

# Meaning explanations of two syntactic islands: Subject islands and Embedded-clause islands

Ted Gibson  
May 19 2018

Project 1: Anne Abeillé; Barbara Hemforth, Elodie Winckel, Ted Gibson  
Project 2: Yingtong Liu, Rachel Ryskin, Richard Futrell, Ted Gibson

# Long-distance dependencies

What makes a long-distance dependency unacceptable?

(Ross, 1967; Chomsky, 1973, 1981, 1982; 1986; 1993; and many others; see Goldberg, 2016; Ambridge & Goldberg, 2008 for a functionalist proposal)

Sometimes long-distance dependencies seem ok:

Ia. What does Susan think that John bought \_\_\_?

Ib. What does Sarah believe that Susan thinks that John bought \_\_\_?

Ic. What does Bill claim that Sarah believes that Susan thinks that John bought \_\_\_?

# Long-distance dependencies

“Islands of extraction”: Ross, 1967

Sometimes long-distance dependencies are not so good:

Complex NPs

\*Who did she see the report that was about \_\_? (cf. She saw the report that was about the senator)

Subjects

\*Who did that Mary liked \_\_ bother him? (cf. That Mary liked x bothered him)

?\*Which sportscar did the color of \_\_ delight the baseball player?

(cf. The color of the sportscar delighted the baseball player)

Adjuncts

??What did she leave the movie because they were eating \_\_?

(cf. She left the movie because they were eating x)

Complements of manner-of-speaking verbs

??What did Mary whisper that Bill liked \_\_? (cf. Mary whispered that Bill liked wine)

Complements of factive verbs

??What did Mary know that Bill liked \_\_? (cf. Mary knew that Bill liked wine)

# Long-distance dependencies

Not only wh-questions: Also other long-distance dependencies within a language

E.g., Relative clauses

Complex NPs (both noun complements and relative clauses)

\*The senator who she saw the report that was about \_ bothered many people on the committee.

Subjects

?\*The sportscar which the color of \_ delighted the baseball player

Complements of manner-of-speaking verbs

?The wine that Mary whispered that Bill liked \_ was a Cabernet.

## **I. Purely syntactic accounts (Chomsky, 1973)**

If meaning were the source of the unacceptability of long-distance extractions, then different extraction-types (with different meanings) should vary in their acceptability, depending on their meaning (e.g., RC vs. WHQ) but they don't

**All that seems to matter is their form**

# I. Purely syntactic accounts (Chomsky, 1973)

Furthermore, initially, it looked like a **phrase-structure measure** could explain many of these unacceptabilities.

**Hence purely form-based accounts**, such as the **Subjacency** Condition (Chomsky, 1973, 1977; cf. Ross, 1967; Chomsky, 1986; Chomsky, 1995):

No rule may move a phrase from position Y to position X (or conversely) in

...X ... [A ... [B ...Y ...] ...] ...X ...

Where A and B are “cyclic nodes”. Maybe what counts as a “cyclic node” varies across languages. Claim: NP and S are cyclic nodes in English

Who<sub>i</sub> [<sub>S</sub> \_\_\_<sub>i</sub> was ignoring Susan ]

Who<sub>i</sub> [<sub>S</sub> was Anthony ignoring \_\_\_<sub>i</sub> ]

\*Who<sub>i</sub> did [<sub>S</sub> she see [<sub>NP</sub> the report [<sub>S</sub> which was about \_\_\_<sub>i</sub> ]]]

\*Which sportscar<sub>i</sub> did [<sub>S</sub> [<sub>NP</sub> the color of \_ ] delight the baseball player]

# I. Purely syntactic accounts (Chomsky, 1973)

Given that it is only the form of these extractions that blocks them (not memory limitations; not meaning) then they can't be learned

These kinds of observations led Chomsky and others (e.g., Phillips, Hornstein at Maryland) to argue that these constraints are a core part of **Universal Grammar (UG)**

# Challenges to the syntactic account

1. Lots of cross linguistic variability in the acceptability of different islands, starting with Rizzi (1982)

2. Some complexities need to be introduced to explain the most basic phenomena:

What does [<sub>S</sub> Susan think that [<sub>S</sub> John bought \_\_\_<sub>i</sub> ] ]?

Chomsky's **“escape hatch”** : S' with no wh-phrase in it (“who”, “which”, “whether”): if you can move from one S' to the next without crossing 2 cyclic / bounding nodes, then the extraction is ok.

Each link crosses only one bounding node:

What<sub>i</sub> does [<sub>S</sub> Susan think [<sub>S'</sub> \_\_\_<sub>i</sub> that [<sub>S</sub> John bought \_\_\_<sub>i</sub> ] ] ]?

What<sub>i</sub> does [<sub>S</sub> Sarah believe [<sub>S'</sub> \_\_\_<sub>i</sub> that [<sub>S</sub> Susan think [<sub>S'</sub> \_\_\_<sub>i</sub> that [<sub>S</sub> John bought \_\_\_<sub>i</sub> ] ] ] ]?

\* Who<sub>i</sub> did [<sub>S</sub> she see [<sub>NP</sub> the report [<sub>S</sub> which was about \_\_\_<sub>i</sub> ] ] ]

\* Which sportscar<sub>i</sub> did [<sub>S</sub> [<sub>NP</sub> the color of \_\_\_ ] delight the baseball player]



## 2. Meaning-based accounts

Some extractions have strange meanings (e.g., Erteschik-Shir 1977; Erteschik-Shir & Lappin 1979; Kuno 1987; Ambridge & Goldberg, 2008; Goldberg, 2016; Ambridge et al., 2014):

### **Conflict in information structure**

Maybe it's not possible to wh-question an element whose function in the declarative form is backgrounded or presupposed, because the function of a wh-question is to obtain focus information (new, foregrounded information)

Complex NPs (both noun complements and relative clauses)

\*Who did she see the report that was about \_\_? (cf. She saw the report that was about x)

Backgrounded clausal subjects:

\*Who did [that she knew] bother him? (Goldberg 2016)

### 3. Processing / memory-based accounts

Non-syntactic factors like plausibility and working memory demands affect extraction difficulty (e.g., Deane, 1991; Kluender, 1992; Sag et al., 2007; Hofmeister & Sag, 2010):

- (2) \*Who did [ you criticize votes for [ the impeachment of \_\_\_]]?
- (3) Who did [ you obtain votes for [ the impeachment of \_\_\_]]? (Deane, 1991)
  
- (4) \*Who did Emma doubt the report that we had captured \_\_\_?
- (5) Which convict did Emma doubt a report that we had captured \_\_\_?  
(Hofmeister & Sag, 2010)

**Maybe some syntactic island effects can be explained by resource complexity or other non-syntactic factors like world knowledge / meaning**

# Constraints on long-distance dependencies

1. Syntactic accounts

2. Meaning-based accounts:

- conflicts in construction meanings
- world knowledge

3. Working-memory accounts

# Structure of presentation

1. Introduction to islands

2. Methods: quantitative acceptability judgments from Mechanical Turk and a French subject pool

3. A discourse-focus explanation of the subject-island constraint (Anne Abeillé, Barbara Hemforth, Elodie Winckel, Edward Gibson)

Contrary to the standard claim, there is a dissociation between acceptability of extractions from WH-questions and relative clauses in English and French

4. S-complement islands (Yingtong Liu, Rachel Ryskin, Richard Futrell, Edward Gibson)

Contrary to the standard view, there is no interaction between sentence type and verb frequency for S-complement islands, suggesting that verb and construction frequency explain the badness of such examples

5. Conclusion: **Maybe all “islands” are meaning- and processing-based, contrary to the UG / syntax claim**

# Acceptability judgments are continuous

1. Lexical knowledge is frequency based: higher frequency words are accessed faster
2. N-gram surprisal is measurable to six orders of magnitude (Smith & Levy, 2013)
3. World knowledge: a continuum from more expected to less expected
4. Grammatical knowledge: probably also a continuum from more expected to less expected: phrase structure rule / construction frequency

Because grammatical knowledge is probably continuous there are an infinite number of possible comparisons on this dimension

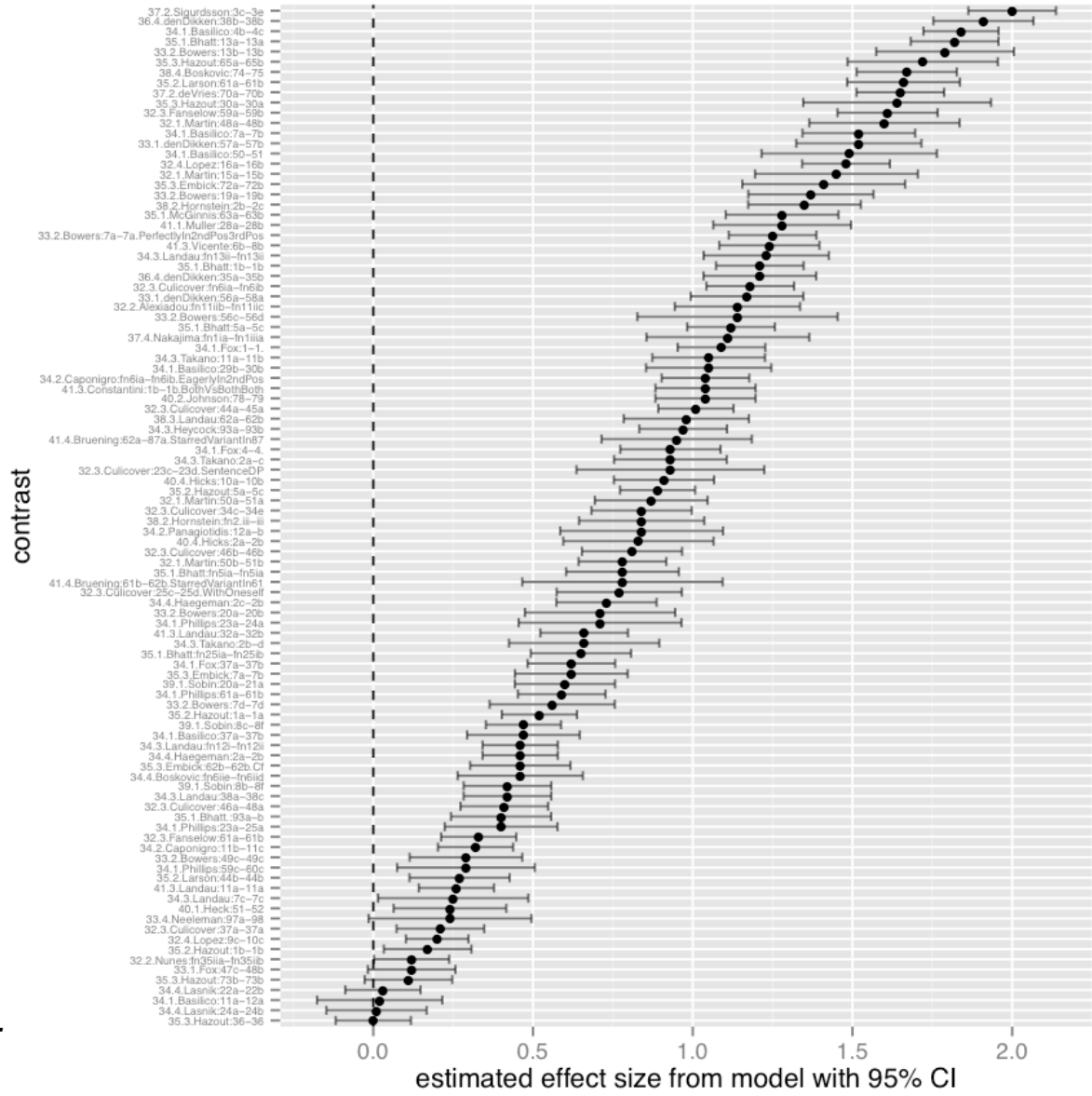
Because of cognitive bias and poor working memory (hard to keep track of more than a few alternatives at once) we eventually need quantitative evidence for a detailed model of the grammar

