Individual differences inform the syntax-processing island debate

Dan Michel
University of California San Diego
Islands in Contemporary Linguistic Theory
November 17, Faculty of Letters (Vitoria-Gasteiz)
University of the Basque Country

Working Memory and Linguistic Data

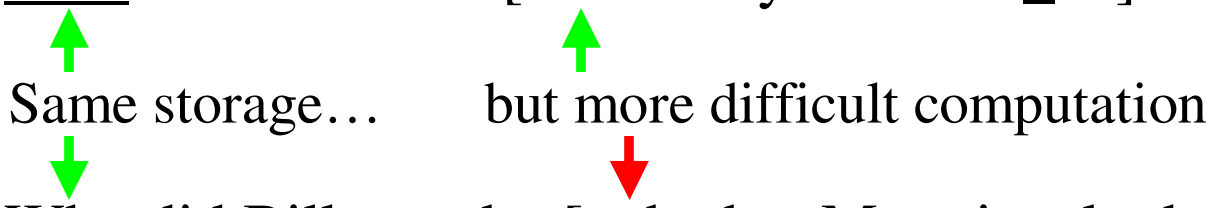
- Center Embedding
  - The dog the rat the cat ate scared ran.
  - Too much for the memory system to handle
  - Grammatical but unacceptable
  - "Languages are not designed for parsability" (Chomsky, 1991)
- Can this approach be used for island phenomena?

Outline

- Working Memory (WM) applied to islands (Kluender and Kutas, 1993)
- Testing WM with individual differences (Sprouse, Wagers & Phillips, 2011)
- A different view of WM in sentence processing (e.g. Lewis and Vasishth, 2005)
- Testing this view of WM with individual differences- a new experiment

Processing Accounts

- Constrained capacity accounts
  - Excessive strain on working memory results in ungrammaticality (e.g. Kluender, 1991, 1998; Kluender and Kutas, 1993b; Hofmeister, 2008; Sag et. al., 2007)
- Working memory
  - There is a single pool of resources used for both **storage** and **computation** (Just and Carpenter, 1992)
- *Whether* islands
  - Who did Bill think [that Mary insulted _ ?]



  - Who did Bill wonder [whether Mary insulted _ ?]

Constrained Capacity Applied to Islands

- Working memory is a capacity for storage and computation.
- Individuals with higher capacity should thus be able to process islands better.
- Do we see co-variation of individual working memory measures and acceptability ratings for islands?

Outline

- Working Memory (WM) applied to islands (Kluender and Kutas, 1993)
- ► Testing WM with individual differences (Sprouse, Wagers & Phillips, 2011)
- A different view of WM in sentence processing (e.g. Lewis and Vasishth, 2005)
- Testing this view of WM with individual differences- a new experiment

Sprouse, Wagers & Phillips

- An island 'violation' needs an 'interaction' of two phenomena:
  - extraction out of
  - a 'certain structure' (such as a *whether* phrase)

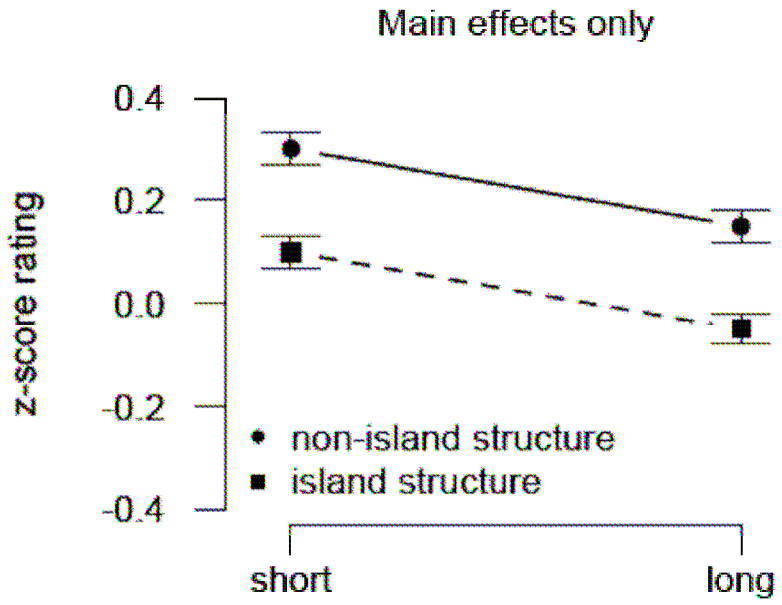
Short, Non: Who _ thinks that John bought a car?
Short, Island: Who _ wonders whether John bought a car?
Long, Non: What do you think that John bought _?
Long, Island: What do you wonder whether John bought _ ?

