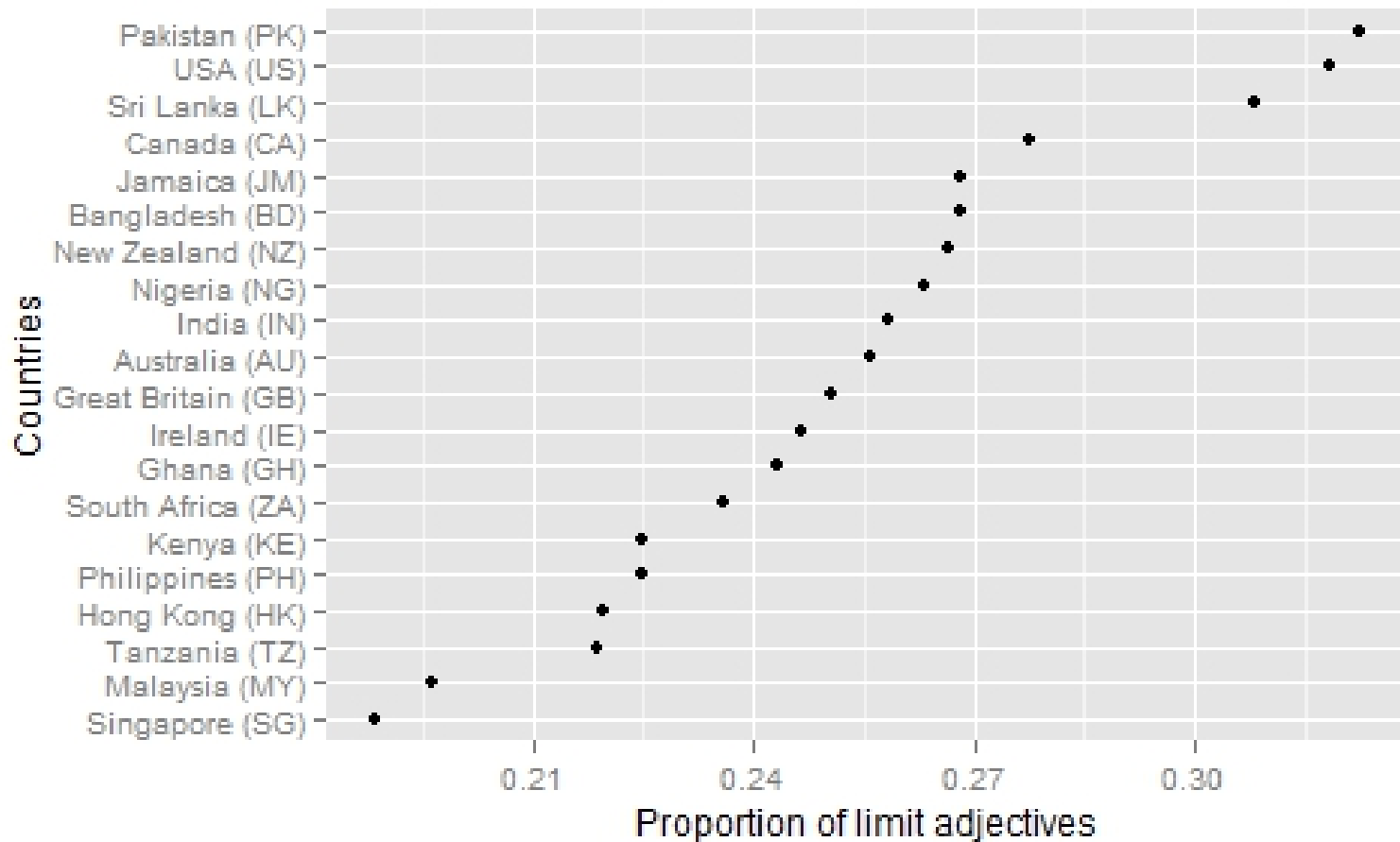
# Distributional cues

Adjectives	Examples	Function of QUITE
<b>Scalar</b>	good, nice, cheap, hard, rich, interesting	moderator (BrE) or booster (AmE)
<b>Limit</b>	different, sure, cooked, wrong, dead	maximizer
<b>Extreme</b>	huge, marvellous, extraordinary, scorching	maximizer (BrE)

(Based on Paradis 1997)