# Investigating the potential categoricity of grammaticality: Mahowald, Graff, Hartman & Gibson (2016)

Method: acceptability judgement 1-7, z-scored within individuals. obtain mean z-score for each item in each contrast, and averaged these to give an overall z-score for the 'acceptable' sentence and for the 'unacceptable' sentence in each contrast.

The effect size is the difference between these two z-scores.

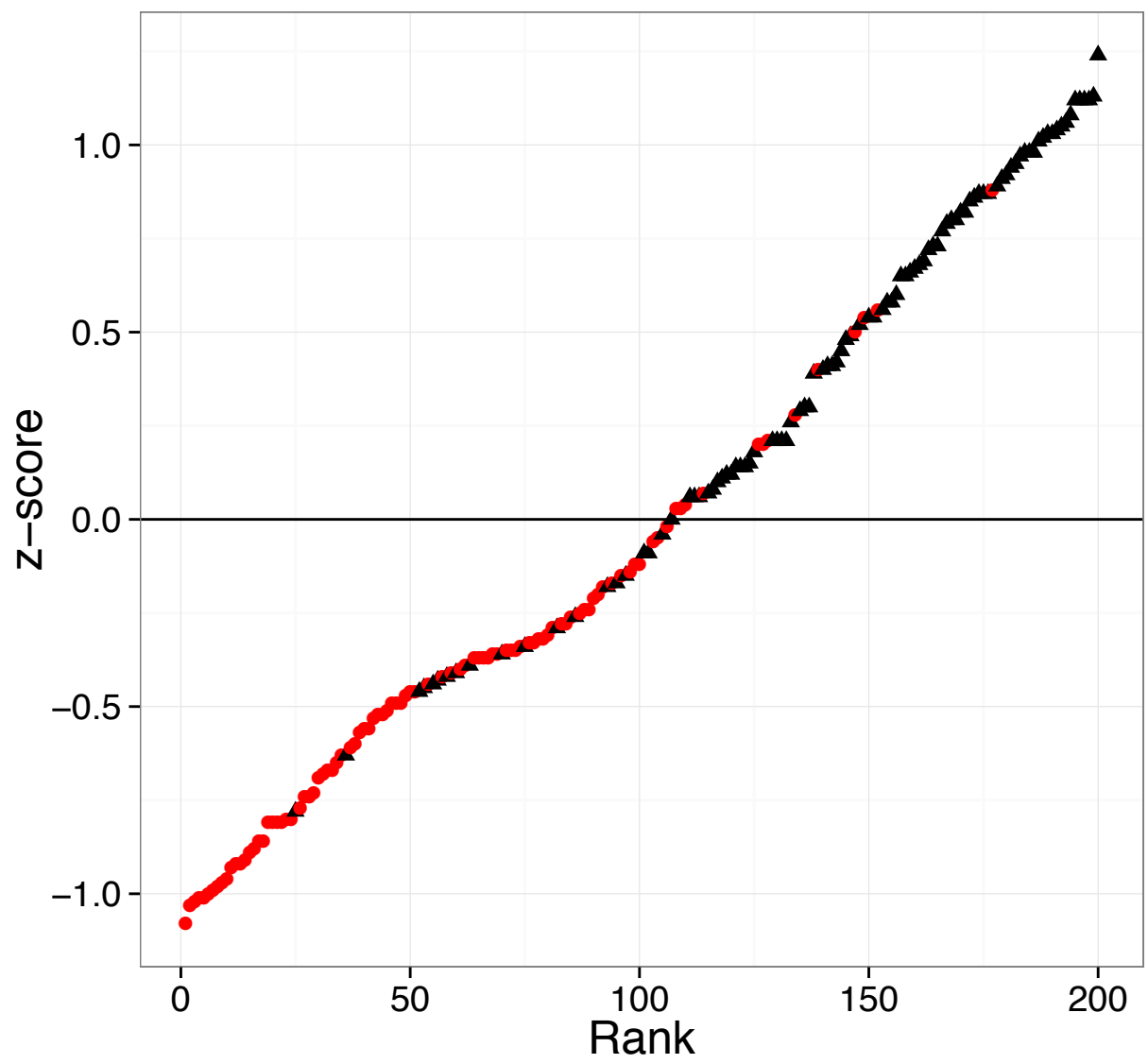
(Effect size: Cohen's *d* is a measure of effect size that is equal to the difference in means between the two conditions, in z-scores.)



z-score:  
(value-  
mean)/  
sd



# Investigating the potential categoricity of grammaticality: Mahowald, Graff, Hartman & Gibson (2016)



z-score:  
(value-  
mean)/  
sd

**GoodOrBad**  
● Bad  
▲ Good



# Acceptability judgments and islands

The critical claim for an island structure is that the acceptability difference for the extraction versions is much more than for some controls without the extraction:

Sprouse, Caponigro, Greco & Cecchetto (2016, NLLT):

object, short:

I voted for the *congressman who you think the lobbyist offended* \_\_\_.

subject, short:

I voted for the *congressman who you think* \_\_\_ offended the lobbyist.

object: long:

I voted for the *congressman who you think the gift from the lobbyist prompted the rumor about* \_\_\_.

subject, long:

I voted for the *congressman who you think the gift from* \_\_\_ prompted the rumor about bribery.

The subject / long condition is worse than the object / long condition.

But critically to show that it's something about the "island" subject and not just any subject position, , it's extra worse compared to the short subject versions: an interaction in the ratings between length and position

**It's hard to have judgments about main effects, let alone interactions, hence we need experiments**



# **Case study I: Subject islands**

## **A discourse-focus explanation of the subject-island constraint**

Anne Abeillé; Barbara Hemforth, Elodie Winckel, Ted Gibson

### **Subject islands:**

*?\*Which sportscar did the color of \_ delight the baseball player?*

In Chomskyan grammar, Adger (2002) list subject-islands as canonically ungrammatical across languages. Szabolcsi (1997, 2006) says subject-islands are “strong” islands. Even Pinker (1994). Other textbooks only mention the sentential subject as an island (e.g., Carnie 2006)

Italian was claimed to be an exception (Rizzi, 1982).

But counter evidence has been found in many languages:

Japanese (Ross 1967), French (Sportiche 1981, Godard, 1992), Italian (Rizzi 1982), Spanish (Torrego 1984), Russian, Norwegian and others (see Stepanov 2007 for a review),

# Experimental evidence

## Extractions from subjects: Sprouse et al. (2016)

### RCs and Wh-questions

#### English RCs:

I voted for the *congressman who you think the lobbyist offended* \_\_\_.

I voted for the *congressman who you think* \_\_\_ offended the lobbyist.

I voted for the *congressman who you think the gift from the lobbyist prompted the rumor about* \_\_\_.

I *voted for the congressman who you think the gift from* \_\_\_ prompted the rumor about bribery.

#### Italian RCs:

Ho incontrato il *giornalista che pensi che il direttore abbia fatto licenziare* \_\_\_. 'I met the journalist that you think that the director fired.'

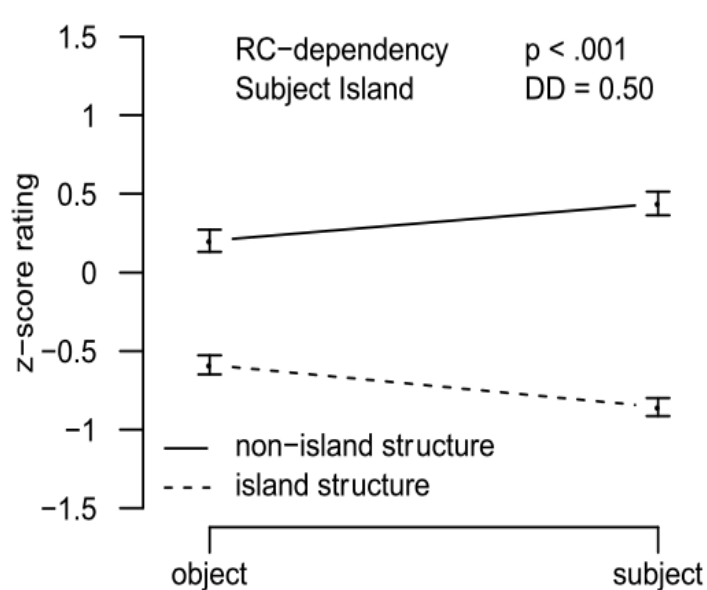
Ho incontrato il *giornalista che pensi che* \_\_\_ abbia fatto arrabbiare il direttore. 'I met the journalist that you think pissed off the director.'