- Difference in Difference score
- Measure of the size of the interaction:

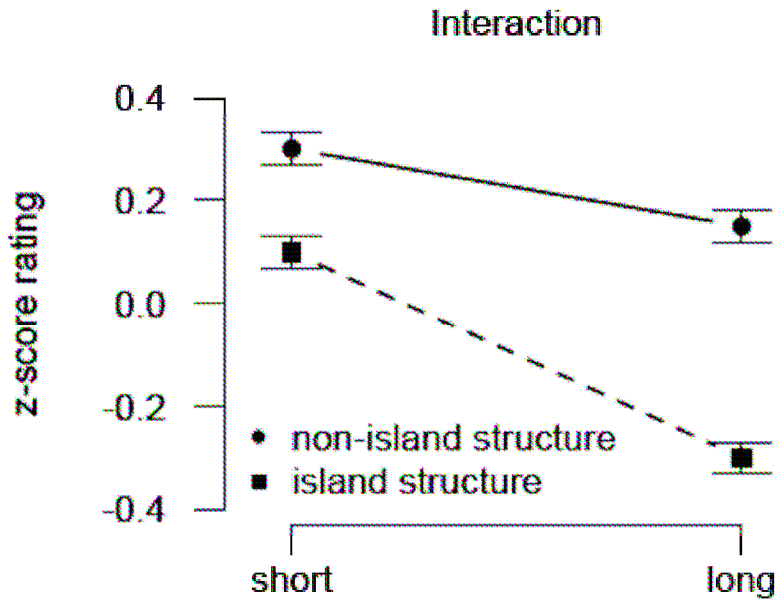
DD = [(long, non island) - (long, island)]

- [(short, non-island) - (long, island)]

a.



b.



Cognitive Measures

- Serial recall
- N-back

Serial Recall

- Subjects presented with a series of words (*bagel, humor, level, magic, novel, topic, tulip, woman*)
- 10 lists with different orders
- Tasked with recalling them in order
- Recall-order measure

N-back

L L K L Z K

- Task: hold targets in memory, continuously update those targets, compare them to a probe.
- Each instance of the n-back (1-, 2-, 3-, and 4-back) is conducted in a separate block.
- Recognition-updating measure

Results

- Comparing the DD score with the WM measures…
- SWP find no robust effects
- Conclude that the constrained capacity account of islands is not supported.