# Limit Adjectives

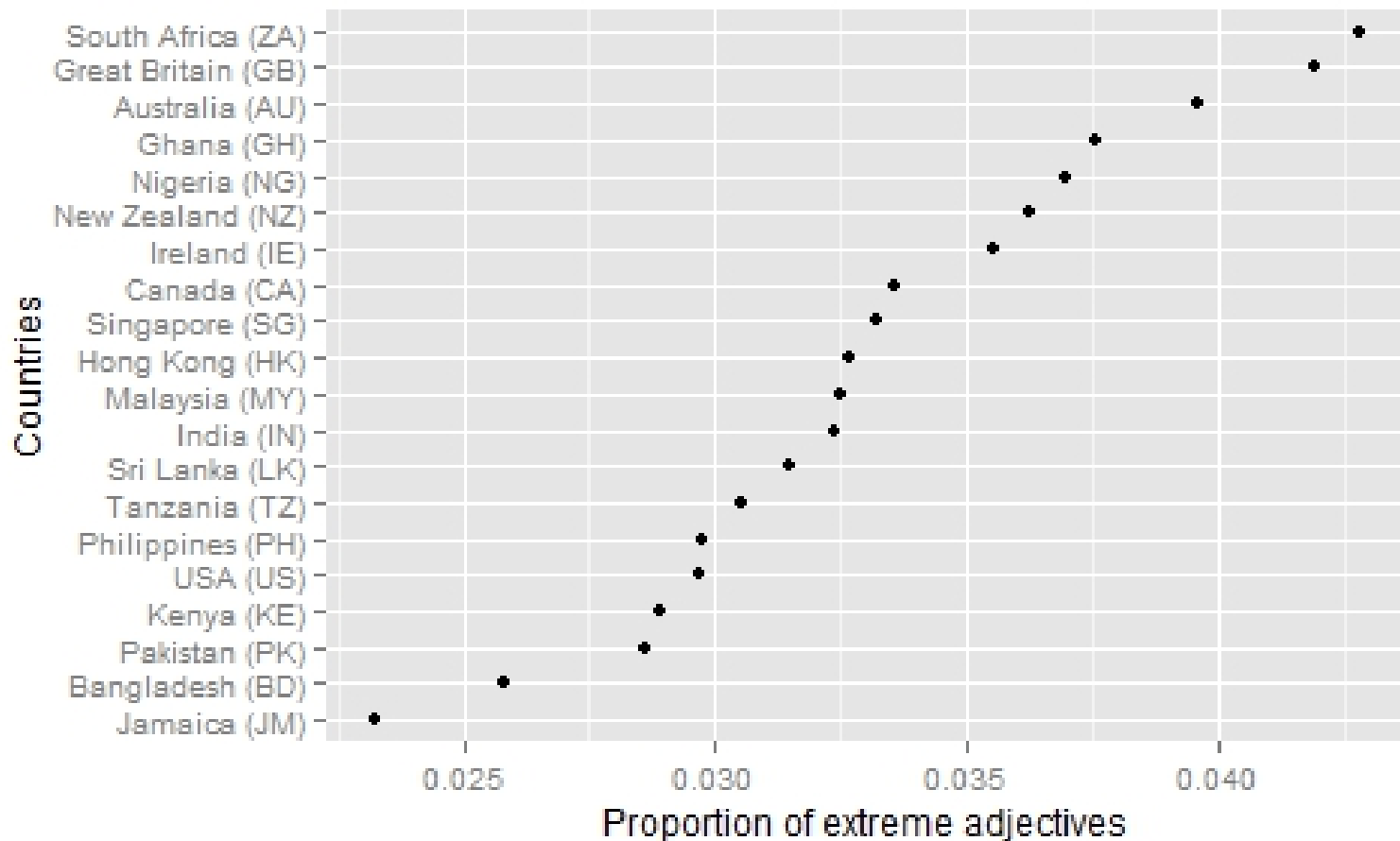
(*quite* = maximizer)



Data from GloWbE

# Extreme Adjectives

(*quite* = maximizer)



Data from GloWbE

# Outline

1. Theoretical background: variation of *quite*

2. Top-down approach:

- GloWbE corpus

- TripAdvisor reviews

3. Bottom-up approach: a clustering model of geographic variants of *quite* + ADJ

4. Conclusions and a warning

# Data from tripadvisor®

- 90 reviews of hotels, restaurants and attractions containing “quite nice”
- Ratings from 1 (terrible) to 5 (excellent)
- User profiles: “lives in” the USA, UK, Australia, Canada, India, Ireland





KDH1976  
Singapore, Singapore

**Contributor**

★ 16 reviews

🍴 4 restaurant reviews

🌐 Reviews in 9 cities

👍 8 helpful votes

*“Quite nice and good value”*

🟢🟢🟢🟢🟢 Reviewed June 2, 2014

Went here to meet my cousin and his family for lunch. A lot of food varieties (starter, salad, especially the main course) to choose from, although not much of special dishes but still mostly taste quite nice, acceptable. For chinese food, it's quite decent. Location is pretty ideal, near the Edinburgh castle, and it start business at 12PM so one can plan to have lunch after visiting the castle in the morning or before visiting the castle.