Ho incontrato il *giornalista del quale pensi che* [l'articolo del direttore] abbia causato il licenziamento \_\_\_. 'I met the journalist who you think the director's article has causes the firing of.'

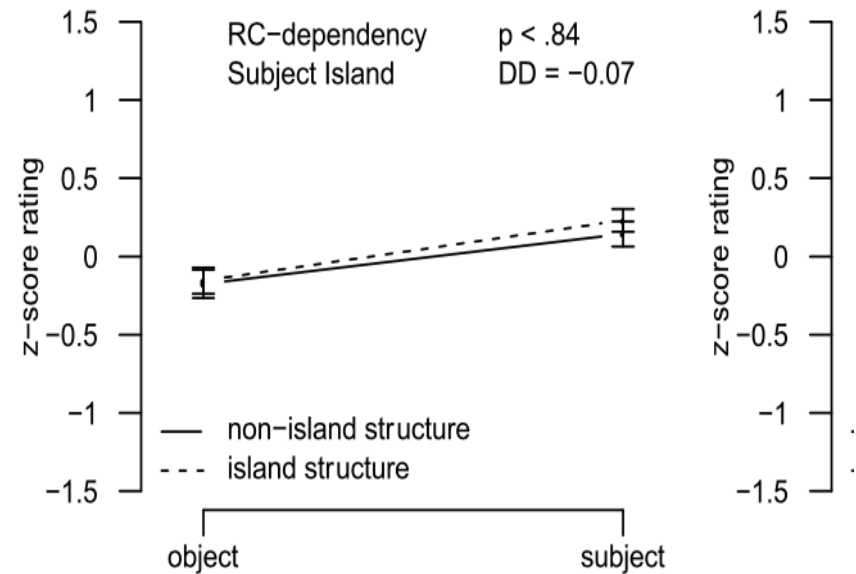
Ho incontrato il *giornalista del quale pensi che* [l'articolo \_\_\_] abbia causato il licenziamento del direttore. 'I met the journalist who you think that the article of caused the firing of the director.'

# Extractions from subjects: Sprouse et al. (2016)

For relative clauses, a subject-extraction penalty only in English:  
subject-islands are in English



English



Italian

## **Differences in Sprouse et al.'s materials: English vs. Italian**

- Preposition stranding in English
- Pied-piping in Italian

The constructions are not parallel

# Experiment I: Extractions from subjects and objects in WH-questions in English

## OBJECT

NP-extracted: Which sportscar did the baseball player love the color of?

PP-extracted: Of which sportscar did the baseball player love the color?

ungrammatical: Which sportscar did the baseball player love the color?

no-extraction: Did the baseball player love the color of the sportscar?

## SUBJECT

NP-extracted: Which sportscar did the color of delight the baseball player?

PP-extracted: Of which sportscar did the color delight the baseball player?

ungrammatical: Which sportscar did the color delight the baseball player?

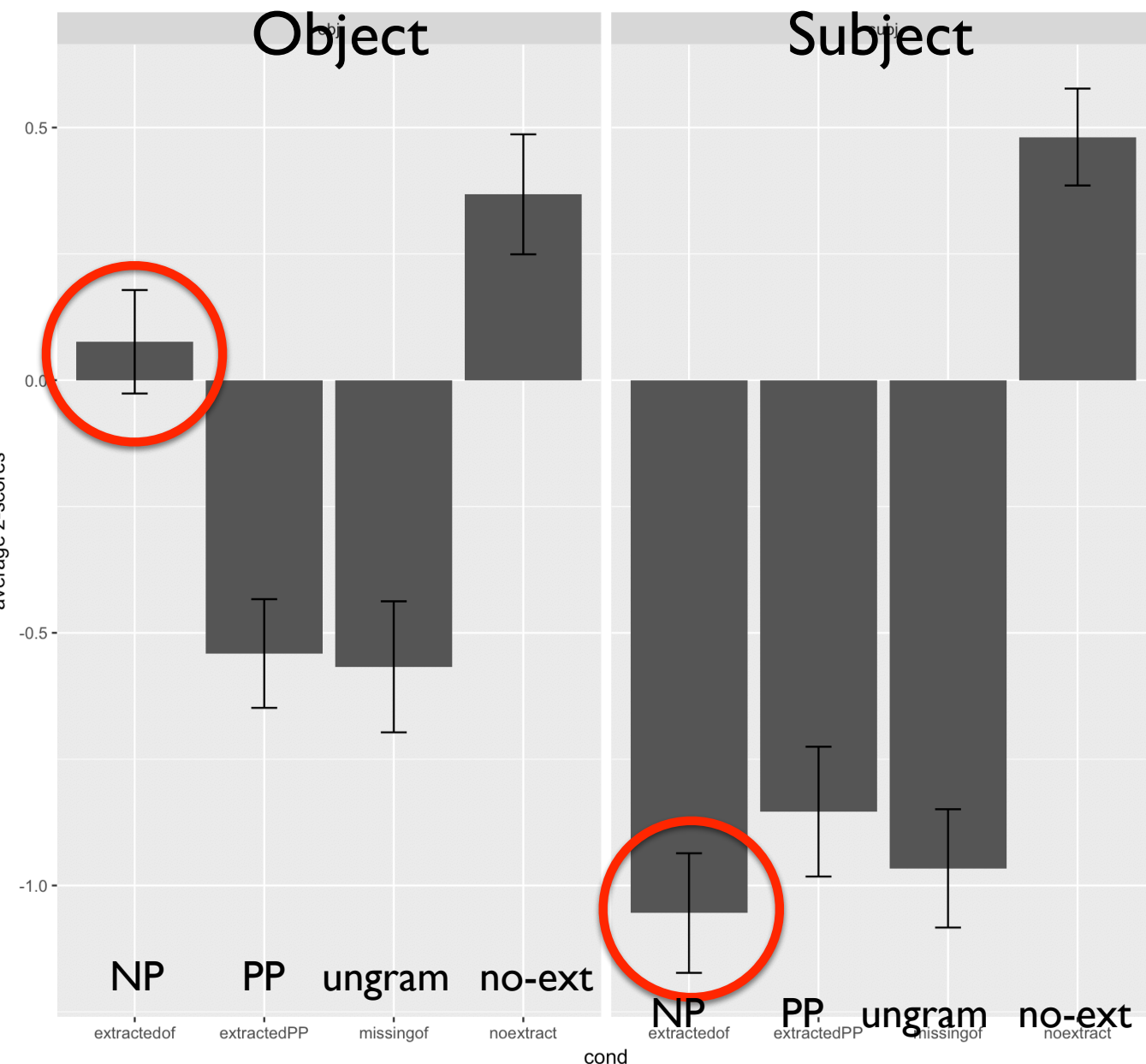
no-extraction: Did the color of the sportscar delight the baseball player?

64 participants from Mechanical Turk rated the acceptability of each trial for 24 target items and 44 distractor items; 7 point scale: Completely unacceptable - completely acceptable

# Project I: English Experiment I: WH-questions Results

object, NP-extracted: Which sportscar did the baseball player love the color of?

subject, NP-extracted: Which sportscar did the color of delight the baseball player?



## Results:

1. only the P-stranded extraction from object condition was near the grammatical baseline
2. All the other extractions (both from subject) are close to the ungrammatical baseline.

**Replication of intuition from literature**

# Let's look at RCs more closely

## **Observation from corpora: Extractions from subjects aren't uncommon for RCs, with the whole PP extracted:**

*There are certain topics that [jokes about \_\_\_] are completely unacceptable. Levine and Sag (2003)*

*... a letter of which every line \_\_\_ was an insult*

(Un)-biasing materials by looking at corpora

What is a good environment for out of subject extractions? (Chaves 2013; cf. Ross, 1967; Chomsky, 2008)

- non animate nouns with *of* complement (so the complement of the noun is not possibly extraposed, contrary to other prepositions)
- non agentive verbs
- the clause must be relevant both for the Noun and its complement
  
- To make sure that the extraction site is in the subject: Property-denoting nouns (*the price of x, the smell of x, etc.*): so it's not extraposed (as for other prepositions)
- Transitive non-agentive verbs, in reversible pairs (*fear/frighten, like/please, love/delight, etc.*)

# Experiment 2: Relative clauses

## OBJECT

coordination: The dealer sold a sportscar, and the baseball player loved the color of the sportscar.

PP-extracted: The dealer sold a sportscar, of which the baseball player loved the color.

in-situ: The dealer sold a sportscar, the color of which the baseball player loved.

ungrammatical: The dealer sold a sportscar, which the color the baseball player loved.

## SUBJECT

coordination: The dealer sold a sportscar, and the color of the sportscar delighted the baseball player.

PP-extracted: The dealer sold a sportscar, of which the color delighted the baseball player.

in-situ: The dealer sold a sportscar, the color of which delighted the baseball player.

ungrammatical: The dealer sold a sportscar, which the color delighted the baseball player.

