Informativity and filler-gap processing
Differences in fillers


- Which article don’t you remember [who wrote ___]? > What don’t you remember [who wrote ___]?  
- That’s the article that you wondered [who was reading ___]. > What did you wonder [who was reading ___]?  
- Which book did which student read ___? > *What did who read ___?
Starting point

Why should the content of an extracted element affect its “extractability”? Why should non-syntactic choices of form affect structural soundness? Why should the choice of wh-phrase form create differences in acceptability?
Preserving structural constraints

- Rizzi (1990) & Cinque (1990): referential (who/every NP/someone/whoever/etc.) vs. non-referential (what/all DP/many DP/some NP)

- only arguments that “refer to specific members of a set in the mind of the speaker or preestablished in discourse” can be exceptions to antecedent government requirements

- This provides an “out” for a distinction between argument versus adjunct extraction behavior in syntactic islands --> leads to distinction between cyclic and unbounded movement
Preserving structural constraints

- On Rizzi & Cinque’s view, the long distance binding possibilities observed with anaphora are shared by “referential” arguments
- Dissonance between indexing system for anaphora and “referential” arguments
  - Bound variable anaphora can be anteceded by nonreferential DPs (e.g. Who divorced his wife?, Each girl loves her mother)
- Chung (1994:33): “Rizzi’s referential indices cannot be identified with the indexing mechanism that is a mainstay of current approaches to anaphora ... if that is so, then we are left wondering whether the use of indices in this theory amounts to more than a diacritic to distinguish the DPs that allow long movement from those that do not.”
D-linking (Pesetsky 1987, 2000): “discourse-linked” wh-phrases are not subject to the normal rules on movement

“Context sets previously mentioned in the discourse qualify a phrase as D-linked, but so do sets that are merely salient (e.g., which book, in a context where speaker and hearer both know that reference is being made to a reading list for a course) and sets whose salience is culturally determined (e.g. what day of the week, which sign of the zodiac).”

Lexical items are not inherently marked for this property; it emerges from context
D-linking

Relationship between larger discourse salience and conditions on movement remains opaque, even to Pesetsky:

“A reliable rule of thumb is that if a wh-word in a multiple wh-question can be felicitously paraphrased with an expression of the form which of the X, it can cause the Superiority effect to disappear. The reason for this link between semantics and syntax is obscure, and will remain obscure even at the end of this book.”
Differences in *wh*-phrases

- Chung (1994) asks: “what is it about the ability to narrow down the domain of *wh*-quantification ‘enough’ that makes it possible for strict locality to be violated?”

- Considering a variety of pragmatic and semantic accounts (e.g. Frampton 1990; Kroch 1989; Comorovski 1989), she ultimately concludes that no current approach offers a unified way of treating all the evident data.
Alternatively . . .

Maybe these differences can best be explained via processing
- We already know that holding a filler in working memory is a relatively onerous task (Wanner & Maratsos 1979; Hawkins 1999, etc.)
- There’s something about having more descriptive content in a filler expression that makes the dependency easier to process
- Of course, that leads to the question of why this should be so

Prediction: significant processing advantage when more informative filler phrases are used in FGDs

Implication: uninformative fillers may be at least partly responsible for certain island effects due to processing difficulty
Albert learned that the managers dismissed the employee with poor sales after the annual performance review.

- Who did Albert learn whether they dismissed after the annual performance review? (BARE)
- Which employee did Albert learn whether they dismissed after the annual performance review? (WHICH)
- Who did Albert learn that they dismissed after the annual performance review? (BASELINE)
F1 (1, 19) = 6.55, p < .05
F2 (1, 23) = 7.725, p < .01
F1 (1, 19) = 15.255, p < .001
F2 (1, 23) = 14.420, p < .001
n.s.
**wh-island reading times**

- Using more explicit *wh*-phrase leads to reading times comparable to baseline’s
- More descriptive/informative *wh*-expression results in easier sentence processing
- No effect of *wh*-type prior to verb where retrieval and integration takes place, i.e. this effect looks like it’s related to retrieval of the *wh*-phrase
wh-island reading times

- These are, however, merely argument extractions, although tensed wh-islands should be strong by most accounts.