Visited May 2014

🟢🟢🟢🟢🟢 Value

🟢🟢🟢🟢🟢 Service

🟢🟢🟢🟢🟢 Atmosphere

🟢🟢🟢🟢🟢 Food

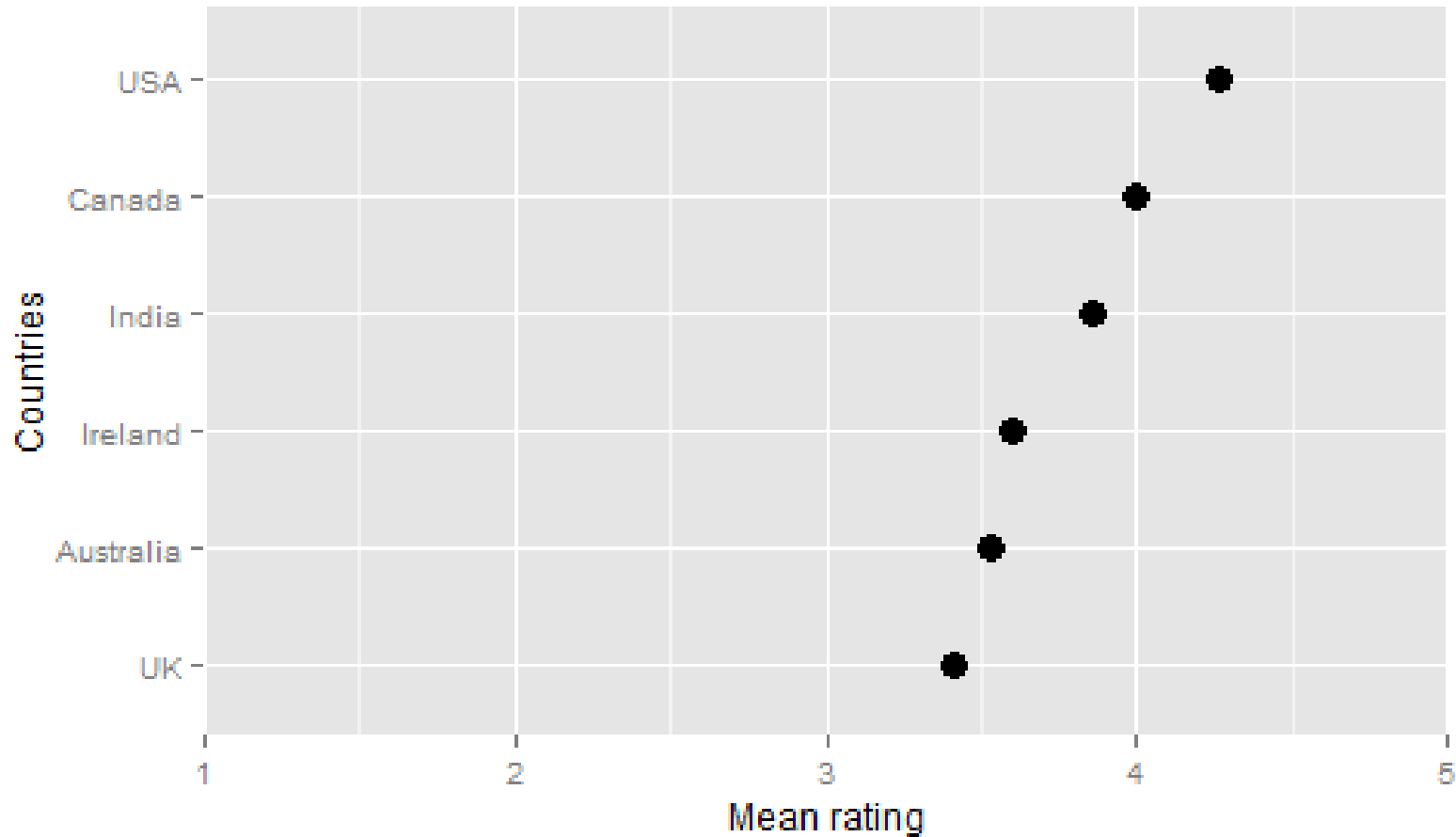
Was this review helpful?



[Ask KDH1976 about Jimmy Chung Lothian Road/Usher Hall](#)

This review is the subjective opinion of a TripAdvisor member and not of TripAdvisor LLC.