64 participants from Mechanical Turk rated the acceptability of each trial for 24 target items and 44 distractor items; 7 point scale: Unacceptable - acceptable



## Experiment 2: Relative clauses

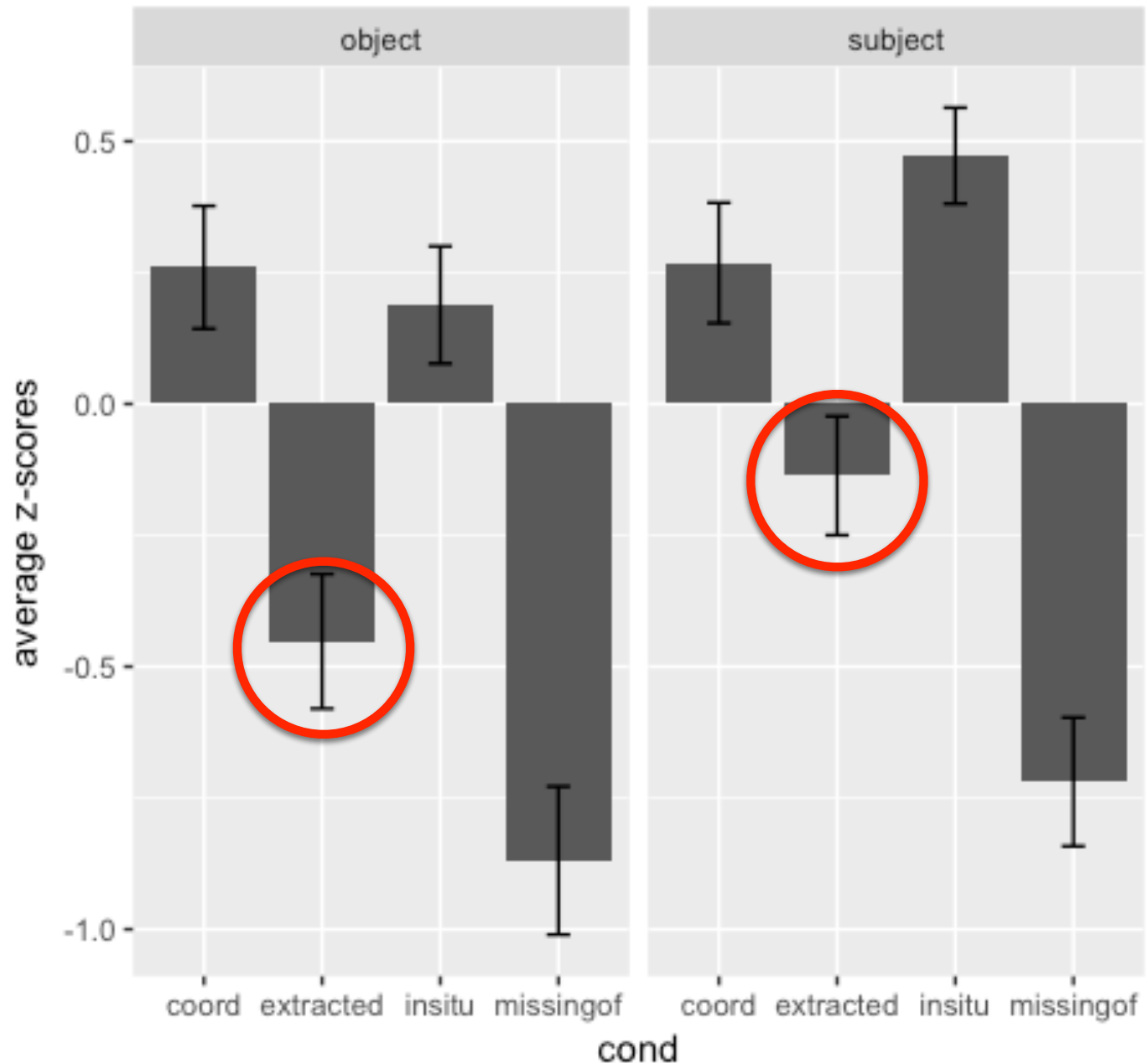
object, PP-extracted: The dealer sold a sportscar, of which the baseball player loved the color.

subject, PP-extracted: The dealer sold a sportscar, of which the color delighted the baseball player.

### Results:

subject-extracted is better than object-extracted!

(Same result for Experiment 3: restrictive RCs, with no commas)



# The Focus constraint (Erteschik-Shir, 1977) applied to questions

Proposal: **The Focus constraint:** (Erteschik-Shir 1977, 1981; cf. Van Valin, 1986; Takami, 1992; Deane, 1992; Goldberg, 2006): It is infelicitous to question an element out of a backgrounded constituent

Wh-questions:

function of extracted element in upper clause: focus (new)

extraction domain: the subject is a natural topic

- The diverging discourse status for the extracted element and the extraction domain is infelicitous

**Thus: the subject extraction penalty in wh-questions is due to a clash of discourse status of the extracted noun and its extraction domain**

More general prediction: questioning out of topic or backgrounded constituents is infelicitous (Goldberg, 2013, Ambridge et al., 2014)

# The Focus constraint

## **Relative clauses:**

function of extracted element in upper clause: any discourse status in the main clause (not in focus position)

extracted element: local topic (inside the RC: what the RC is about)

So no penalty for out of subject extraction (or object extraction)

## **A possible advantage for relativization out of subject**

Compatible with the general advantage for subject relative clauses (Gibson 2000; Lewis & Vasishth, 2005)

Or perhaps a corollary to Keenan & Comrie 1973's accessibility hierarchy ?

Relativizing of the subject possessor easier than relativizing of the object possessor (Hawkins 2004, Hale 2003)

# French Experiment I: Wh-questions

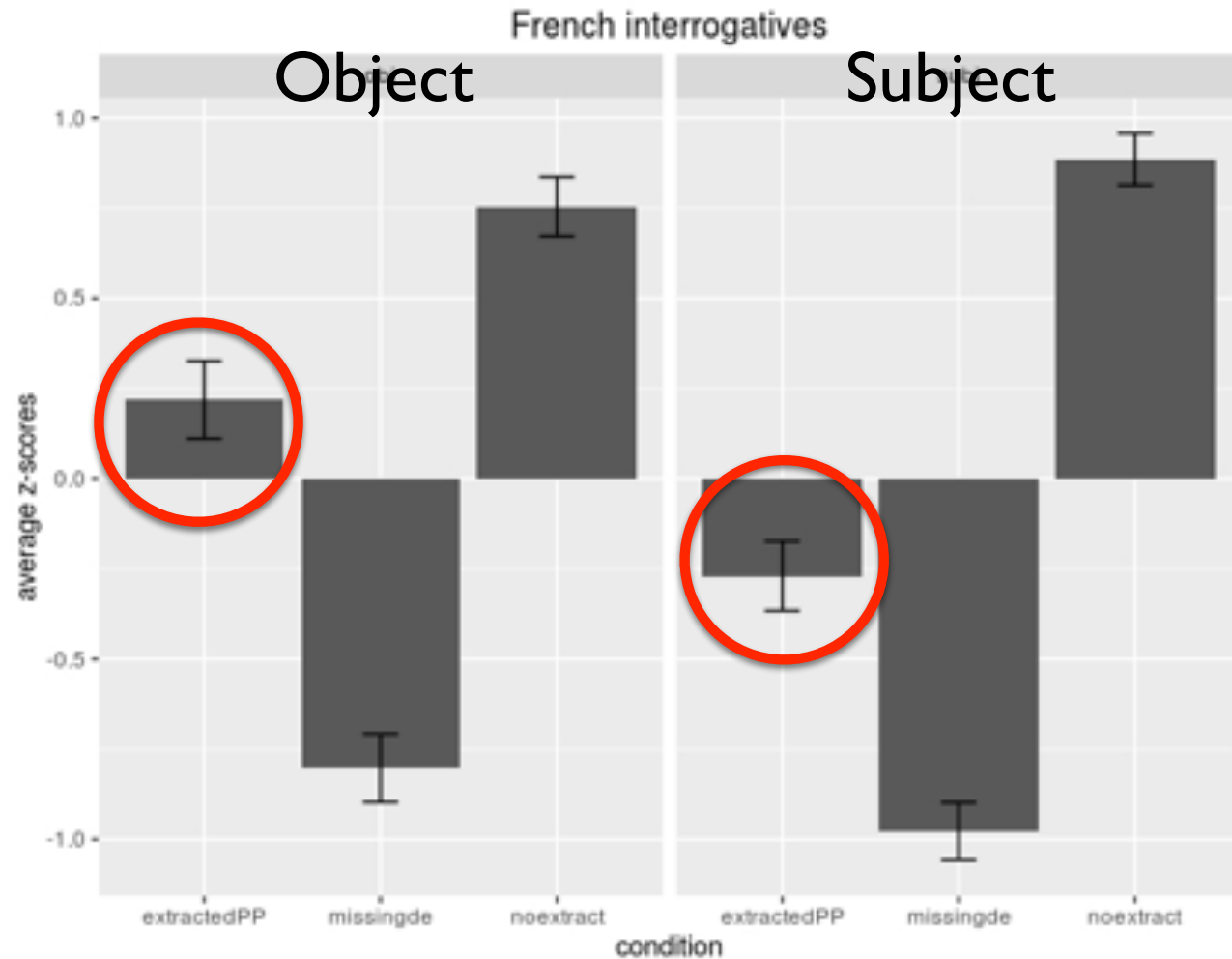
object extraction: De quelles activités est-ce que les employés craignent certains aspects pour leur santé?

(Of which activities did the employees fear certain aspects for their health?)

subject extraction: De quelles activités est-ce que certains aspects menacent la santé des employés?

(Of which activities do certain aspects endanger the employees' health?)

