The ‘Reductionist’ Approach

at least some island phenomena … owe their character to the accumulation of performance-based difficulties that rises above some threshold to create a perception of ungrammaticality 

Hofmeister and Sag 2009: 3

island-violating constructions involve numerous processing pressures that aggregate to drive processing difficulty above a threshold so that a perception of unacceptability ensues 

Hofmeister and Sag 2009: Abstract

Independently Known Sources of Processing Difficulty

- ‘long’ (i.e. multiclausal) filler-gap dependencies are harder to process than ‘short’ (monoclausal) filler-gap dependencies
- this effect manifests itself, in part, in measurably reduced acceptability ratings for examples such as (1b) as compared with (1a):
  
  (1) a. Which restaurant should we go to?  
       b. Which restaurant did they say that we should go to?

- island configurations frequently involve local configurations which are more complex in some sense than other kinds of configurations (whether versus that, for instance)
- this effect also manifests itself, in part, by way of measurably reduced acceptability ratings for examples such as (2b) as compared with (2a):
  
  (2) a. Who thinks that people should be sacked?  
       b. Who wonders whether people should be sacked?

neither of which is an island violation as classically understood.
- the coming together of these two effects could produce in consultants the experience of discomfort which we have come to name the ‘island effect’.

acceptability reflects the contribution of grammatical principles (competence-based factors), as well as constraints imposed by resource-limitations (performance-based factors) 

Hofmeister & Sag 2009: 2

Cumulativity and Interaction

But mere additivity is not enough. The quantifiable effect we call the ‘island effect’ is not the result of simply adding the unacceptability level linked with each of the two factors above (distance and complexity). If the difference in acceptability between (3a) and (3b):

(3) a. Who thinks that John bought a car?  
    b. What do you think that John bought?

is, say, -1, and the difference in acceptability between (3a) and (4):

(4) Who wonders whether John bought a car?

is, say, -2, then naive or rational people might expect that (5):

(5) What do you wonder whether John bought?

would come out as -3 (the simple sum of the two supposedly contributing factors) in its level of acceptability.

But that is not what we observe in our experiments. (5) is in fact consistently judged lower in acceptability than one would expect given linear additivity for the two factors. What we in fact see is super-additivity—in which (5) is assessed at an acceptability level of, say, -5 rather than -3.

How to Understand the Fact of Super-additivity?

For what Hofmeister and Sag call ‘competence theories’ of islandhood, super-additivity just is the observational correlative of the grammatical phenomenon of islandhood.

For ‘reductionist’ theories:

super-additivity must reflect ‘an interaction based on limited resources’.

For non-reductionist (grammatical, or competence-based) theories of islandhood, on the other hand, the degree of unacceptability of (5) reflects a fact about that structure alone (the crossing of an island boundary); appeal to any interaction of factors is irrelevant.

We therefore have a way of experimentally distinguishing the two approaches.
Distinguishing the two Theories

As processing resources are increased:

- **the competence theory** leads us to expect that there should be no corresponding increase of acceptability for (5), since on that theory the degree of unacceptability of (5) reflects a property of the competence grammar rather than any undue pressure on processing resources.
- **the reductionist theory** leads us to expect that as more resources become available, the strength of the interaction should decrease, and the measured degree of acceptability of (5) should increase correspondingly. That is:

  - As processing resources are increased, the ‘island effect’ should come to seem additive rather than super-additive.

But what would it mean to ‘increase processing resources’ in a context like this?

As it happens, individuals are known to differ with respect to their working memory capacity and:

- these differences are known to affect sentence processing as measured in reading times and in error rates.
- there are at least two standard ways of measuring an individual’s working memory capacity:
  A by way of a verbal serial recall task
  B by way of the n-back test

Therefore we are now in a position to empirically distinguish the two hypotheses. We can:

- recruit a bunch of subjects
- measure the working memory capacity of each of them
- get them to provide acceptability judgments on examples like (1)–(5)
- analyze the results to see if those with higher working memory capacity evaluate examples like (5) in a way that differs systematically from those with lower working memory capacity. Specifically:
  * if those with higher working memory capacity judge (5) and similar examples more acceptable than do those with lower working memory capacity, that would support the ‘reductionist’ theory
  * if there is no amelioration effect for those with high working memory scores, that would suggest something more like the grammatical theory of islandhood.

The Experiments

- 173 undergraduates at Irvine
- monolingual native speakers of English
- working memory capacity measured both by verbal serial recall and n-back tests
- four islands tested:
  a Whether-Island
  b n-complement case of the CNPC
  c Subject Condition
  d Adjunct Island Condition
- four conditions each:
  a short, no island (**who thinks John bought a car**)
  b long, no island (**what do you think that John bought**)
  c short, island (**Who wonders whether John bought a car**)
  d long, island (**What do you wonder whether John bought**)
- magnitude estimation for measuring acceptability

Three Analyses of the Data

**Analysis One: a Group Analysis**

- Take the top-performing 12.5% (those with the highest working memory capacity)
- Take the lowest-performing 12.5% (those with the lowest working memory capacity)
- Ignore the middle 75%
- Look for evidence that the high-performers do not show super-additivity.

**Outcome:** There is no such evidence.
Analysis Two: a Linear Regression Analysis

But by doing a group analysis, we throw away 75% of the data. How can we keep all of the data and still look for an effect of working memory capacity on perceived island strength?

Calculate (for each subject and for each island) the strength of the island-effect by calculating a differences-in-difference score (ΔΔ)—an index of the strength of the island-effect (a number reflecting for each subject and for each island how much a degradation in acceptability results when the filler-gap dependency extends across the island-boundary). Then:

We can plot each subject's island-strength-score (their ΔΔ) against the measure of their working memory capacity.

Reductionist theories presumably predict that ΔΔ’s should decrease as working-memory-capacity increases.

Grammatical theories predict that there should be no relationship between the two scores.

Linear regression analysis suggests that there is no relationship.

Analysis Three: a Bootstrap Analysis

But we're still throwing away a lot of potentially important data, because the score for each subject in each condition in each island is an average over four tokens (four judgments).

Don't average. Choose one judgment out of each quadruple at random and derive a different ΔΔ.

Each such randomly-chosen quadruple can be viewed as an outcome from the experiment.

If we do this for each subject, we will come up with a different result for the linear regression.

This process can be repeated 10,000 times, in effect doing 10,000 different (simulated) re-runs of the experiment.

Outcome: Again there doesn't seem to be any effect of resource capacity on the interaction effect, even at this very fine-grained level of analysis.