# Mean ratings



Kruskal-Wallis  $\chi^2 = 23.747$ ,  $df = 5$ ,  $p < 0.001$

# Interpretation

- American and South Asian varieties, which emerged earlier, have the highest relative frequency of limit adjectives (the maximizer function). With scalar adjectives, *quite* seems to be closer to a booster.

# Interpretation

- American and South Asian varieties, which emerged earlier, have the highest relative frequency of limit adjectives (the maximizer function). With scalar adjectives, *quite* seems to be closer to a booster.
- The moderating function and the use with extreme adjectives are the most prominent in British English and those varieties, where English was 'exported' relatively late (AU, NZ, ZA) (Trudgill & Hannah 1982).

# Interpretation

- American and South Asian varieties, which emerged earlier, have the highest relative frequency of limit adjectives (the maximizer function). With scalar adjectives, *quite* seems to be closer to a booster.
- The moderating function and the use with extreme adjectives are the most prominent in British English and those varieties, where English was ‘exported’ relatively late (AU, NZ, ZA) (Trudgill & Hannah 1982).
- This ties in with historical evidence that the maximizer function is the earliest one, occurring already in Shakespeare and Milton’s works, whereas the moderating function is attested only from the mid-19<sup>th</sup> century (Stoffel 1901; [www.etymonline.com](http://www.etymonline.com)).

# Outline

1. Theoretical background: variation of *quite*
2. Top-down approach:
  - GloWbE corpus
  - TripAdvisor reviews
3. Bottom-up approach: a clustering model of geographic variants of *quite* + ADJ
4. Conclusions and a warning

# A Semantic Vector Space model

- Step 1. Take the frequencies of adjectives that fill in the slot of *quite* + ADJ in all geographic varieties of the GloWbE.

# A Semantic Vector Space model

- Step 1. Take the frequencies of adjectives that fill in the slot of *quite* + ADJ in all geographic varieties of the GloWbE.
- Step 2. Weight the frequencies of adjectives in 20 variants of *quite* + ADJ (here: Positive Pointwise Mutual Information).
  - ➡ 20 vectors with weighted frequencies that represent the geographic variants of *quite* + ADJ