**Replication of English results:  
object extraction is better  
than subject extraction**

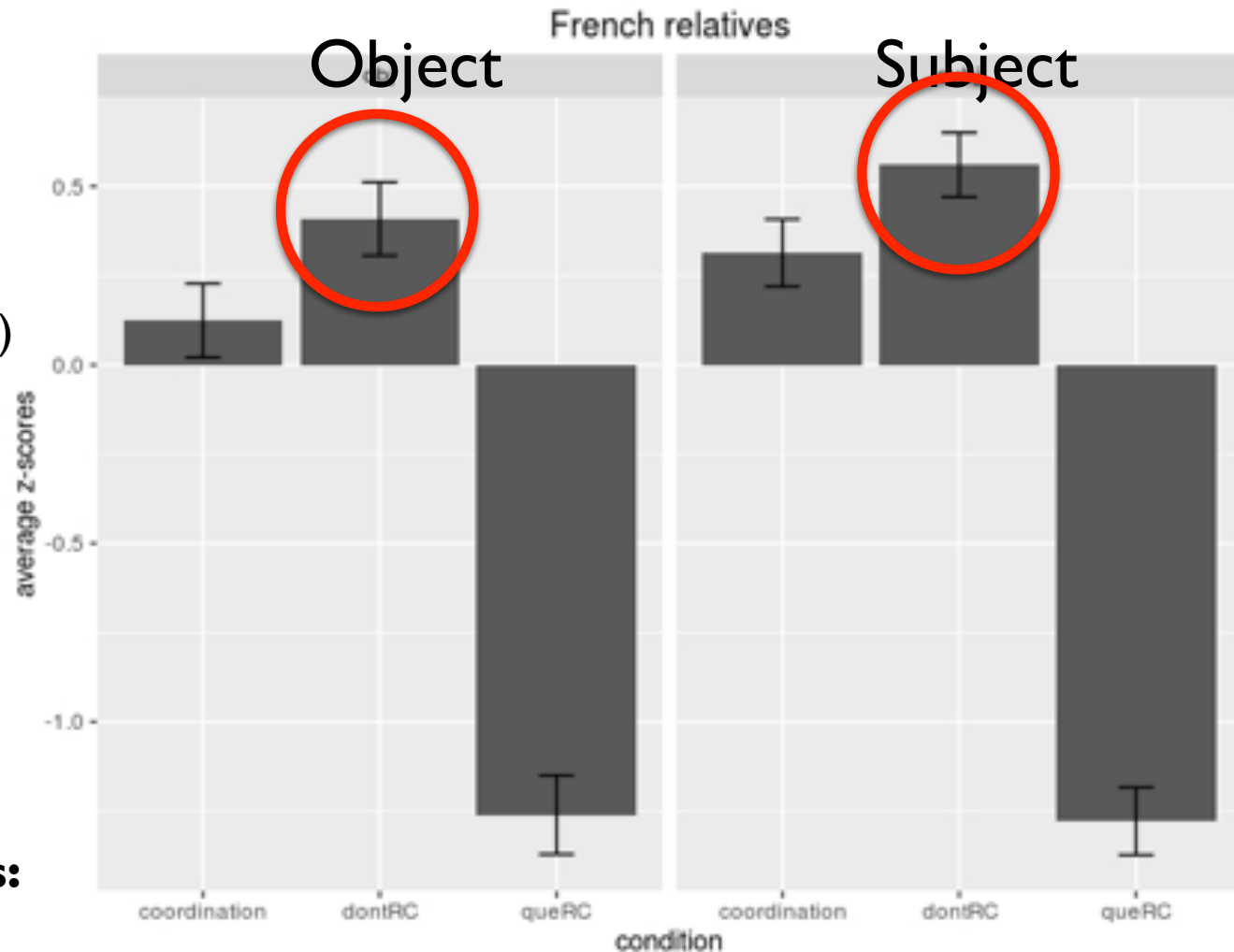


# French Experiment 2: Relative clauses

object extraction: L'entreprise organise des activités, dont les employés craignent certains aspects pour leur santé.  
(The company organizes some activities, of which the employees fear certain aspects for their health.)

subject extraction: L'entreprise organise des activités, dont certains aspects menacent la santé des employés.  
(The company organizes some activities, of which certain aspects endanger the employees' health.)

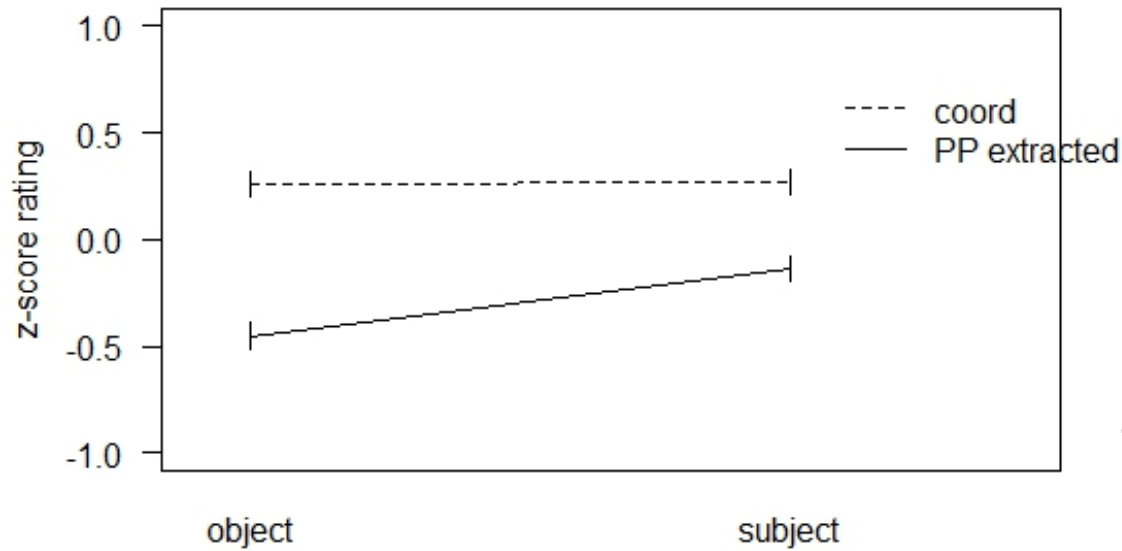
**Replication of English results:  
subject extraction is better  
than object extraction**



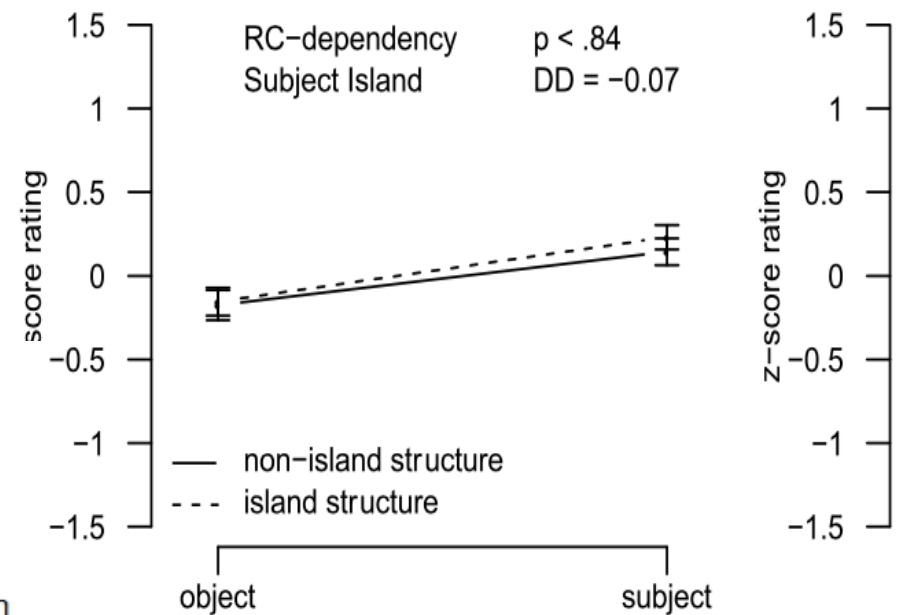
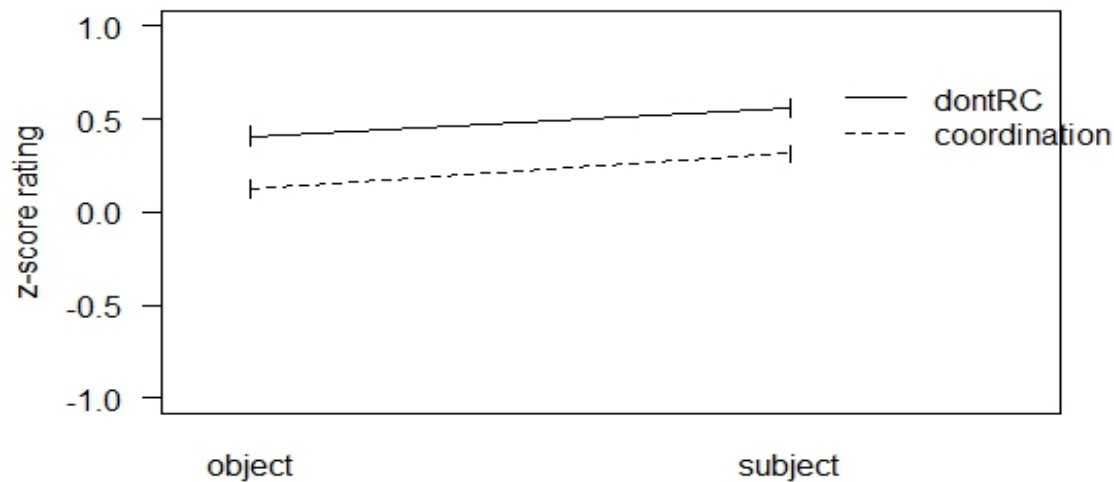
# English, French and Italian results

## Relative clauses

### English relative clauses



### French relative clauses

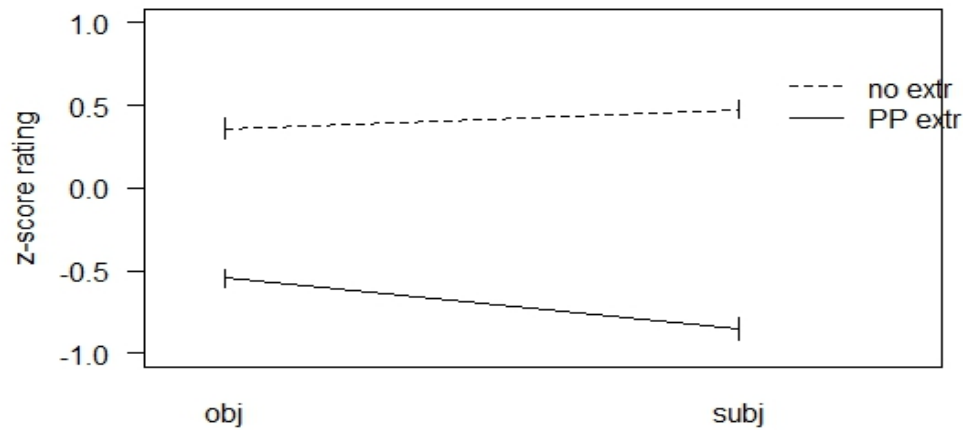


Italian  
Sprouse et al. (2016)

