

On subject islands

Rui P. Chaves

*Department of Linguistics
University at Buffalo – SUNY*

June 28th, 2012

**<< Ling Lunch >>
Paris Diderot**

Overview

- 1 Background
- 2 Subject islands
- 3 New evidence
- 4 Proposal
- 5 Conclusions

Overview

- 1 Background
- 2 Subject islands
- 3 New evidence
- 4 Proposal
- 5 Conclusions

The basic contrast

- (1) a. A book about war is ultimately an essay on human nature.
- b. *What is a book about ultimately an essay on human nature?
- c. ?What is a book about war ultimately an essay on?

Specified Subject Constraint (Chomsky'77)

- (2) *Which cars were the hoods of damaged by the explosion?
(Ross'67)
- (3) *Who did stories about terrify John?
(Chomsky'77)
- (4) *Who was a picture of lying there?
(Kayne'81)
- (5) *Who do you think pictures of would please John?
(Huang'82)
- (6) *Which books did talking about become difficult?
(Cinque'90)
- (7) *Who did my talking to bother Hilary?
(Pollard & Sag'94)
- (8) *Which book did a review of appear in the Times?
(Jackendoff'02)
- (9) *Which candidate were posters of all over the town?
(Lasnik & Park'04)

Extraction from objects is (claimed to be) less restricted

(Ross'67, Kuno'76, McCawley'81, Deane'91, ...):

- (10) a. Who did you see a picture of?
- b. Who did you hear stories about?
- c. The man who I read a statement about was sick.
- d. Nixon was one president that they had no trouble getting votes for the impeachment of.

- (11) a. How much money are you making the claim that the company squandered?
- b. Which rebel leader did you hear rumors that the CIA assassinated?
 (Pollard & Sag'94)

Sentential Subject Constraint (Ross'67)

- (12) a. *Who did that Bill married surprise you?
b. *The hat which that I brought seemed strange to the nurse was a fedora.
(Ross'67)
c. *Who did he say that for Bill to marry was a surprise?
(Huang'82)
d. *Who does the claim that Mary likes upset Bill?
(Lasnik & Saito'92)

Not universal (with caveats)

The following languages are known to allow extraction from subject phrases:

- Swedish
- Turkish
- Russian
- Japanese
- Hindi
- German

See Stepanov'07 for an overview.

Some recent accounts:

- SYNTAX

- **Takahashi'94**

- Chains must be uniform.
 - Make the shortest move.

- **Nunes & Uriagereka'00**

- Linearizing X and Y requires an asymmetric c-command.
 - Once X is linearized, its internal structure is opaque.

- PRAGMATICS

- **Erteschick-Shir'07**

- SUBJECT_{top} [... X ...]_{loc}
 - Extraction must target FOC.

- PROCESSING

- **Kluender'04**: cognitive resource limitations.

The performance-based story (roughly)

- Acceptability = $\frac{1}{\text{Processing difficulty}}$
- Processing difficulty =
Storage + Intervener processing + Retrieval

Evidence (for a variety of different constructions):

- referential processing of noun phrases and verbs incurs a processing cost.
(Gibson'98'00)
- clause boundaries incur a processing cost, and depend on the complementizer type.
(Kluender & Kutas'93)
- unspecific fillers (e.g. *who*) are more subject to memory decay, and therefore are harder to retrieve than more specific fillers (e.g. *which fireman*)
(Hofmeister'07, Hofmeister & Sag'10)

Independent motivation: acceptability of center embedding.

- (13) a. The boy started crying.
b. The boy the cat scratched started crying.
c. The boy the cat the dog bit scratched started crying.

The intelligibility of double center-embeddings can improve:

- (14) a. The bees that the hives that the farmer built housed stung the children.
(Stolz'67)
b. The rat that the cat I saw chased squeaked.
(Kac'81)

Some cases are fully acceptable, others are incomprehensible:

- (15) a. The movie everyone I know loved was *Inception*.
(Chaves'12)
b. People people people left left left.
(Rogers & Pullum'11)

Typical traits of performance-based low acceptability

- Graded acceptability.
- Can be additively ameliorated by
 - processing-light elements (e.g. pronouns)
 - prosodic cues
 - contextual bias
 - reactivation
 - coherence
- Set of exceptions do not form a natural grammatical class.

- 1 Background
- 2 Subject islands**
- 3 New evidence
- 4 Proposal
- 5 Conclusions

Complication 1

Subject Islands appear to be selective:

- (16) a. Of which cars were the hoods damaged by the explosion?
(Ross'67)
- b. That is the lock to which the key has been lost.
- c. A house of which only the front has been painted will be on your left at the second light; you make a right turn there.
(Levine & Hukari'06)
- d. Of which car was the driver awarded a prize?
(Chomsky'08)

Attested object PP extractions from subjects

(17) They have eight children [of whom] five _ are still living at home.