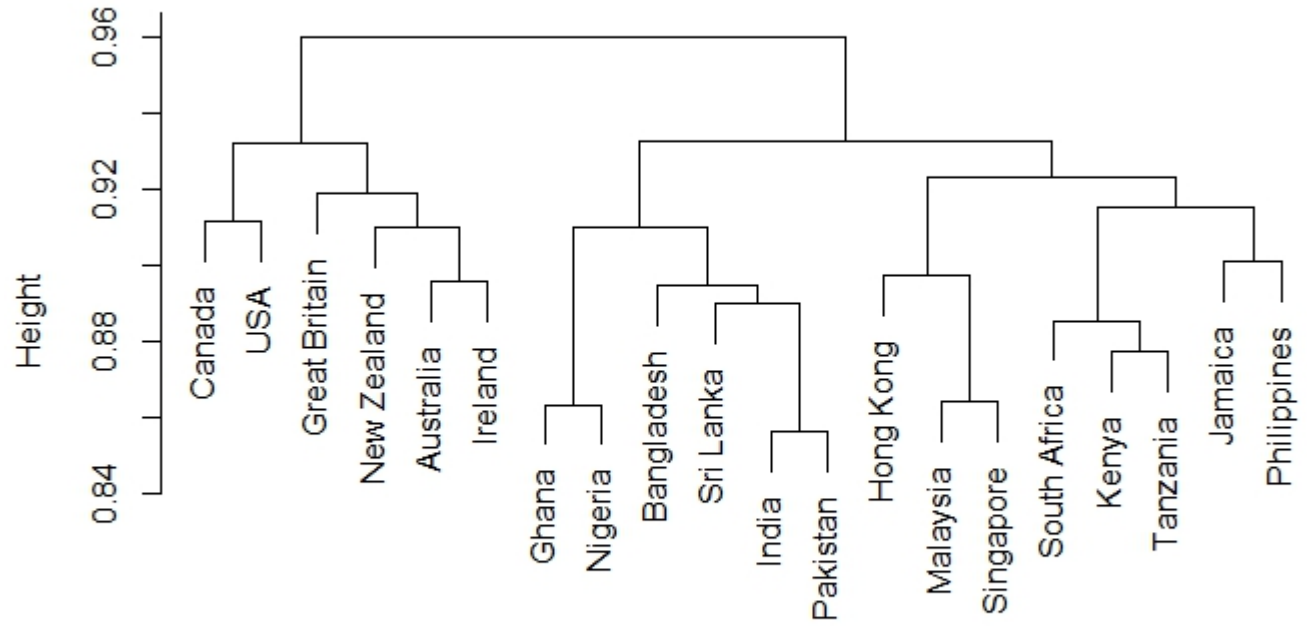
# A Semantic Vector Space model

- Step 1. Take the frequencies of adjectives that fill in the slot of *quite* + ADJ in all geographic varieties of the GloWbE.
- Step 2. Weight the frequencies of adjectives in 20 variants of *quite* + ADJ (here: Positive Pointwise Mutual Information).
  - ➡ 20 vectors with weighted frequencies that represent the geographic variants of *quite* + ADJ
- Step 3. Compute distances between all possible pairs of 20 vectors (here:  $1 - \cos(x, y)$ ). These distances represent (dis)similarities between the variants wrt. the distribution of adjectives.

# A Semantic Vector Space model

- Step 1. Take the frequencies of adjectives that fill in the slot of *quite* + ADJ in all geographic varieties of the GloWbE.
- Step 2. Weight the frequencies of adjectives in 20 variants of *quite* + ADJ (here: Positive Pointwise Mutual Information).
  - ➡ 20 vectors with weighted frequencies that represent the geographic variants of *quite* + ADJ
- Step 3. Compute distances between all possible pairs of 20 vectors (here:  $1 - \cos(x, y)$ ). These distances represent (dis)similarities between the variants wrt. the distribution of adjectives.
- Step 4. Use a clustering algorithm (here: agglomerative hierarchical clustering).

### Cluster Dendrogram



quite.dist  
hclust (\*, "complete")

# Top 10 most distinctive collexemes in 6 clusters

L1 varieties	West Africa	South Asia	South East Asia	South and East Africa	Jamaica and Philippines
<i>right</i>	<i>unfortunate</i>	<i>evident</i>	<i>nice</i>	<i>chilly</i>	<i>popular</i>
<i>happy</i>	<i>reserved</i>	<i>natural</i>	<i>packed</i>	<i>high</i>	<i>expensive</i>
<i>sure</i>	<i>interesting</i>	<i>popular</i>	<i>expensive</i>	<i>hot</i>	<i>affordable</i>
<i>wrong</i>	<i>intriguing</i>	<i>common</i>	<i>disappointed</i>	<i>encouraging</i>	<i>fine</i>
<i>correct</i>	<i>commendable</i>	<i>simple</i>	<i>fun</i>	<i>steep</i>	<i>erect</i>
<i>different</i>	<i>presumptuous</i>	<i>opposite</i>	<i>good</i>	<i>taxing</i>	<i>commendable</i>
<i>extraordinary</i>	<i>challenging</i>	<i>famous</i>	<i>pricey</i>	<i>cool</i>	<i>torque</i>
<i>separate</i>	<i>understandable</i>	<i>impossible</i>	<i>surprised</i>	<i>malnourished</i>	<i>well-liked</i>
<i>ready</i>	<i>grave</i>	<i>obvious</i>	<i>spacious</i>	<i>impressive</i>	<i>young</i>
<i>likely</i>	<i>encouraging</i>	<i>confident</i>	<i>troublesome</i>	<i>rocky</i>	<i>challenging</i>