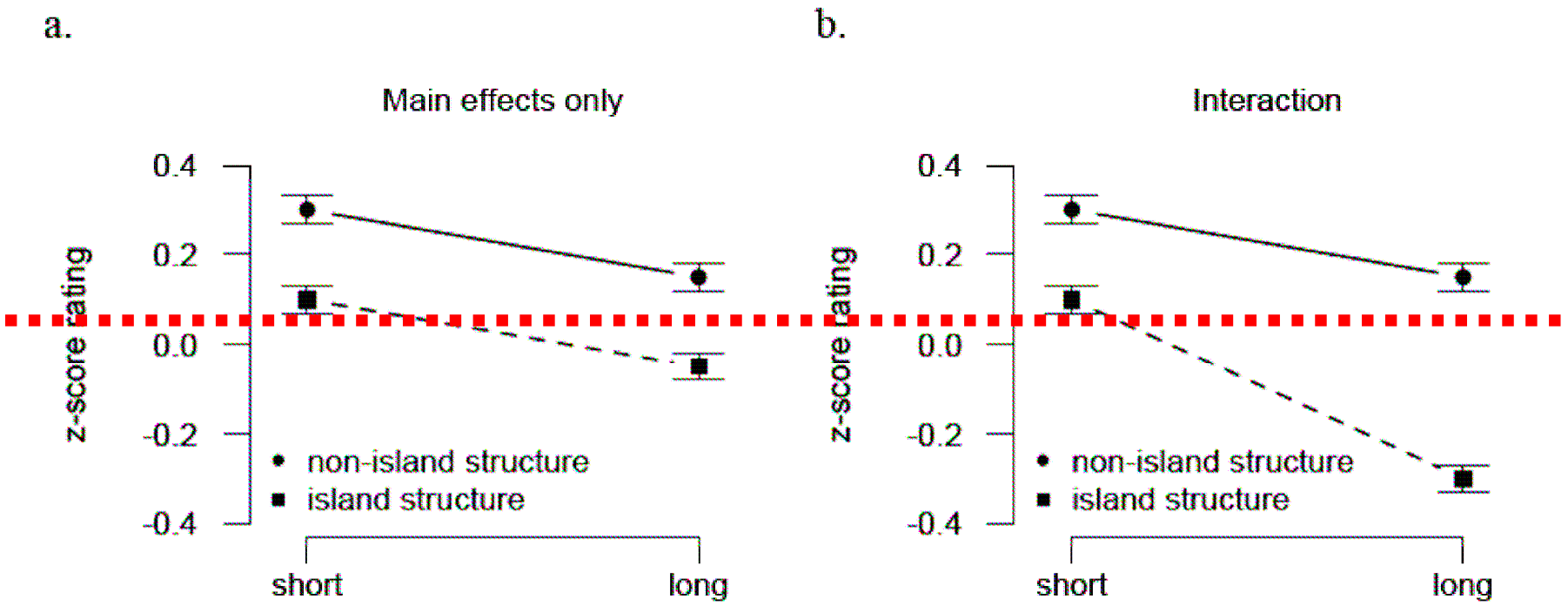
Issues

1) Compared only the statistical interaction (DD score) to the WM tasks
   - It may be possible that a *statistical* interaction is not the defining characteristic of an island effect (Ross, 1987)
2) Null effects are difficult to interpret
   - Could be a lack of power issue
3) Measured WM with two different tasks, but neither included interference

Addressing the issues

1) Compared only the statistical interaction (DD score) to the WM tasks

- Why this is an issue:
  - It is unclear that a *statistical* interaction is required



Addressing the issues

2) Null effects are difficult to interpret

- Why this is an issue:
  - SWP (following Kluender and Kutas, 1993) state that the effects of processing **extractions** and **island structures** tax WM
    - This can't be shown in a DD score
    - Perhaps WM effects can't be observed in acceptability judgments
- The current experiment will examine multiple dimensions (effects of **extraction**, **island structure** and the **interaction** between them)

Addressing the issues

3) Measured WM with two different tasks, but neither included interference

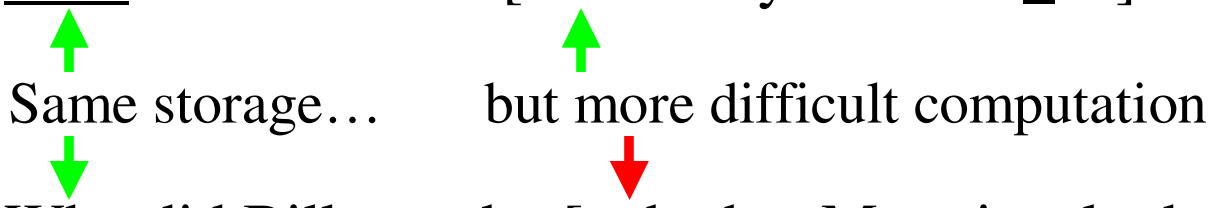
- Why this is an issue:
  - Just as there is more than one 'grammatical' or 'syntactic' approach to islands, there can be more than one processing/ WM-based approach
    - Kluender and Kutas (1993) is based on a constrained capacity view of WM
    - Another major view of WM is in terms of **similarity-based interference**
- The current experiment will include individual measures that include **memory-interference**

Outline

- Working Memory (WM) applied to islands (Kluender and Kutas, 1993)
- Testing WM with individual differences (Sprouse, Wagers & Phillips, 2011)
- ► A different view of WM in sentence processing (e.g. Lewis and Vasishth, 2005)
- Testing this view of WM with individual differences- a new experiment

Processing Accounts

- Constrained capacity accounts
  - Excessive strain on working memory results in ungrammaticality (e.g. Kluender, 1991, 1998; Kluender and Kutas, 1993b; Hofmeister, 2008; Sag et. al., 2007)
- Working memory
  - There is a single pool of resources used for both **storage** and **computation** (Just and Carpenter, 1992)
- *Whether* islands
  - Who did Bill think [that Mary insulted _ ?]



  - Who did Bill wonder [whether Mary insulted _ ?]

Processing Accounts

- Similarity-based interference accounts
  - Excessive strain on working memory results in ungrammaticality (e.g. Gordon, Hendrik and Johnson, 2001; Gordon, Hendrick and Levine, 2002; Lewis and Vasishth, 2005; Lewis, Vasishth and Van Dyke, 2006; Van Dyke and McElree, 2006)
- Working memory
  - Words have **features** and are retrieved from recent memory based on those features. If features in recent memory are confusable/overlap **similarity-based interference** occurs.
- *Whether* islands
  - Who did Bill think [that Mary insulted _ ?]

Nothing special about this storage… but more difficult retrieval due to interference

  - **Who** did Bill wonder [**whether** Mary insulted _ ?]

