Paul Deane: “Limits to attention: a cognitive theory of island phenomena”
-- Highlights handout prepared by M. Wagers, UCSC
-- LING229/Seminar in Syntactic Theory/Islands, 24 February 2010

(1) **A cognitive linguistics treatment of island constraints**
Island constraints are not abstract conditions on structures or transformation. They directly reflect constraints on “general cognitive processing.”

(2) **Specific thesis**: Island constraints reflect limits on *attention.*
-- Extraction, by Deane’s hypothesis, requires paying simultaneous attention to two portions of a sentence.
-- Island violations occur when other parts of the sentence structure divert the limited resource of attention.

(3) **Data & Analysis:**
-- a wide-range of patterns in extraction from NP modifiers
-- the exceptions to the Coordinate Structure Constraint
-- pragmatic control of variation

**Complex NPs**

(4) (a) Which book did you see pictures of?
   *Which book did you destroy pictures of?

   *Potential solution*: reanalysis \[VP V [NP N PP] \] \[VP V NP PP \]
   cf. Ross’s ‘Modal’ constructions

   *Problem*: how do you characterize when reanalysis occurs?

   (b) What does he see a picture of?
   (c) *What does he see a book about?

   (d) *It’s the club foot that I know the man with
   (e) It’s John that I enjoy a good fight with

(5) Bolinger (1972): *characterization*  
an NP can be extracted if it ‘characterizes’ the head of the NP that contains it

   (a) Which house do you own the furniture in?
   (b) *Which garage do you own the car in?
   (c) It’s Jordan’s I know the people at
   (d) *It’s the front door that I know the people at

(6) Cattell (1979): *part-whole licensing*  
an NP can extracted if head is a part, NP is the whole

p. 1/10
(a) Which car do you like the gears in?
(b) *Which car do you like the girls in?

(7) Kuno (1987): attribution (head N is an ‘attribute’ of extracted NP)

(a) This is the kind of solution which I fully recognize the defect with
(b) *This is the kind of defect that I don’t like solutions with

(8) Kuno (1987): context effects

(a) *Who did they destroy more pictures of?
(b) A: Right after Mao died, they started taking pictures of Committee members off the walls.
   B: Who did they destroy more pictures of, Chairman Mao or Jiang Qing?

   What does a pattern like this mean?
   (b') ?What did the museum vandals destroy paintings of?

(9) Kuno (1987): topicality of extracted NP

What aspects of political activities did you write a book on (about Nixon)?

(10) Deane:
Q: Are we using semantic/pragmatic support to nudge around marginal Ss?
   In cases above, we cross 2 bounding nodes. Suppose 2 = marginal.
   A: No. There are cases of ‘deep extraction’: 3+ bounding nodes are crossed.

(11) N = 3
(a) Which NPs are there unusual possibilities for extraction from?
(b) This is one newspaper that the editor exercises strict control over the publication of.
(c) Which laws do you advocate an end to the enforcement of?
(d) Which issues are you prepared to discuss the full range of opinions about?
(e) Which games have you discovered workable strategies for victory in?
(f) Which buildings do you have keys to the apartments in?
....

Ross (1967)

Which reports does the government prescribe the height of the lettering on?

(12) N = 4 Which committee did he have aspirations for appointment to the chairmanship of?

(13) N = 5 The ‘dowager’ story (p. 11, ex. 35)
... This is the only committee that I have seen fit to extend recognition to your aspirations for appointment to the chairmanship of

... This is the only committee that I have seen fit to extend recognition to your aspirations for appointment to the chairmanship of

cf.
My dear sir, I have not seen fit to extend recognition to your aspirations for appointment to the chairmanship of these committees.

Questions:

What is the structure of these NPs? Are they really stacked?

?? Is there a reverse length effect on acceptability ratings?

Deane points out the clear contrast with:

* Who did they destroy a book about?
* Who did you photocopy stories about?

COORDINATE STRUCTURE CONSTRAINT

Ross (1967); Goldsmith (1985); Lakoff (1986): ‘framing’

(14) What did Harry go to the store and buy ___? L’s ‘A’ scenario: Expected sequence, extraction from last conjunct

(15) How much can you drink ___ and still stay sober? ‘B’ scenario: counter-expectations extraction from first conjunct

(16) That’s the stuff the guys in the Caucasus drink and live to be 100

‘C’ scenario: cause-result

(17) (a) What kind of cancer can you eat herbs and not get ___? ‘A’
(b) What kind of herbs can you eat ___ and not get cancer? ‘B’

Are these not really coordinations?