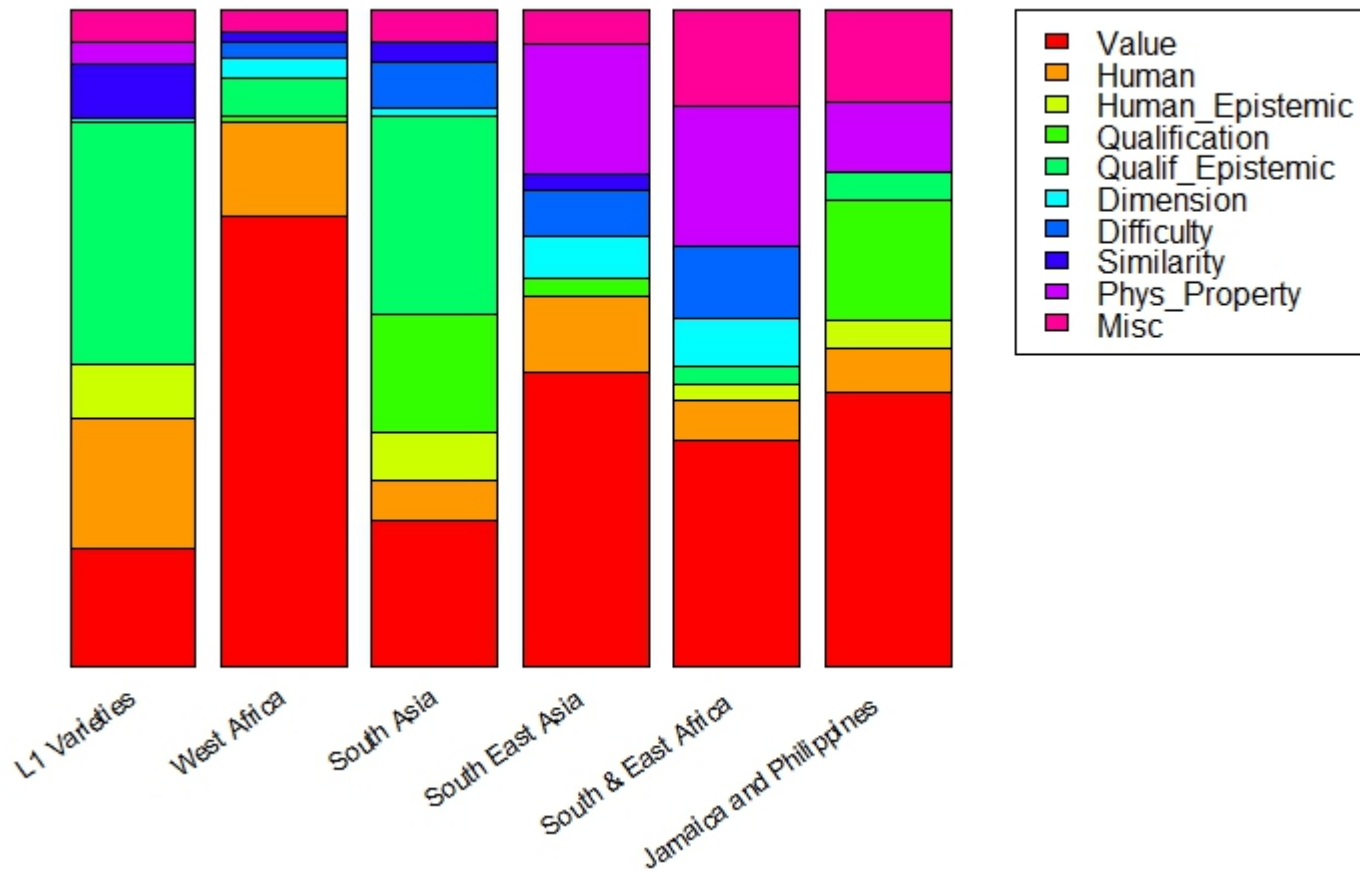
All  $p < 0.001$   
 see Gries & Stefanowitsch (2004)

# Adjective classes

Class	Examples
Value	<i>good, bad, lovely, perfect, odd, lucky</i>
Human Propensity	<i>jealous, happy, kind, clever, eager</i>
Human Epistemic	<i>aware, sure, ignorant</i>
Qualification	<i>usual, normal, appropriate</i>
Qualification Epistemic	<i>true, right, evident, obvious</i>
Dimension	<i>small, large, tall, wide, deep</i>
Difficulty	<i>easy, difficult, tough, simple</i>
Similarity	<i>similar, dissimilar, identical, other</i>
Physical Property	<i>wet, rough, clean, hot, dead, sick</i>
Miscellaneous	<i>young, slow, far, blue</i>