Measurable differences

- Similarity-based interference accounts invoke a notion of **similarity-based interference** while constrained capacity accounts do not.
  - This can be discerned by using a WM task that includes similarity-based interference with one that does not
- Constrained capacity accounts invoke a notion of **active memory cost** while similarity-based interference accounts do not.
  - The verbal span task encourages this strategy

Outline

- Working Memory (WM) applied to islands (Kluender and Kutas, 1993)
- Testing WM with individual differences (Sprouse, Wagers & Phillips, 2011)
- A different view of WM in sentence processing (e.g. Lewis and Vasishth, 2005)
- ► Testing this view of WM with individual differences- a new experiment

Whether island- acceptability judgments

Methods:

- 81 subjects (45 female), age 17 to 27 (avg 20.4)
- 4 WM/attention tasks
- 1-7 Likert scale
- Paper and pencil
- Manipulated extraction (long, short) x structure (island, non-island)
- 8 items per condition (32 total, 168 fillers)
- Latin square design, plus reverse order (8 lists)

Improvements upon SWP

- Analyzing all factors/conditions; not just the interaction (i.e. DD score)
- Additional individual difference measures
- Materials are designed to extend to future self-paced reading and ERP experiments (sentence position controls)

Types of Working Memory Measures

- N-back
  - Recognition, updating, interleaved dependencies
  - General (not specifically verbal) working memory capacity
- Verbal span
  - Recall, ‘natural’ language processing
  - Constrained capacity
- Memory-interference
  - Recognition, semantic and phonological lures
  - Similarity-based interference
- Flanker
  - Selective attention, no memory component
  - ‘Pure’ attention

Cognitive Measures

- N-back

L L K L Z K

  - Task: hold targets in memory, continuously update those targets, compare them to a probe.
  - Eight letters one at a time (500 ms fixation 1000 ms presentation)
  - Each instance of the n-back (1-, 2-, 3-, and 4-back) is conducted in a separate block of 30 items (10 correct positive responses) with its own instructions.

Cognitive Measures

- Verbal Span

  John saw Mary at the train station.

  The trees have grown since last spring.

  The policeman found the suspect under the bridge.

  - Developed to specifically target verbal working memory capacity (Just and Carpenter, 1992).
  - Participants read a set of sentences aloud and then attempt to recall the last word of each sentence at the end of the set.
  - Sets range in number of sentences from three to five. Participants read five examples of each set.

Cognitive Measures

- Memory-interference
  - Task: recognize an item from memory in the presence of either a phonological or semantic lure.

Study:	Test:	
jaguar	cheetah	(semantic lure)
grass	glass	(phonological lure)
snow	boat	(false)
pineapple	pineapple	(true)

  - 3 blocks. 10 study words per block, 10 test words (all 500 ms fixation, 1500 ms presentation). Total 5 each lures, false conditions.

Cognitive Measures

- Flanker attention task (Eriksen and Schulltz, 1979)
  - assess reaction time and a participant's ability to screen out competing information (selective attention).

congruent

incongruent

- Response to congruent stimulus = reaction time
- Incongruent stimuli - congruent stimuli = ability to suppress competing information (lower number indicating better)
- Participants were presented a total of 32 trials (500 ms fixation, 1000 ms presentation)

Whether island

1) Non-island- short extraction Who had _ on Tuesday assumed [that the decorator annoyed the carpenter …
2) Island- short extraction Who had _ on Tuesday inquired [whether the decorator annoyed the carpenter …
3) Non-Island- long extraction Who had the carpenter assumed [that the decorator annoyed _ on Tuesday …
4) Island- long extraction Who had the carpenter inquired [whether the decorator annoyed _ on Tuesday …
… when the deadline was missed?]

Whether island

Adjunct insertion – no main effect or interactions

1) Non-island- short extraction Who had _ **<u>on Tuesday</u>** assumed [that the decorator annoyed the carpenter …
2) Island- short extraction Who had _ **<u>on Tuesday</u>** inquired [whether the decorator annoyed the carpenter …
3) Non-Island- long extraction Who had the carpenter assumed [that the decorator annoyed _ **<u>on Tuesday</u>** …
4) Island- long extraction Who had the carpenter inquired [whether the decorator annoyed _ **<u>on Tuesday</u>** …
… when the deadline was missed?]

Whether island predictions- acceptability judgments

1) Non-island- short extraction Who had _ on Tuesday assumed [that the decorator annoyed the carpenter …
2) Island- short extraction Who had _ on Tuesday inquired [whether the decorator annoyed the carpenter …
3) Non-Island- long extraction Who had the carpenter assumed [that the decorator annoyed _ on Tuesday …
4) Island- long extraction Who had the carpenter inquired [whether the decorator annoyed _ on Tuesday …
… when the deadline was missed?]