(Huddleston & Pullum'02)

(18) a. Already Agassiz had become interested in the rich stores of the extinct fishes of Europe, especially those of Glarus in Switzerland and of Monte Bolca near Verona, [of which], at that time, only [a few _] had been critically studied.

b. (...) a letter [of which] [every line _] was an insult (...)
(Santorini'81)

Complication 2

Parasitism (Ross'67,Engdahl'83,...)

- (19) a. *What did the attempt to repair ultimately damage the car?
b. What did the attempt to repair ultimately damage?
- (20) a. *Who did the rivals of shoot Castro?
b. Who did the rivals of shoot?

Problems for parasitism as null pronominal resumption

- Co-reference is not necessary:

(21) There are certain heroes that long stories about are always very easy to listen to.

(Jacobson'94, Pollard & Sag'94)

- Non-nominal fillers:

(22) This is a table ON WHICH anyone who puts some books must subsequently put some magazines as well.

(Levine & Hukari'06)

- Non-parasitic extractions:

(23) a. What were pictures of seen around the globe?

(Kluender'98)

b. That piano, which the boy's loud playing of drove everyone crazy, was badly out of tune.

(Ross'67)

Complication 3

Non-parasitic exceptions

- For Levine'01 and Levine & Sag'03 (24) is **passable**:
(24) There are certain topics that jokes about are completely unacceptable.
- For Kluender'98/04 (25) is **acceptable**:
(25) What were pictures of seen around the globe?

In sum, any theory needs to explain:

- island effects (and graded acceptability)
- selectivity effects
- parasitism effects
- non-parasitic exceptions

Kluender'04: complex subjects impose greater processing costs than complex complement phrases.

- the longer the subject the shorter the verb phrase, and vice versa. (Bloom'90, Kemper'87)
- adults between 70 and 90 years of age have more difficulty repeating sentences with complex subjects than with complex objects (Kemper'86)
- subject phrases are subject to more disfluencies than object phrases (Clark & Wason'87)
- processing open-class words at the beginning of sentences entails greater processing effort (N400 amplitudes). (Garnsey'85, Kutas et al.88, Petten & Kutas'91)
- subject phrases with complex subjects lead to longer speech initiation times. (Ferreira'91, Tsiamtsiouris & Cairns'09)
- center-embeddings in subjects are harder to process than center-embeddings in objects.
(Amy & Noziet'78, Eady & Fodor'81)

Kluender'04:

- Has no explanation for parasitism.
- Is silent about how selectivity emerges.
- How exactly can the contrast of (26a) be due to cognitive resource (memory) limitations?
(26) a. *Who was a picture of posted online?
 b. Who did Kim post a picture of online?

- 1 Background
- 2 Subject islands
- 3 New evidence**
- 4 Proposal
- 5 Conclusions

Acceptable Specified Subject Island violations:

- (27) a. [Which president] [would the impeachment of] [cause outrage]?
- b. [Which movie] [will the soundtrack of] [never be popular]?
- (28) a. [Which problem] [will no solution to] [ever be found]?
- b. [Which disease] [will the cure for] [never be found]?
- c. [Which question] [will the answer to] [never be known]?
- d. [Which word] [will the misspelling of] [never be noticed]?
- e. [Which poison] [will the antidote to] [never be discovered]?
- f. [Which book] [will the author of] [never be known]?

Hypothesis

If Subject Island effects are performance and/or pragmatic conditions, then they might be attenuated.

Three factors which may be relevant

- 1 Filler specificity (... , Kluender'98, Hofmeister'07)

(29) The lieutenant could not remember [which soldier/person] the commander that was deeply respected ordered _ to scout the area ahead.

- 2 Relevance (Erteschik-Shir'73, Kuno'87, Deane'92)

(30) a. Who did you write a book about?
b. ?Who did you lose a book about?

- 3 Prosodic phrasing (see Fodor'02 for overview and discussion)

(31) That wine was found in my car surprised me.

Experiment

Sentence acceptability study

Filler specificity × Subject definiteness

Design:

2×2 design, 40 quadruples of sentences, 4 lists

80 distractors (3 kinds), 43 paid participants

Method: Magnitude Estimation.