- How far does this effect of informativity go? Does it work only for arguments or “referential” fillers?
Adjunct extractions in *wh*-islands

Task: judge whether it’s possible to answer the question

Andrew overheard the daycare staff discussing how they wanted to get away from the children for a few hours.

**BARE:** How long did Andrew hear whether the children had played during the daycare’s afternoon recess?

**LONG:** How many hours did Andrew hear whether the children had played during the daycare’s afternoon recess?

**BASELINE:** How long did Andrew hear that the children had played during the daycare’s afternoon recess?
did Andrew hear whether/that the children had

F1 (1, 27) = 12.136, p < .01
F2 (1, 23) = 12.339, p < .01
The advantage of longer and more informative fillers applies not only to arguments but adjuncts as well.

Distinction between weak and strong islands may nevertheless be real inasmuch as storing and retrieving adjunct phrases may be harder.

Is this a more general phenomenon that applies to the processing of all sorts of filler-gap dependencies, or is this some sort of relic of island processing?
The driver who was friends with Jill came in first in the 2005 NASCAR race, but finished last the previous year.

**SIMPLE:** What did a driver that Jill knew **win** after losing the previous year?

**WHICH:** Which race did a driver that Jill knew **win** after losing the previous year?

**COMPLEX:** Which NASCAR race did a driver that Jill knew **win** after losing the previous year?

{1) 2005 NASCAR race; 2) 2006 NASCAR race ; 3) 2006 bobsled race}
Results: Reading times @ verb

- Which-phrases led to significantly faster reading times at V2
- Comp-simple: $F_1(1,34) = 5.315, p < .05$; $F_2(1,19) = 6.634, p < .05$
- Which-simple: $F_1(1,34) = 9.473, p < .01$; $F_2(1,19) = 10.392, p < .01$
Results: Reading times after verb

- Bare *wh*-phrase leads to slower reading times (~40 ms)
- Comp-simple: $F_1(1,34) = 9.610, p < .01, F_2(1,19) = 5.486, p < .05$
- Which-simple: $F_1(1,34) = 9.202, p < .05, F_2(1,19) = 7.446, p < .05$
- Effect does show up on previous word (V1), but not earlier
Beyond wh-phrases: indefinites

16 items, 3 conditions:

a. It was a communist who the members of the club **banned** from ever entering the premises. (SIMPLE)
b. It was an alleged communist who the members of the club **banned** from ever entering the premises. (MID)
c. It was an alleged Venezuelan communist who the members of the club **banned** from ever entering the premises. (COMPLEX)
Reading times at verb

30 ms difference between complex and simple conditions

Near significant by subjects $F(1,34) = 5.351, p = .063$

Significant by items $F(1,15) = 3.690, p < .05$

Beyond wh-phrases: indefinites
**wh-island reading times**

- Evidence from a number of filler-gap dependencies that more informative filler-phrases facilitate processing at the retrieval and integration site.

- In *wh*-islands, the less informative filler phrase results in a serious slowdown in processing, which is potentially pushing difficulty past a threshold where the sentence becomes unacceptable.

- Again, what principle should account for this?
  - Note: there doesn’t have to be just one!
Memory facilitation hypothesis

Linguistic elements that encode more information (lexical, semantic, syntactic, etc.)

- Facilitate their own memory retrieval later

Intuitively, the idea is that poorly characterized/less distinctive discourse entities are harder to recall

Example: *which therapist* encodes more information than *who* → *which therapist* leads to faster retrieval (all else being equal)
Memory facilitation hypothesis

- Bradshaw & Anderson (1982): sentence recall improves when presentation occurs with supporting facts (see also Anderson & Reder, 1979; McDaniel et al. 1988)
- Elaboration increases the number of possible retrieval paths to the target concept
- Multiple retrieval paths improves chances of successful recall
- All effects observed here were localized to retrieval sites
Memory facilitation hypothesis

Distinctiveness (reduced interference)
- As more features get encoded, corresponding stored item stands out more from any competitors in memory
- Upon retrieval, interference effects will be reduced
- Note that this type of processing explanation can explain observed differences across referential and non-referential phrases
Memory facilitation hypothesis

Some general principle of memory and comprehension is responsible for variation in acceptability in island contexts

- One more factor that, in combination with other processing hurdles, results in unacceptability of certain types of filler-gap dependencies