(18) (a) What did he go to the store, buy ___ , load ___ in his car, drive home, and unload ___?
(b) How many courses can you take ___ for credit, still remain sane, and get all A’s in ___?
(d) This is the kind of brandy that you can sip ___ after dinner, watch TV for a while, sip some more of ___, work a bit, finish off ___, go to bed, and still feel fine in the morning.

(19) P-gaps?
(a) Sam is not the kind of guy you can just sit there, listen to ___ , and not want to punch ___ in the nose
(b) *Sam is not the kind of guy you can just sit there while listening to ___ without wanting to punch ___ in the nose
(MW: My judgment that there is a contrast isn’t so clear.)

**Semantic, Pragmatic and Functional Factors**

**Domain-Attribute**

<table>
<thead>
<tr>
<th>Extracted NP</th>
<th>Extraction Site NP Head</th>
</tr>
</thead>
</table>

Langacker (87): “cognitive domain”

- **(20)** (a) Which newspapers do we maintain strict editorial control over?
  (b) *How much editorial control do you publish a newspaper with?*

- **(21)** (a) Which girls did you notice shapely legs on?
  (b) *What kind of legs do you like girls with?*

- **(22)** (a) What products did you praise the high quality of?
  (b) How high a quality will you buy sugarcane of?

- **(23)** Q: Is this a complement/modification distinction?
  A: No, consider:
  (a) Which store did you buy [the furniture in]?
  (b) *Which crate did you buy [the furniture in]?

**Other Domain-Attribute Asymmetries: metonymy**

- **(24)** (a) The ten million dollar inheritance just walked in.
  (b) #The heiress is being held at Prudential-Bache Securities. (meaning the $$)

- **(25)** (a) The biggest engine just won the race and is having its tires changed.
  (b) #The car is two feet wide and three feet long. (meaning the engine)

**Other Domain-Attribute Asymmetries: anaphoric peninsulas**

- **(26)** (a) I saw headlights coming straight at me, but I was able to get out its (=the car’s) way.
  (b) #As the car came at me, I noticed they (=the headlight’s) were bright.

  (a) Look at her Dolly Partons.
  (b) When Dolly Parton comes in, they jiggle

**Framing Effects**

- **(28)** The deep extraction examples involve highly stereotyped/conventional information.
  (a) advocating the enforcement of laws
  (b) the editor exercising control over the publication of a newspaper
  (c) keys to the apartments in a building
  (d) discussing opinions about a subject

... (MW: Is this letting reanalysis in the back door?)
(29) (a) Which posts did you seek appointment to?
(b) Which posts did you refuse appointment to?
(c) ?Which posts did you appreciate appointment to?
(d) ?Which posts did you discuss appointment to?
(e) *Which posts did you describe appointment to?
(f) *Which posts did you study appointment to?

(30) Narrative frames with coordinates that permit extraction:
--- preparatory actions
--- scene setters
--- internal causes
--- incidental events

Functionalist Accounts in the Literature

(31) Erteschik-Schir and Lappin (1979)
Takami (1988)
Kuno (1987)

(32) E-S: You can extract constituents only when they can be construed as dominant.

*a constituent C of a sentence S is dominant in S iff the speaker intends to direct attention of his hearers to the intention of C, by uttering S.*

**Lie-Test**

Bill said: John believes that Orcutt is a spy.
(a) which is a lie -- he doesn’t
(b) which is a lie -- he isn’t

Bill said: John carefully considered the possibility that Orcutt is a spy.
(c) which is a lie -- he didn’t (consider it)
(d) #which is a lie -- he isn’t (a spy)

Bill said: That Sheila knew all along is likely.
(e) which is a lie -- it isn’t.
(f) *which is a lie -- she didn’t.

(33) Kuno: *Topichood condition for extraction*— Only those constituents in a sentence that qualify as the topic of the sentence can undergo extraction processes.

(a) Henry plays the lute and sings madrigals.
(b) *The lute which Henry plays ___ and sings madrigals is warped.

(c) John married someone who dislikes this child.
(d) *This is the child who John married someone who dislikes.
(c’) John actually married someone who doesn’t like the child. (MW: ‘about’ child)
(d’) ??This is the child that John actually married someone who doesn’t like.

(e) ##Speaking of violence, Snead is an Englishman who condones it.
(f) Speaking of these languages, I have a friend who knows them.
(g) *Violence is something that Snead is an Englishman who condones.
(h) You study languages that you have a friend who knows.

(34) Takami: Extraction occurs from the focus phrase – I.e., what provides new information.

(a) John gave the book to a young girl.
(b) The gang opened the safe with a drill.
(c) John was still a small boy in 1950.

(d) Which girl did John give the book to?
(e) What did the gang open the safe with?
(f) *Which year was John still a small boy in?