- i. [SPECIFIC FILLER / DEFINITE SUBJECT]
Which politician did the opponents of organize a protest?
- ii. [SPECIFIC FILLER / INDEFINITE SUBJECT]
Which politician did opponents of organize a protest?
- iii. [NON-SPECIFIC FILLER / DEFINITE SUBJECT]
Who did the opponents of organize a protest?
- iv. [NON-SPECIFIC FILLER / INDEFINITE SUBJECT]
Who did opponents of organize a protest?

LMER results

– specific fillers more acceptable than non-specific fillers

($\beta = -0.086$, $t = -4.369$, $p < 0.0001$).

– definiteness not significant

($\beta = -0.003$, $t = -0.162$, $p = 0.871$).

– no specificity/definiteness interaction

($\beta = -0.019$, $t = -0.712$, $p = 0.476$).

LMER results

– specific fillers more acceptable than non-specific fillers

($\beta = -0.086$, $t = -4.369$, $p < 0.0001$).

– definiteness not significant

($\beta = -0.003$, $t = -0.162$, $p = 0.871$).

– no specificity/definiteness interaction

($\beta = -0.019$, $t = -0.712$, $p = 0.476$).

Mechanical Turk replication: 7-point Likert-scale methodology.

92 participants, 20 pairs items, 2 lists.

LMER: 757 data points ($\beta = 0.65$, $t = 8.25$, $p < 0.001$).

Satiation ('linguist's disease')

- CNPC violations can sometimes satiate (Snyder'00, Goodall'05, Sprouse'12)
- *Whether* island violations always satiate (Snyder'00, Hiramatsu'00, Braze'02)
'What do you wonder whether John likes?'

Satiation in Subject Island violations is controversial:

- Significant (Hiramatsu'00a, Francom'09)
- Marginally significant (Snyder'00)
- Not significant (Sprouse'09, Hiramatsu'00b, Crawford'11)

Satiation ('linguist's disease')

- CNPC violations can sometimes satiate (Snyder'00, Goodall'05, Sprouse'12)
- *Whether* island violations always satiate (Snyder'00, Hiramatsu'00, Braze'02)

'What do you wonder whether John likes?'

Satiation in Subject Island violations is controversial:

- Significant (Hiramatsu'00a, Francom'09)
- Marginally significant (Snyder'00)
- Not significant (Sprouse'09, Hiramatsu'00b, Crawford'11)

Satiation found in both experiments:

ME experiment: $\beta = 0.01$, $t = 6.66$, $p < 0.001$.

MT Likert replication: $t = 4.415$, $p < 0.0001$.

Filled-gap effect

A.k.a. 'Active Filler Hypothesis' (Frazier, 1987)

Comprehenders tend to posit a gap as soon as possible:

(32) What did you say _

Filled-gap effect

A.k.a. 'Active Filler Hypothesis' (Frazier, 1987)

Comprehenders tend to posit a gap as soon as possible:

(32) What did you say _ John thinks _

Filled-gap effect

A.k.a. 'Active Filler Hypothesis' (Frazier, 1987)

Comprehenders tend to posit a gap as soon as possible:

(32) What did you say _ John thinks _ I read _

Filled-gap effect

A.k.a. 'Active Filler Hypothesis' (Frazier, 1987)

Comprehenders tend to posit a gap as soon as possible:

(32) What did you say _ John thinks _ I read _ a lot about _?

Filled-gap effect

A.k.a. 'Active Filler Hypothesis' (Frazier, 1987)

Comprehenders tend to posit a gap as soon as possible:

(32) What did you say _ John thinks _ I read _ a lot about _?

This filled-gap effect **does not exist for subject phrases**.

(Stowe'86, Pickering et al.'94, Frazier'87, Ellis'91)

Phillips'06 finds a filled gap effect in **infinitival** subjects:

- (33) a. *The outspoken environmentalist worked to investigate what the local campaign to preserve _ had harmed the annual migration
- b. *The outspoken environmentalist worked to investigate what the local campaign that preserved _ had harmed the annual migration

Attested Sentential Subject violations

- (34) a. There are people in this world that to describe as despicable would be an understatement.

(<http://www.thegatewaypundit.com/2011/10/obama-weekly-address-millionaires-need-to-pay-the-same-tax-rate-as-plumbers-or-bus-drivers/>)

- b. Chief Magistrate Elizabeth Bolton said to describe Clohesy's behaviour as despicable would be an understatement.

(<http://www.intmensorg.info/australia2.htm>)

- c. In his bedroom – which to describe as small would be a gross understatement – he has an audio studio setup.

(<http://pipl.com/directory/name/Frohwein/Kym>)

- d. They amounted to near twenty thousand pounds, which to pay would have ruined me.