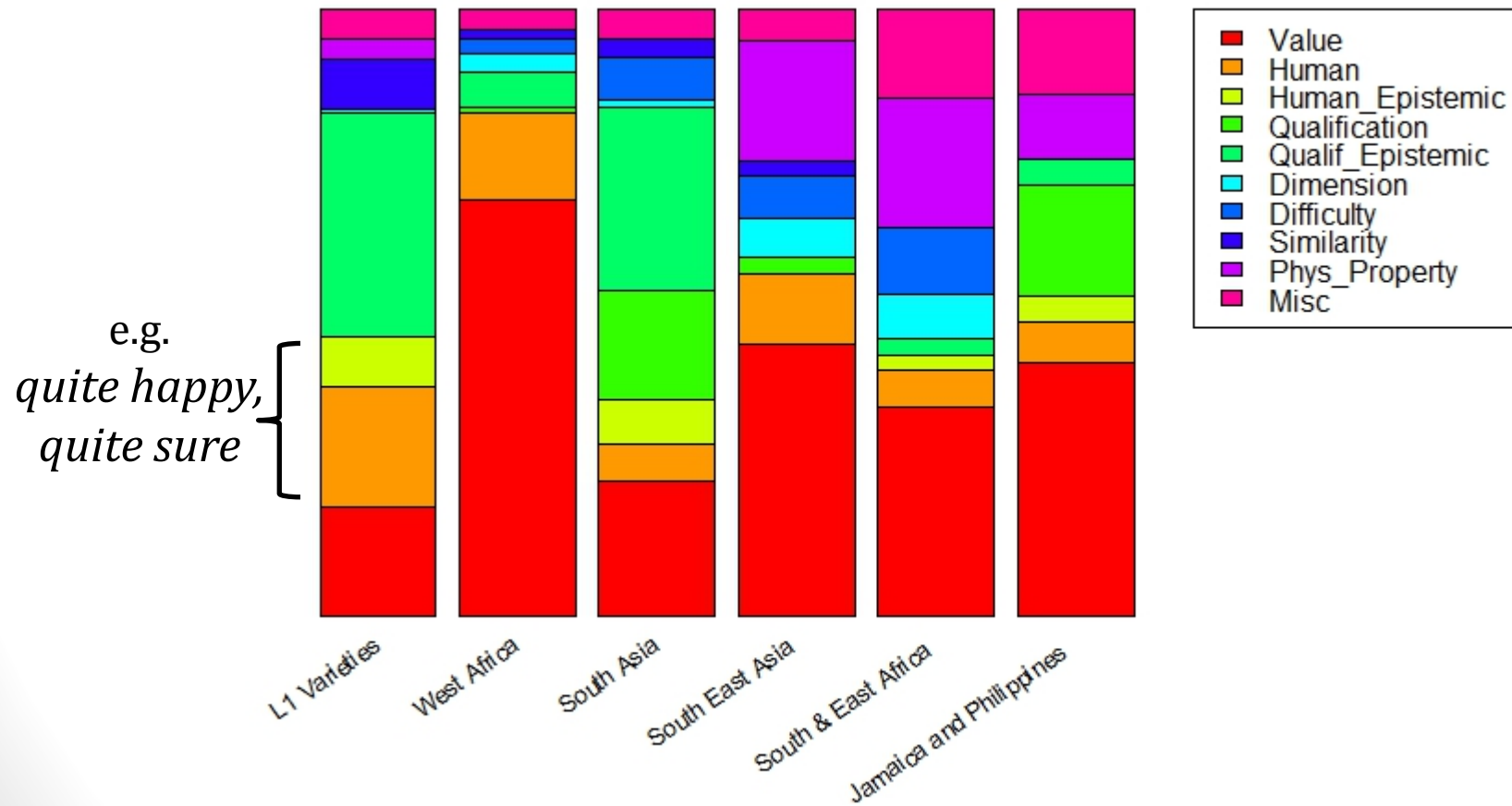
Based on Dixon (2004), with some modifications