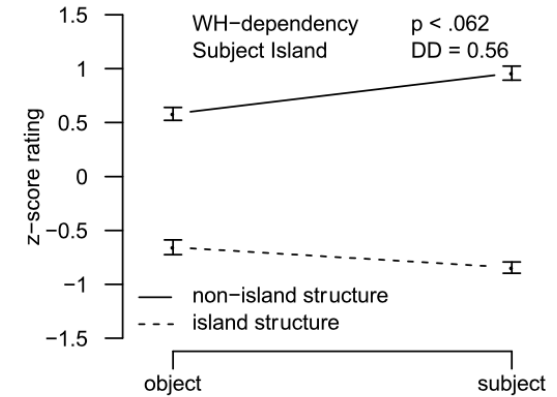
# English, French and Italian results

## Wh-questions

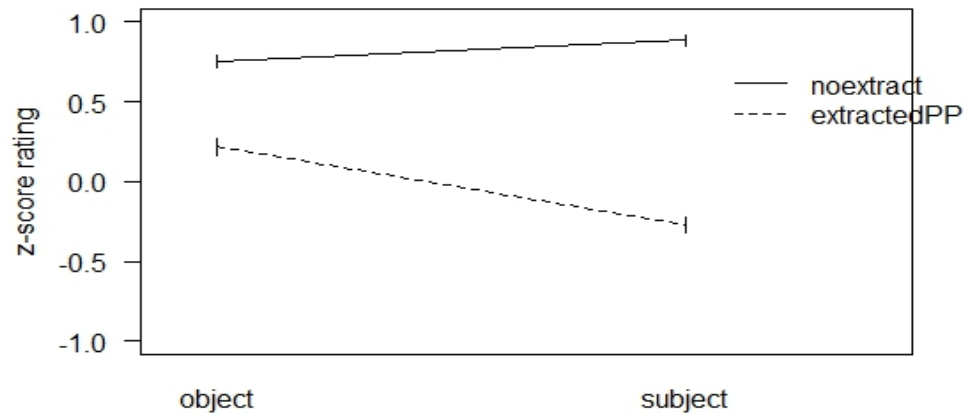
English  
Interrogatives with pied-piping



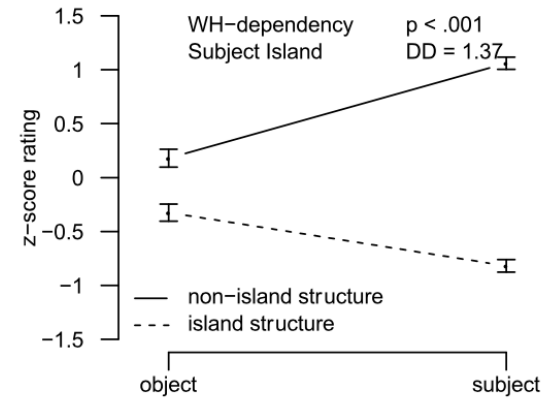
English  
Sprouse et al. (2016)



French questions



Italian  
Sprouse et al. (2016)



Results from Abeillé et al. (in progress)

# Project I: Subject-islands: **Conclusions**

Anne Abeillé; Barbara Hemforth, Elodie Winckel, Ted Gibson

**The Focus constraint:** It is infelicitous to question an element out of a backgrounded constituent

Wh-questions: function of extracted element in upper clause: focus (new)  
extraction domain: the subject is a natural topic

**Thus: the subject extraction penalty in wh-questions is due to a clash of discourse status of the extracted noun and its extraction domain**

Relative clauses: function of extracted element in upper clause: any discourse status (not focus)  
extracted element: local topic (inside the RC: what the RC is about)

So no penalty for out of subject extraction (or object extraction)

The contrast in judgements between WH-Qs and RCs is contrary to syntax-only accounts

This is the first example of an acceptability judgement difference across extractions from Wh-questions and RCs (also observed by Sprouse et al. but not noted for what it is)



# Project I: Subject-islands: **Methodological conclusions**

Anne Abeillé; Barbara Hemforth, Elodie Winckel, Ted Gibson

1. Without experiments, you can't investigate island structures (inherently interactions)
2. Need controlled materials across languages
3. Experiments aren't sufficient: Still need to avoid confounds (p-stranding vs. no p-stranding)

We find here remarkable similarity across English and French, which also looks very similar to Italian

Only difference appears to be p-stranding in English

## Project 2: S-complement islands

Yingtong Liu, Rachel Ryskin, Richard Futrell, Ted Gibson

Contrast in acceptability of extraction across different S-complement verbs:

“bridge” verbs: (“think”, “say”, “believe”)

What did Mary say that Bill liked\_\_?      What did Mary think that Bill liked\_\_?

Manner-of-speaking verbs:

??What did Mary whisper that Bill liked\_\_?

(cf. Mary whispered that Bill liked wine.)

Factive verbs

??What did Mary know that Bill liked\_\_?

(cf. Mary knew that Bill liked wine.)

Syntax-only account: Different structures: adjuncts for manner of speaking (Baltin, 1982); embedded NPs for factives (Kiparsky & Kiparsky, 1971)

# Maybe frequency can explain these differences?

Perhaps bad S-complement extractions are just low frequency S-complements? (Dabrowska, 2004;Verhagen, 2006, cf. Trueswell et al. 1993; MacDonald et al. 1994)

e.g., “think”, “say”: high frequency S-comp verbs

“whisper”, “moan”: low-frequency S-comp verbs

Ambridge & Goldberg (2008): this is unlikely because factives are high-frequency S-complement verbs (e.g., “know”), and extractions are difficult out of factive verbs

# Information-Structure based accounts Ambridge & Goldberg, 2008

Perhaps bad extractions are cases where *backgrounded information* is extracted.

***Backgrounded constructions are islands*** (BCI) account: *Backgrounded constituents may not serve as gaps in filler-gap constructions.*

A relative clause (RC) is backgrounded information. So extraction from RCs is disallowed:

\*Who did Mary see the report that was about \_?

Sentential subjects are backgrounded, so no extraction from sentential subjects:

\*Who did [that she knew] bother him? (Goldberg 2013)

\*What did [Bill's stealing ] shock Sue? (Ambridge et al., 2014)

# Information-Structure based accounts Ambridge & Goldberg, 2008

***Backgrounded constructions are islands*** (BCI) account: *Backgrounded constituents may not serve as gaps in filler-gap constructions.*

Claim: In manner of speaking verbs, the main assertion is in the manner, not the embedded clause. So extraction should be difficult, except in supportive contexts, where manner is already mentioned. Then extraction should be ok.

In factive verbs, the complement of the factive verb is presupposed, so it should be impossible to extract from.

Prediction: 3 classes of Verbs:

Factive verbs (realize, remember, notice, know): should not allow extraction

manner-of-speaking verbs (whisper, stammer, mumble, mutter): less likely to allow extraction

‘bridge verbs’ (say, decide, think, believe): most likely to allow extraction

# Information-Structure based accounts

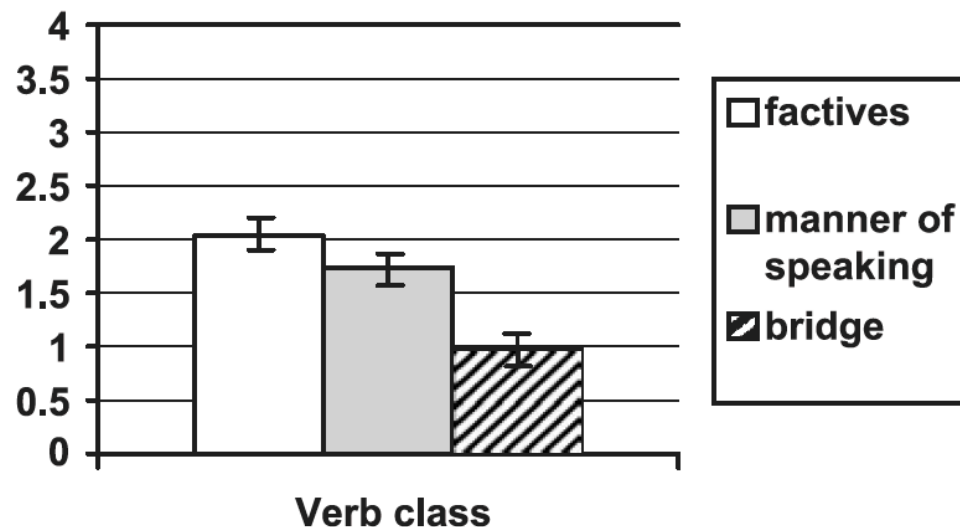
## Ambridge & Goldberg, 2008

Experiment: Acceptability of wh-question vs. declarative

Factive verbs (realize, remember, notice, know): should not allow extraction

manner-of-speaking verbs (whisper, stammer, mumble, mutter): less likely to allow extraction

‘bridge verbs’ (say, decide, think, believe): most likely to allow extraction



*Mean difference (dispreference-for-extraction-question) scores. Higher scores indicate greater ungrammaticality of the question form (relative to the corresponding declarative)*

# Attempted replication and extension of Ambridge & Goldberg (2008): Liu, Ryskin, Futrell & Gibson (in progress)

Replication: Rate acceptability of wh-extraction, and declarative control  
1-5 scale; 24 verbs, in 3 classes; 4 versions of each (96 items); 120 AMT participants

Bridge example

Declarative: Maria said that Frank bought something

Wh-question: What did Maria say that Frank bought?

Manner example:

Declarative: Maria whined that Frank bought something.

Wh-question: What did Maria whine that Frank bought?

Factive example:

Declarative: Maria knew that Frank bought something.

Wh-question: What did Maria know that Frank bought?