Whether island predictions- acceptability judgments

1. Extraction: short > long

1) Non-island- **short extraction** Who had _ on Tuesday assumed [that the decorator annoyed the carpenter …
2) Island- **short extraction** Who had _ on Tuesday inquired [whether the decorator annoyed the carpenter …
3) Non-Island- long extraction Who had the carpenter assumed [that the decorator annoyed _ on Tuesday …
4) Island- long extraction Who had the carpenter inquired [whether the decorator annoyed _ on Tuesday …
(e.g. King and Kutas, 1995; Fiebach, Schlesewky and Friederici, 2002; Phillips, Kazanina and Abada, 2005, among others) … when the deadline was missed?]

Whether island predictions- acceptability judgments

2. Island hood: Non-Island > Island

1) <u>Non-island</u>- short extraction Who had _ on Tuesday assumed [that the decorator annoyed the carpenter …
2) Island- short extraction Who had _ on Tuesday inquired [whether the decorator annoyed the carpenter …
3) <u>Non-Island</u>- long extraction Who had the carpenter assumed [that the decorator annoyed _ on Tuesday …
4) Island- long extraction Who had the carpenter inquired [whether the decorator annoyed _ on Tuesday …
… when the deadline was missed?] (e.g. Kluender and Kutas, 1993b)

Whether island predictions- acceptability judgments

3. Island violation: (4) deemed least acceptable

1) Non-island- short extraction Who had _ on Tuesday assumed [that the decorator annoyed the carpenter …
2) Island- short extraction Who had _ on Tuesday inquired [whether the decorator annoyed the carpenter …
3) Non-Island- long extraction Who had the carpenter assumed [that the decorator annoyed _ on Tuesday …
4) <u>Island- long extraction</u> Who had the carpenter inquired [whether the decorator annoyed _ on Tuesday …
… when the deadline was missed?]

Whether island results

- Predictions 1-3 were borne out:

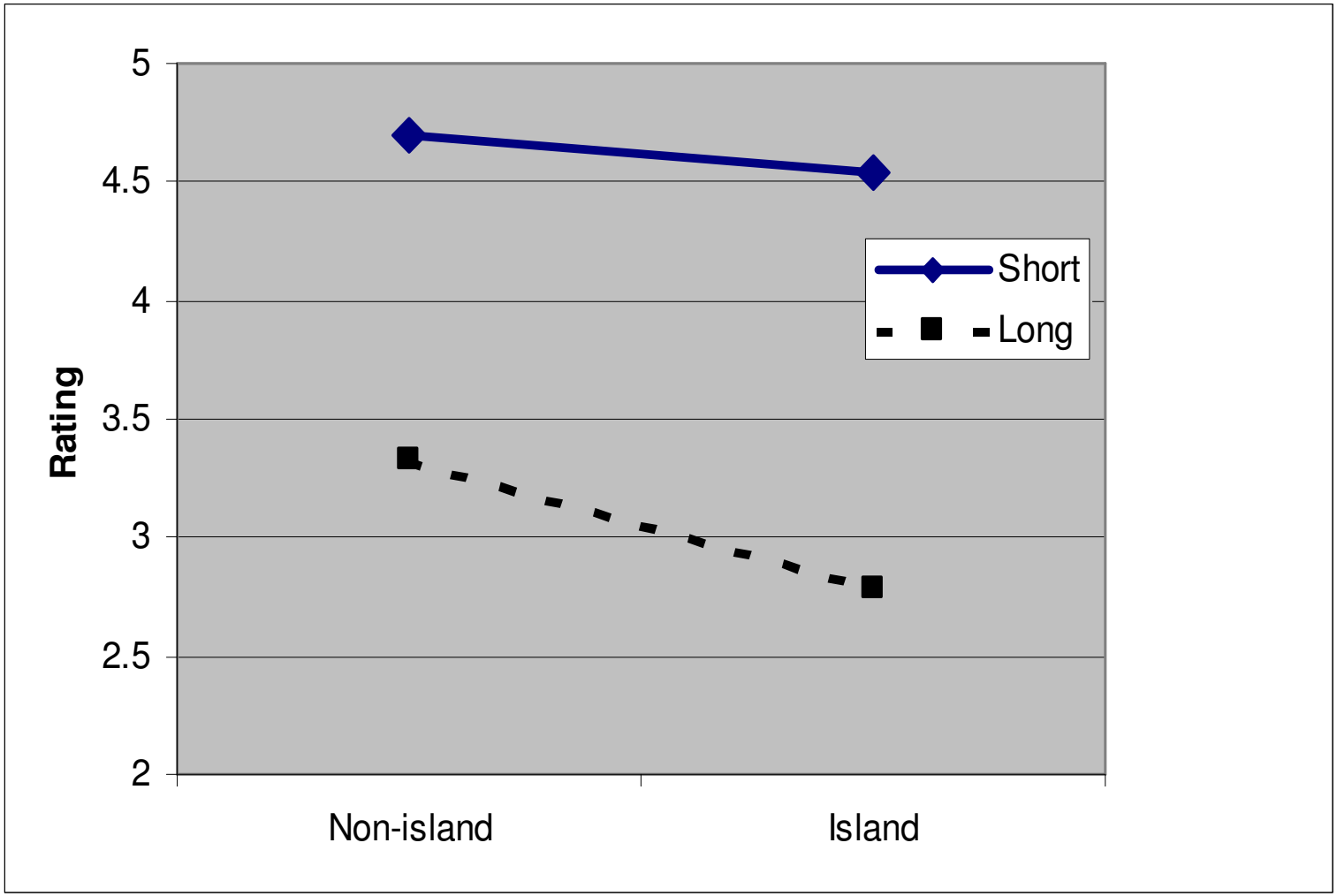
  1) Short extraction was rated as more acceptable than long extraction