(g) Did John give the book to a young girl? (no – to a grownup)
(h) Did the gang open the safe with a drill? (no – with dynamite)
(i) Was John still a small boy in 1950? (#no – in 1940)

(i) Which party did John write the letter after?
(j) *Which party did John bury the letter after?

(35) Deane: None of these accounts is, on its own, sufficient, but they point to a common source explanation, attention.

Attention and Extraction

(36) (a) Extraction consists in the establishment of a long-range grammatical relation.
(b) An obvious prerequisite to establishing a relation between two concepts is that one be paying attention to both concepts at the same time.
(c) This presents no problem with local grammatical relations; [they’re] directly adjacent.
(d) In long distance extraction, the two concepts to be linked are separated far enough from one another that some means must be provided to focus attention on both.
(e) And what means would be more natural than if the two concepts were ones which commanded attention anyway?
(37) \[\text{attention} \rightarrow \text{NP} \rightarrow \text{[XP ... ]} \rightarrow \text{topic} \rightarrow \text{focus} \rightarrow \text{dominance}\]

(See, however, Foraker & McElree (2007): Linguistically-focused constituents are not maintained in the focus of attention.)

(38) **Spreading activation account**

*Divergent spread from core to periphery/parts*
- domain to attributes
- topic to propositions about it
- discourse entity to its properties

*Convergent spread from periphery/parts to core*
- attributes to domain
- propositions about a topic to topic
- properties to entity

(39) **Activation of Whole**
- Salient (max)
- Active (~ ½)
- Inactive (0)

**Activation of Parts**
- Active
- Inactive
- Inactive
(40) **Activation of Parts**
- Salient
- Active
- Inactive

**Activation of Whole**
- Salient
- Active
- Inactive

*Syntactic Spatialization of Form*  (cf. Lakoff 1987)

*Phrase structure trees as spreading activation networks*

(41) **Upwards search**

```
  A
 B  C
D E F G
```

(42) **Down**

```
 S'  
  |  |
  S  |
  |  |
  V  |
  |  |
  NP |
  |  |
  NP |
```

The further assumption is made that activation from an XP is essentially copied to X.

Searching down essentially ‘activates’ all the heads on the c-command path below it.

(43) If we view conjuncts as headless, then activation will ‘stop’ spreading at conjoined VPs.
This predicts that conjuncts will be islands unless ‘some other factor intervened’
But it seems a strange assumption that somehow headless structures stop activation spread.
(44) Conclusion (p. 57, §5.3)

(I) Grammatical processing is conditioned by the distribution of attention;
(II) The distribution of attention depends on spreading activation to be precise:
   i. The essential locality of syntactic extraction follows from patterns of activation spreading in syntactic tree structures; these patterns follow in turn from the way activation spreads between part and whole in an conceptual representation.
   ii. In some cases, activation can spread via conceptual representations of sentence meaning: this is essentially what happens in sentences where an attribute noun facilitates extraction.
(III) Long-range extraction depends upon another aspect of attention-distribution: the fact that certain concepts (potential topics and intrinsic foci) automatically command attention. Grammatical processes that relate phrases expressing such concepts are thus possible regardless of syntactic distance.

(45) A couple of basic concerns (MW) –

(a) This style of explanation implies very very dense concurrent representations. There must be nodes that share connectivity with syntactic and semantic representations, so that many different kinds of factors can interact.
(b) The resource demands of spreading activation representation of phrase structure are high – demanding potential connectivity between any set of nodes that could be related.
   -- Given standard productivity/systematicity arguments, this means lots of potential connectivity.
   -- Implemented spreading activation models all use some sort of gating circuits between constituents. Activation spread is not passive in these cases, undermining some of D’s assumptions.
(c) What happens with parallel analyses?
(d) D wants attention to be limited resource. But his account of spreading activation is not. (Activation copied from XP to X).

(46) A psycholinguistic concern –

(a) We typically characterize the processing of filler-gap relationships as active. The filler seems to initiate a search for potential gap sites in the input, before any direct evidence of where a gap might be (i.e., before a missing constituent is indicated).
(b) There is reasonable evidence that there are two phases: search and retrieval.
(c) The search step of linking fillers to gap sites does not seem to depend on divided attention in D’s sense. Each part of the input is ‘focal’ for a brief amount of time
during the search process, so attention isn’t *necessarily* ‘distracted’ by intervening material, once a potential gap site is located.

(d) Another way of putting it – it’s not clear there’s a lot of concurrent structure that could distract the linking of filler and gap, when that linking occurs. See Berwick & Weinberg (1984) for similar discussion, of how certain processing strategies render distant structure effectively local.

(e) There may be ways, though, of translating Deane’s ideas into more current thought.