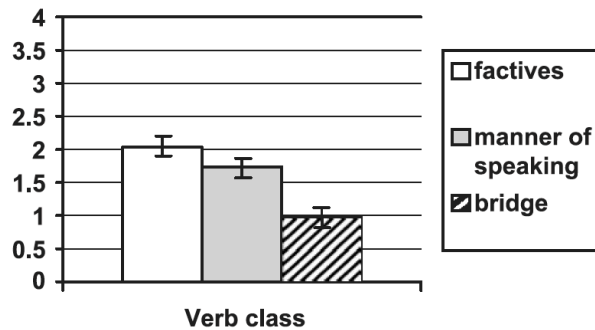
Prediction: Difference in the wh-questions, but not in the declaratives: **interaction**  
between sentence type (WHQ, Decl) and verb type

# Liu, Ryskin, Futrell & Gibson (in progress)

## Experiment I

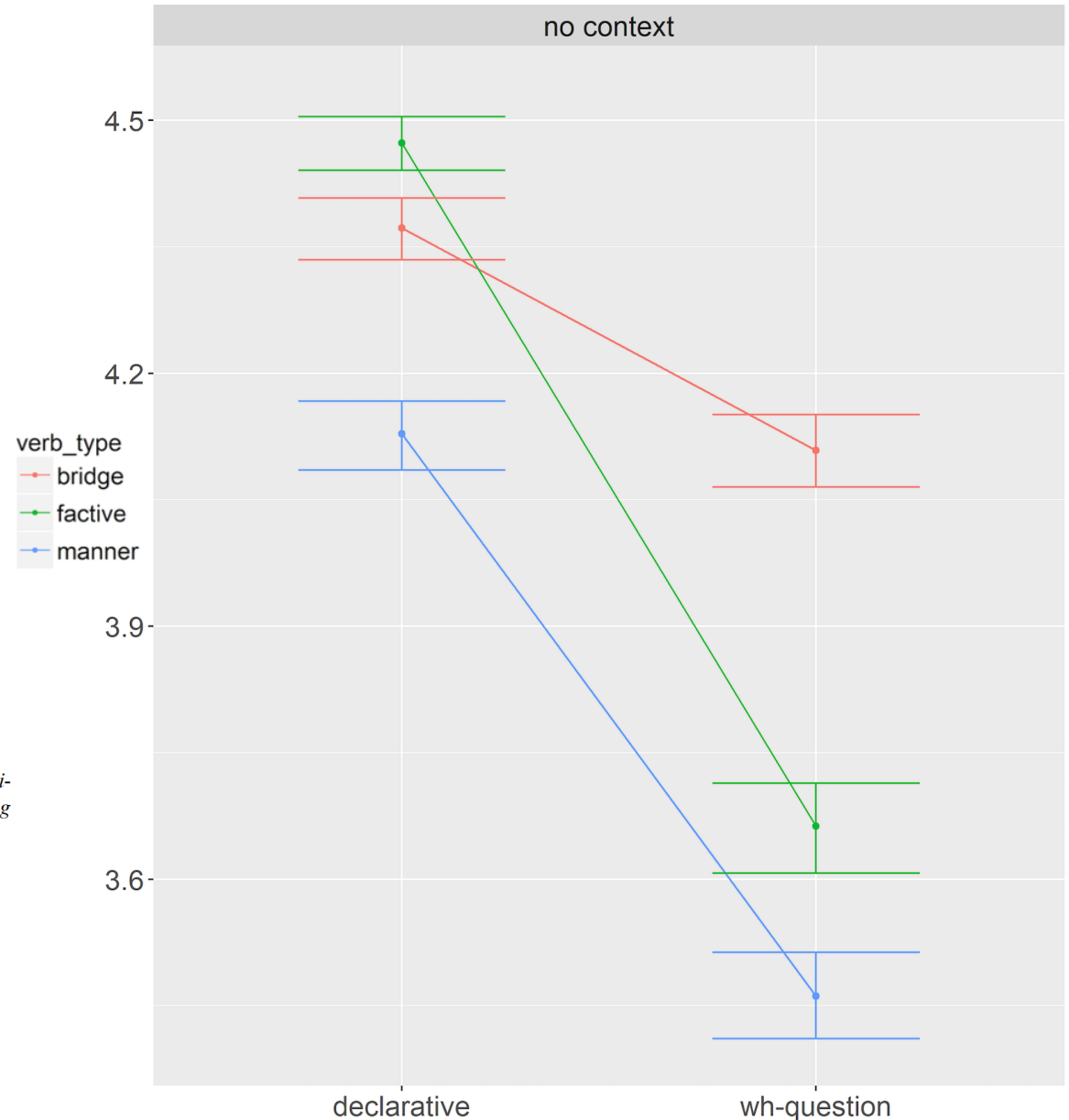
Prediction 1 (replication): 3 classes of verbs, factive worst; manner intermediate; bridge best

Yes, in the difference scores: But similar effects in the declaratives as whqs



Mean difference (dispreference-for-extraction-question) scores. Higher scores indicate greater ungrammaticality of the question form (relative to the corresponding declarative)

Is the normality assumption correct, in these data? No: near top of scale for declarative. So interaction is suspect. (cf. Jaeger, 2008, for binary data analyzed by linear regressions)





# Liu, Ryskin, Futrell & Gibson (in progress)

## Experiment I

Attempted analyses of the data with ordinal regression don't converge well: probably not enough data

So convert the ordinal 1-5 data to a binary scale and analyze with logistic regression

1,2,3 = unacceptable

4,5 = acceptable

```
model: rating_transformed1 ~ sentence_type*log_fre+(1+sentence_type*log_fre|
WorkerId)+(1+sentence_type|matrix_verb)
```

```
result:
```

```
result:
```

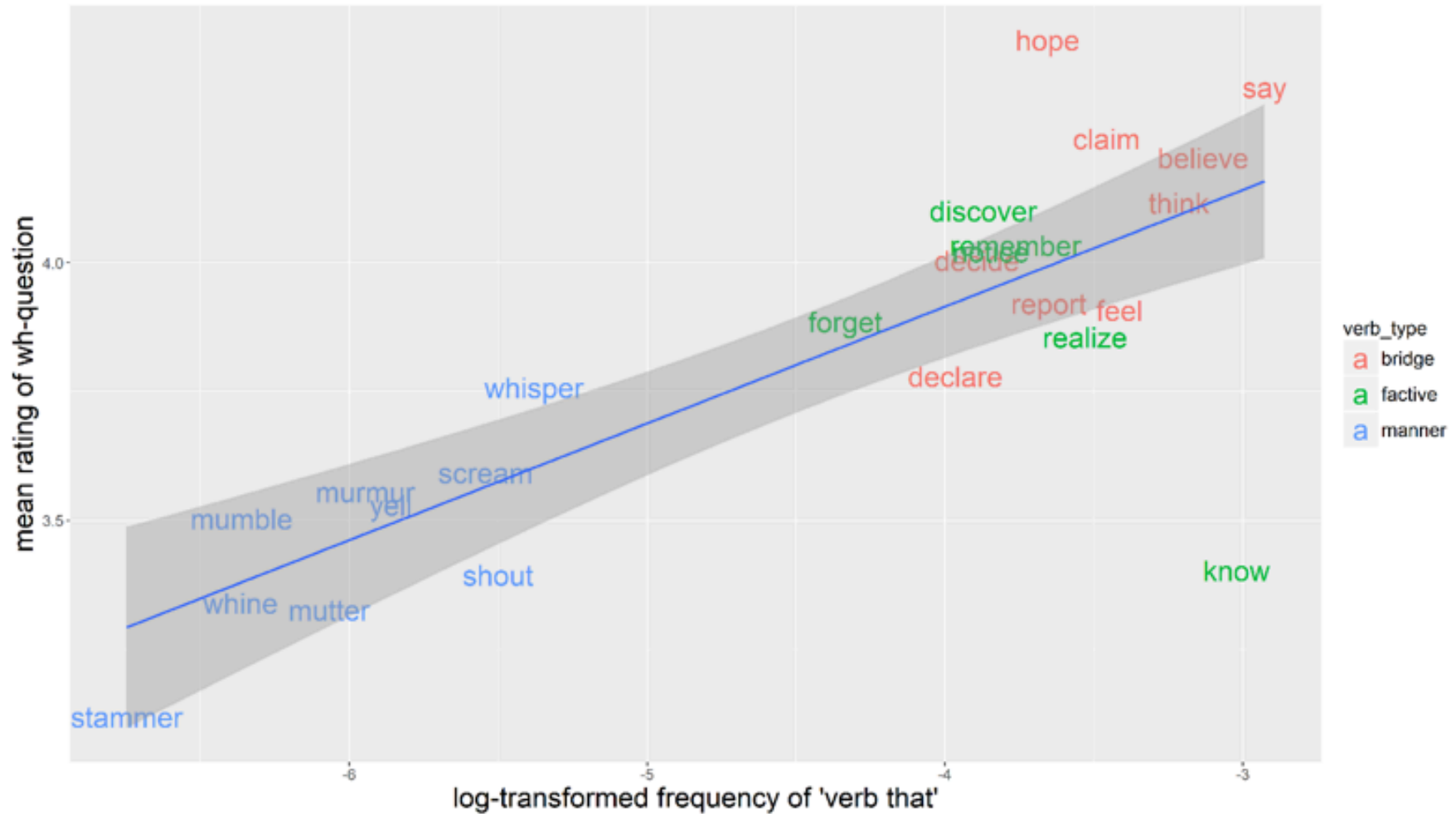
```
# Fixed effects:
```

#	Estimate	Std. Error	z value	Pr(> z )
# (Intercept)	6.97773	0.91984	7.586	3.3e-14 ***
# sentence_typeint	-2.44451	0.99772	-2.450	0.0143 *
# log_fre	0.49077	0.17860	2.748	0.0060 **
# sentence_typeint:log_fre	-0.08659	0.19694	-0.440	0.6601

no significant interaction! Just two main effects

# Liu, Ryskin, Futrell & Gibson (in progress)

## Experiment I: Verb frequency vs. acceptability scores



A strong relationship between frequency and wh-question acceptability (Spearman rho = .80;  $p < .001$ )

**NOTE: lots of overlap between bridge and factive**

# Liu, Ryskin, Futrell & Gibson (in progress)

## Experiment I

Simple theory of difficulty of extracting across manner and factive verbs:

Verb frequency + construction frequency

People learn verbs and constructions with exposure

So WH-questioning across an infrequent verb is more difficult

# Liu, Ryskin, Futrell & Gibson (in progress)

## Experiment 2

### **Is there something different about “know”?**

Function of a wh-question: to find out what your conversation partner **knows** about.

Therefore it is often pragmatically odd to ask what someone else knows about certain events:

if that person knows it, then it is simpler to ask directly about the events.

e.g. What did Susan know that Anthony liked?

in most contexts, it is simpler to ask “what did Anthony like?”

Test of this idea: Switch to clefts or relative clauses. The meaning of a cleft / relative clause is that both speaker and hearer already know this information.