  2) Non-islands were rated as more acceptable than islands

  3) The long extraction out of an island condition was rated lowest

    - Did not show a clear statistical interaction
    - $p_1 = 0.14$, $p_2 < 0.01$
    - Supporting a threshold view over a statistical interaction view of island unacceptability?

Whether island results



Results are presented here on the original 7-point scale, but analyses were also run on z-scores with no difference in the pattern of results.

Whether-island and WM Results

- WM analysis: multiple ANOVAs each comparing three factors:
  - 'extraction distance' (long and short) x
  - 'island condition' (island and non-island) x
  - 'cognitive measure' (high and low)
- ANOVAs run separately for each cognitive measure.
- Cognitive measure groups were formed by median split.

Whether island and WM results

- No effects for the interaction (i.e. DD-type measure)
- No effects for the island structure vs. non-island structure conditions
- Only effects for distance manipulation

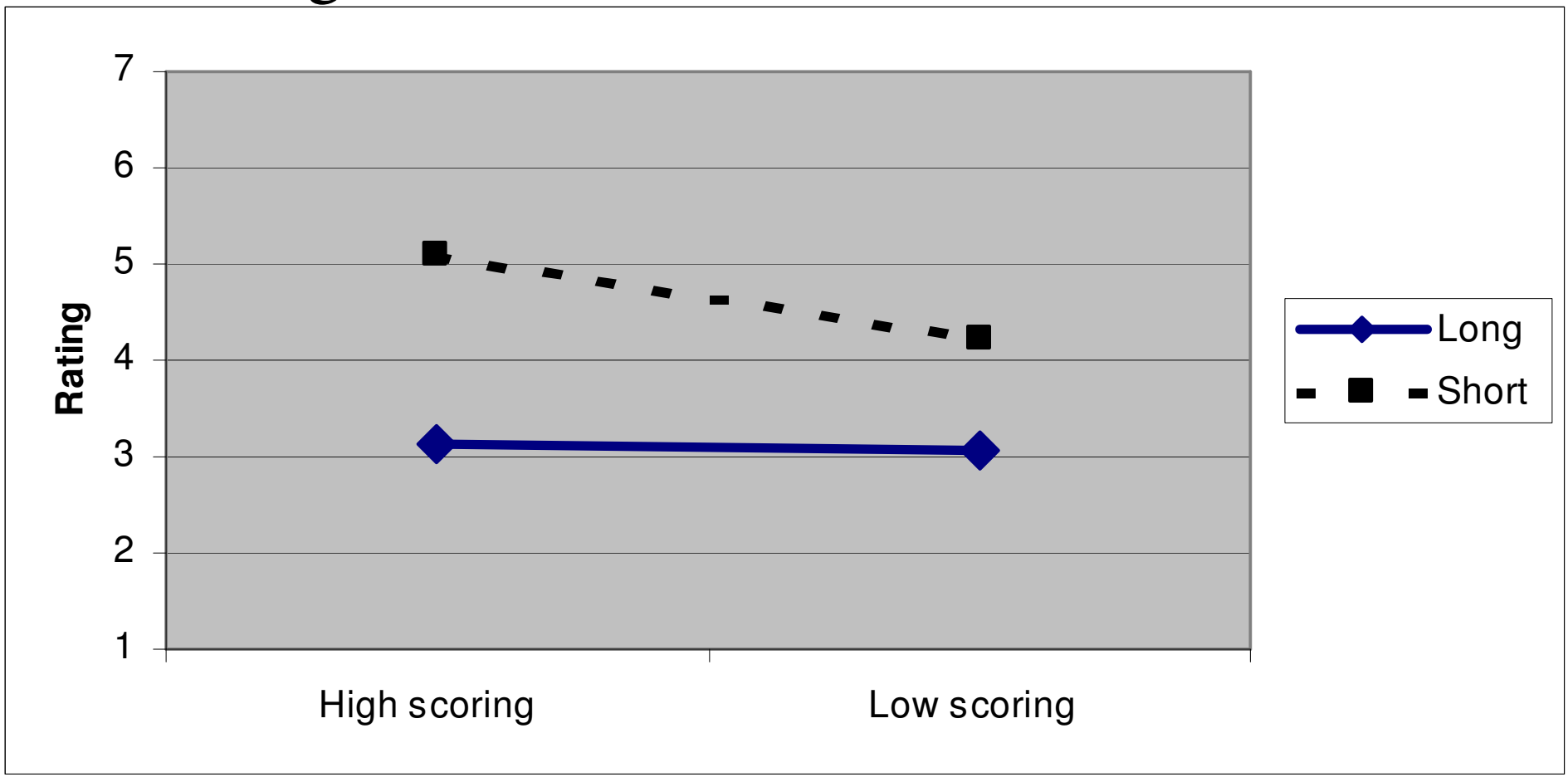
- SWP replicated (constrained capacity account for islands not supported)
- Additionally, similarity-based interference account for islands not supported