# Semantic classes of distinctive collexemes ( $p < 0.001$ )



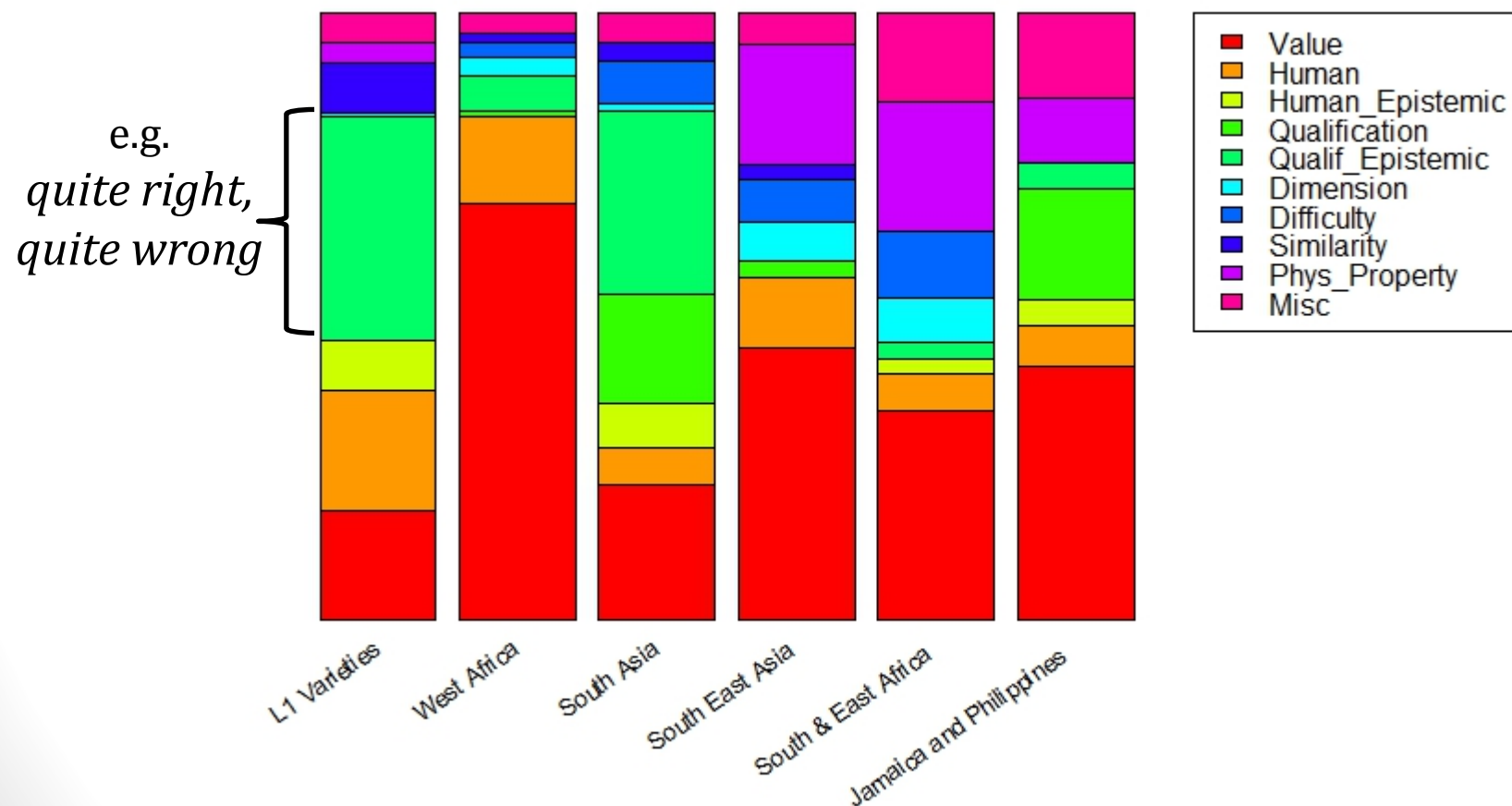
Adjective classes are represented by the sums of the distinctiveness scores of the adjectives that belong to each class.

# Human Propensity Classes



Adjective classes are represented by the sums of the distinctiveness scores of the adjectives that belong to each class.

# Epistemic Qualification Class



Adjective classes are represented by the sums of the distinctiveness scores of the adjectives that belong to each class.



# Interpretation

- Human Propensity classes: the use of English in L1 varieties may be more subjective and intersubjective than in other Englishes, which are used primarily in public and impersonal domains.
- Qualification Epistemic class: A closer look at the contexts suggests that the construction is often used to express agreement (*You're quite right*) or soften the disagreement (*I don't think you're quite right*) with the interlocutor. Again, such uses are intersubjective.

# Outline

1. Theoretical background: variation of *quite*
2. Top-down approach:
  - GloWbE corpus
  - TripAdvisor reviews
3. Bottom-up approach: a clustering model of geographic variants of *quite* + ADJ
4. Conclusions and a warning

# Conclusions

- The geographic variation of *quite* reflects the historical development of its functions from that of a maximizer to the function of a moderator, as well as the timeline of the expansion of English all over the world.

# Conclusions

- The geographic variation of *quite* reflects the historical development of its functions from that of a maximizer to the function of a moderator, as well as the timeline of the expansion of English all over the world.
- However, the geographic differences are not restricted to the ones explained by the historical factor. The use of *quite* seems to be influenced by cultural constraints, such as the domain of English use.

# Conclusions

- The geographic variation of *quite* reflects the historical development of its functions from that of a maximizer to the function of a moderator, as well as the timeline of the expansion of English all over the world.
- However, the geographic differences are not restricted to the ones explained by the historical factor. The use of *quite* seems to be influenced by cultural constraints, such as the domain of English use.
- Further research is needed to pin down more subtle differences and filter out potential corpus noise.

# A warning

10 Things That Americans Don't Realize are Offensive to Brits

...

## **Describing something as “quite” good**

Using “quite” to mean “very” is an English language anachronism that's thriving in the U.S. A British friend recently told me about the time his American partner joined him for a family Christmas back home. Thinking she was offering up high praise, she managed to upset her U.K. hosts by describing dinner as “quite good.” To their ears, “quite” in this context detracted from the “good”: the meal was passable but could have been better.

<http://www.bbcamerica.com/mind-the-gap/2013/01/29/10-things-americans-dont-realize-are-offensive-to-brits/>

# Thank you!

 Quite British

A rather spiffing  
webcomic