# Liu, Ryskin, Futrell & Gibson (in progress)

**Experiment 2:** Rate acceptability of clefts and declarative controls

Factive example

**Declarative:** Susan knew that Anthony liked the pie.

**Cleft:** It was the pie that Susan knew that Anthony liked.

Manner example:

**Declarative:** Angela mumbled that Kevin took the necklace.

**Cleft:** It was the necklace that Angela mumbled that Kevin took.

Bridge example:

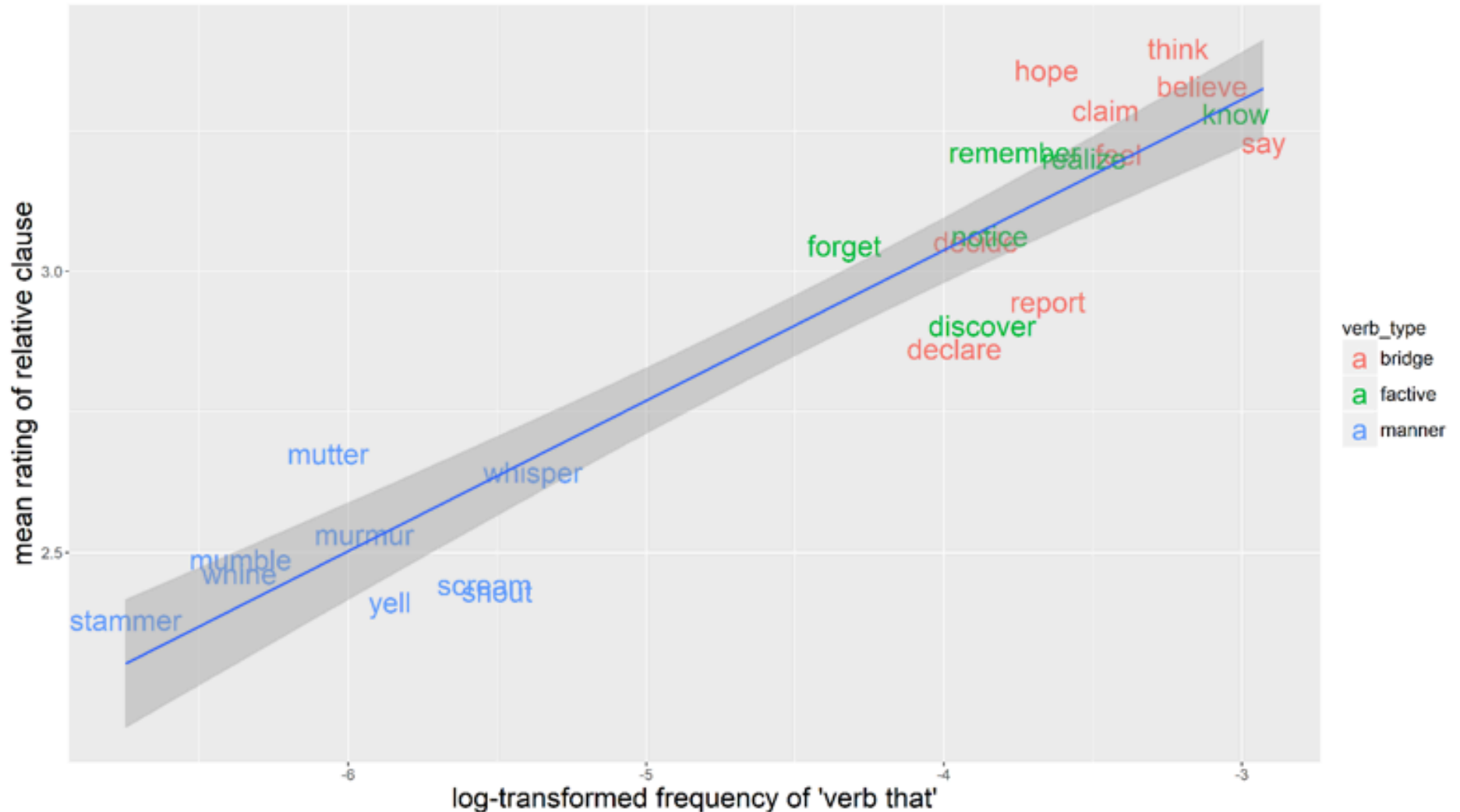
**Declarative:** Amanda believed that Jeffrey took the bread.

**Cleft:** It was the bread that Amanda believed that Jeffrey took.

As in WHQs, when analyzed as logistic regression, two main effects, no interaction.

# Frequency explains cleft extraction complexity

Liu, Ryskin, Futrell & Gibson (in progress)



A strong relationship between frequency and cleft acceptability (Spearman rho = .93;  $p < .001$ )

“know” is no longer an outlier

A similar effect in the declarative versions

# Liu, Ryskin, Futrell & Gibson (in progress)

## Experiment 3

Like Experiment 1, but with more verbs:

48 verbs; 6 versions of each (288 items); normed for plausibility first

120 AMT participants

Binary rating: Acceptable / unacceptable

High frequency example

Declarative: **Cindy said that Brian sold something.**

Wh-question: **What did Cindy say that Brian sold?**

Medium frequency example:

Declarative: **Christine concealed that David bought something.**

Wh-question: **What did Christine conceal that David bought?**

Low frequency example:

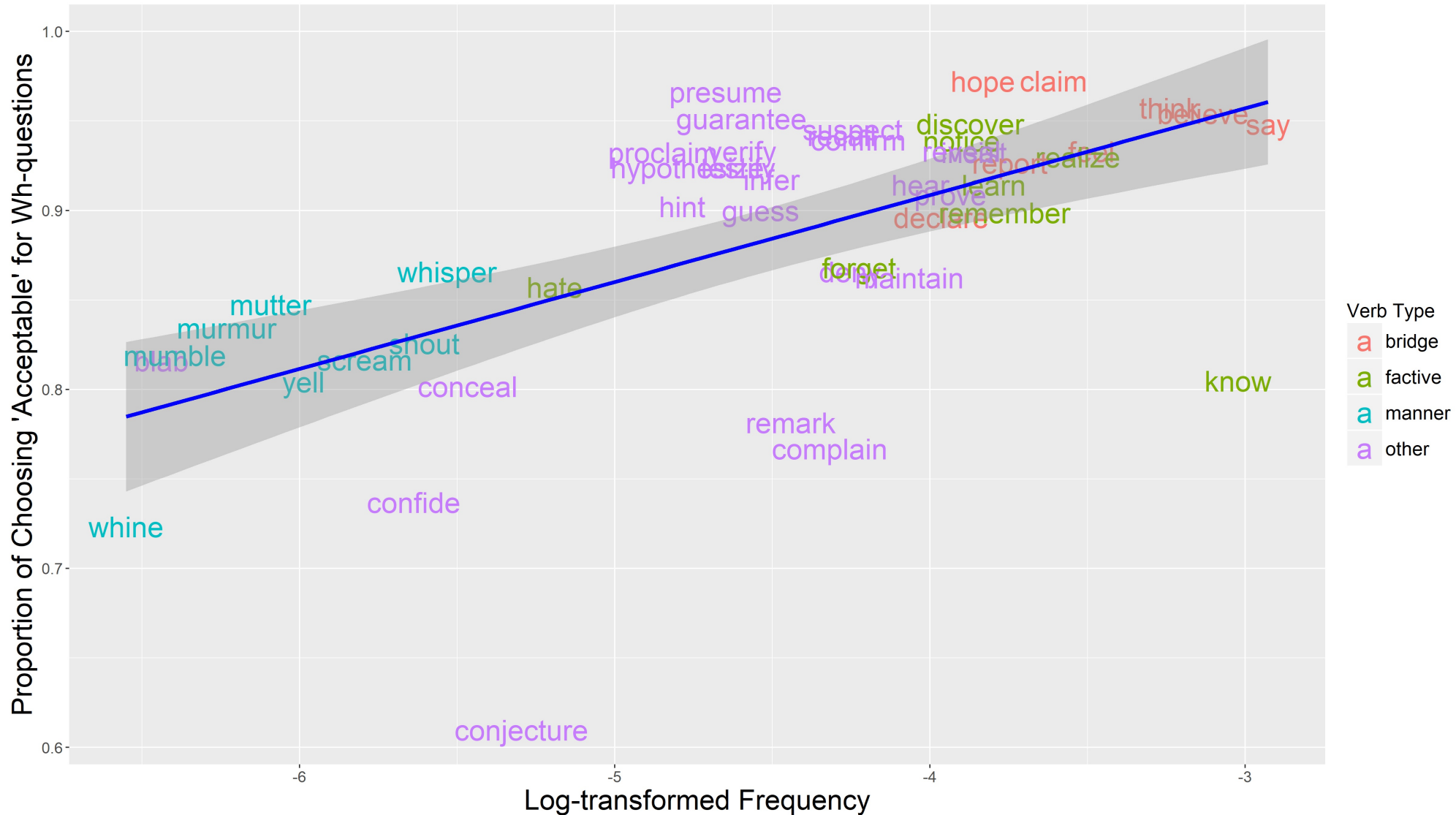
Declarative: **Mary whined that William bought something.**

Wh-question: **What did Mary whine that William bought?**

Prediction: Main effects of verb frequency and sentence type (Declarative / WHQ)

# Frequency explains WH-Q extraction complexity

Liu, Ryskin, Futrell & Gibson (in progress)



A strong relationship between frequency and WHQ acceptability (Spearman rho=0.60;  $p < .001$ ) (similar figure for declarative)



# Liu, Ryskin, Futrell & Gibson (in progress)

## Summary

Conclusion from these studies:

Verb frequency + construction frequency explains most of the variance

no interaction between structure and verb type, so no “islands” at all

# Summary and conclusions

1. Without experiments, you can't investigate island structures (inherently interactions)
2. Need controlled materials across languages
3. There is a dissociation between acceptability of extractions from WH-questions and relative clauses in English and French
4. A remarkable similarity across English, French, and Italian: the subject extraction penalty in wh-questions is due to a clash of discourse status of the extracted noun and its extraction domain.
5. S-complement islands: verb and construction frequency explain the unacceptability of such examples

**Overall conclusion: Maybe all “islands” are meaning- and memory-based, contrary to the UG syntax claim.**