Whether island and WM results

- Only the **memory-interference** task showed a significant main effect and interaction with a linguistic manipulation.
- Driven by the **phonological/ orthographical lures**:
- Main effect of score on acceptability rating ($p_1 = 0.002$, $p_2 = 0.01$)
  - high scorers (4.11 Mean, 1.41 SD)
  - lower scorers (3.64 Mean, 1.15 SD)
- Interaction of recognition-interference with extraction distance ($p_1 = 0.01$, $p_2 = 0.03$)

	High scoring	Low scoring
Long extraction	3.12 (1.022)	3.058 (0.928)
Short extraction	5.103 (0.985)	4.227 (1.053)

Phonological-interference and Extraction Distance



Low score indicates susceptibility of phon/orth interference

1) High scorers do not rate ‘difficult to process’ sentences any higher

2) High scores may be better on the ‘easier’ sentences

(alternatively, the low scorers have trouble on both).

Difficulties for low scorers

- The lack of boost for the high scorers in the difficult condition, but a penalty for the low scorers in the easier condition has been reported previously in the ERP literature
  (King and Kutas, 1995; Münte, Schiltz and Kutas, 1998; Fiebach, et al, 2002)

	Sentence Regions			
Sentence Type	*Pre-relative Clause*	*Early Relative Clause*	*Late Relative Clause*	*Main Verb Phase*
SO	The reporter who	the senator	harshly attacked	admitted the error
SS	The reporter who	harshly attacked	the senator	admitted the error

King and Kutas (1995): Positive drift associated with ease of processing

Good comprehenders show more positive drift to easy condition.