(Benjamin Franklin et al. 1834. *Memoirs of Benjamin Franklin*, vol 1. p.58)

Summary: subject island acceptability

- 1 has graded acceptability;
... in some cases, acceptability is fairly **high**.
- 2 can exhibit satiation;
- 3 hinges on pragmatic and prosodic factors.

- 1 Background
- 2 Subject islands
- 3 New evidence
- 4 Proposal**
- 5 Conclusions

Hypothesis

Subject island violations are grammatical, but speakers avoid sentences with subject-internal gaps.

REASONS:

Complexity avoidance:

- Complex subject phrases are harder to process than complex object phrases.
- Discourse contexts for which subject-internal gaps would be felicitous are necessarily very complex.
- P-stranding causes the grammatical function of a NP filler to be more difficult to determine than of a PP filler.

Pragmatic restrictions:

- Odds are that most complex subject phrases will not satisfy filler-verb relevance constraints.

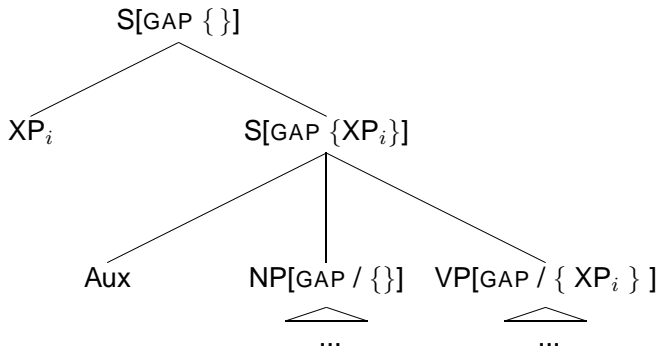
Consequences

- 1 Extreme low frequency of subject-internal gaps leads to a grammatical heuristic: *expect gap-less subjects*.

verb-lexeme $\Rightarrow / \left[\text{ARG-ST} \left\langle \left[\text{GAP } \{\} \right], \dots \right\rangle \right]$

(cf. with Fodor'78, Berwick & Weinberg'84, Hawkins'99)

- 2 Corollary: the gap is expected to be in the VP.



Explaining the phenomena:

- when the gap is not subject-internal, the heuristic guides the parser to the correct parse. The heuristic works!

(35) Who did the rivals of Castro shoot _ ?

Explaining the phenomena:

- when the gap is not subject-internal, the heuristic guides the parser to the correct parse. The heuristic works!

(35) Who did the rivals of Castro shoot _ ?

- when there are two gaps (one subject-internal the other VP-internal), only one expectation is violated.

(36) Who did the rivals of _ shoot _ ?

Explaining the phenomena:

- when the gap is not subject-internal, the heuristic guides the parser to the correct parse. The heuristic works!

(35) Who did the rivals of Castro shoot _ ?

- when there are two gaps (one subject-internal the other VP-internal), only one expectation is violated.

(36) Who did the rivals of _ shoot _ ?

- when the gap is subject-internal, two expectations are violated. This creates a strong garden-path due to a 'digging-in' effect **unless** the correct parse is cued by prosody, specific *wh*-phrases, and and highly relevant filler-gap verb dependencies.

(37) *Who did the rivals of _ shoot Castro?

Why are PP Subject island violations usually better than NP Subject island violations?

- (38) a. Of which cars were the hoods _ damaged by the explosion?
b. *Which cars were the hoods of _ damaged by the explosion?
(Ross'67)

Less ambiguity: a PP filler has information about the grammatical role of the phrase, whereas an NP filler does not.

Further evidence for processing-based rules:

- 1 language-specific relative clause attachment preferences.
 - High attachment preference
(Spanish, French, Japanese)
 - Low attachment preference
(Arabic, English, Norwegian, Romanian, Swedish)
- 2 NP and PP ordering in English VPs
 - NP-PP (canonical)
 - PP-NP (when NP is significantly longer)

Wasow'09: this canonical but violable ordering tendency should be seen as part of the grammar of English.

- 1 Background
- 2 Subject islands
- 3 New evidence
- 4 Proposal
- 5 Conclusions**

Conclusions

Subject Islands

- May be due to a processing heuristic that reflects the very low frequency of subject-internal gaps, caused by their processing complexity and restrictive pragmatics.
- This is consistent with:
 - graded acceptability;
 - selective satiation;
 - selective filled-gap effects;
 - NP/PP filler contrasts;
 - parasitism effects;
 - non-parasitic exceptions;

Acknowledgement: experimental research funded by Professional Development Award Program grants, awarded by the New York State/United University Professions Joint Labor-Management Committees.