Poor comphrehenders appear to be processing both sentences with difficulty

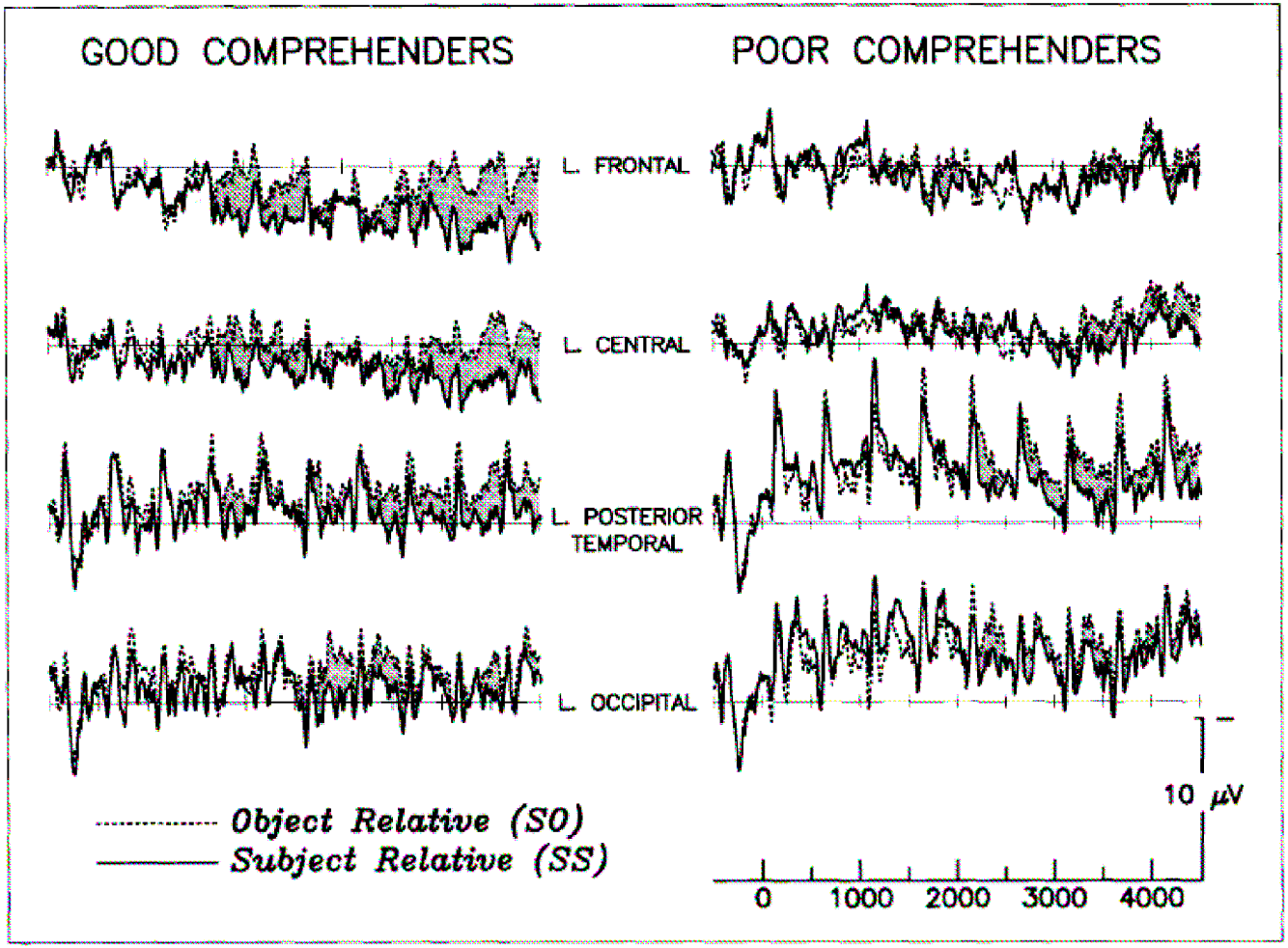


Figure 2. Average ERPs from representative left Hemisphere Sites from Good (n = 12) and Poor (n = 12) Comprehenders for Object Relative (SO) and Subject Relative (SS) sentences. Waveforms are aligned on the first word of each sentence type and include the response to words up to and including the final word of the main verb phrase.

Possible Implication(s)

- High scorers are not differentiating themselves from low scorers in the difficult conditions
  - Are the difficult conditions difficult for *everyone*? i.e. beyond the range of possible human WM ability?
    - (no humans are expected to be able to parse 4 levels of center embedding)
  - Is the limit on the difficult sentences grammatical rather than processing related?
    - We see little difference between groups in the difficult sentences
    - But we do see differences in the easier sentences which are not proposed to be limited by the grammar

Summary of Findings

1) In comparing four different cognitive measures of individual variation, none of them were found to co-vary with the ratings of the island-violation effect

2) Only one of these measures, the memory-recall task (specifically phonological), co-varied with any linguistic manipulation (extraction distance)

Discussion

- SWP present null results and conclude that the constrained capacity approach to islands is not supported
- The current experiment presents positive results
  - It is possible to detect co-variation between cognitive measures and linguistic acceptability judgments
  - The null results reported here are more straightforwardly interpretable
- The lack of influence of any of the individual measures on the island violations indicate a lack of support for both
  - Constrained capacity approach to islands
  - Similarity-based interference approach to islands

Discussion

- The positive result from memory-interference co-variation with extraction distance:
  - Indicates the importance of similarity-based interference for long distance dependencies
  - Gives support to similarity-based interference as the proper way to conceptualize WM in sentence processing
- Phonological lure sub-portion of the similarity-based interference memory task is specifically implicated.
  - Phonological shape a cue to [+WH]?
  - Would different island types provide different results?
  - Would non-*wh* dependencies provide different results?

Thank you

Thanks to:

- Robert Kluender
- Grant Goodall
- Ivano Caponigro
- Seana Coulson
- Marta Kutas
- Language and Brain Lab
  http://grammar.ucsd.edu/brainlab/
- Experimental Syntax Lab
  http://grammar.ucsd.edu/syntaxlab/