

Discourse-linking and long-distance syntactic dependency formation in real-time

Oliver Boxell
University of Essex

QITL 4

Humboldt-Universität zu Berlin, March 2011

What are long-distance dependencies?

A dependency which spans more than one clause

[CP What do you think [CP (*that*) he'll say ~~what?~~]]



A diagram consisting of a horizontal line with an upward-pointing arrow at the left end and a downward-pointing arrow at the right end, connecting the 'What' in the outer clause to the 'what?' in the inner clause.

BUT locality principles mean that long-distance dependencies occur **successive-cyclically**. This means they involve multiple within-clause steps. Note the intermediate antecedent representation.

[CP What do you think [CP what (*that*) he'll say ~~what?~~]]



A diagram consisting of two horizontal lines with upward-pointing arrows at their left ends and downward-pointing arrows at their right ends. The first line connects the 'What' in the outer clause to the 'what' in the inner clause. The second line connects the 'what' in the inner clause to the 'what?' in the inner clause. The 'what' in the inner clause is circled in red.

Is *wh*-dependency formation constrained?

Island constraints (Chomsky 1962; Ross 1967) define domains within which a fronted *wh*-phrase may not grammatically set-up an association.

A violation of the locality constraints which drive successive-cyclicity creates a *wh*-island violation:

***Who** did the doctor wonder [CP **what** he should send *what* to *who*?]



A noun phrase changes everything...

Wh-dependencies with a *which-N* antecedent (e.g. *which horse*) rather than a bare antecedent (e.g. *who*) seem to have peculiar properties. This has been widely observed: Karttunen (1977); Maling and Zaenen (1982); Pesetsky (1987, 2000); Comorovski (1989); Cinque (1990); Rizzi (1990).

Island type	Bare <i>wh</i> -phrase	Which-N <i>wh</i> -phrase
Wh-island	Who did the doctor wonder what he should send to?	Which lady did the doctor wonder what he should send to?
Adjunct island	When did you buy a cake from Mr. Bun on?	Which Tuesday did you buy a cake from Mr. Bun on?
Conjunct island	What did you eat some cakes and drink?	Which coke did you eat some cakes and drink?

Table 1: Island violations with bare and *which-N* antecedents

Which-N forms are D(iscourse)-linked

Well, they are according to Pesetsky (1987).

D-linked forms are those with lexical specificity. This specificity can limit the set of possible referents for the *wh*-expression to those which may be extrapolated from the discourse-context.

Why should this mean that D-linked dependencies are less sensitive to constraints like islands?

D-linking and the Binding Hypothesis

Pesetsky (1987) proposed that D-linked antecedents (**optionally**) take scope over their base-generated representation via a binding operation.

Binding is not thought to unfold successive-cyclically. Thus, it leaves no intermediate representation:

Which lady_i did the doctor wonder [CP what he should send to ~~*which lady*~~_i]?

Islands are constraints on *wh*-**movement**-type dependencies. Thus, if D-linked dependencies are formed by **binding** they would be precluded from sensitivity to such constraints.

What do we know already about the processing of D-linked *wh*-phrases?

D-linked antecedents seem to have **an early-established discourse prominence** (Radó 1998; Frazier and Clifton 2002; Diaconescu and Goodluck 2004).

D-linked antecedents take longer to reconstruct at underlying verb positions. This is probably due to the **increased lexical information which needs reconstructing** (de Vincenzi 1996; Shapiro et al. 1999; Shapiro 2000; Piñango et al. 2001; Piñango and Burkhardt 2005; Schumacher et al. 2010).

The current study

Hypothesis:

Real-time instantiation of an intermediate representation of an antecedent should be **restricted to non D-linked *wh*-dependencies** if D-linked dependencies form in a single (binding) step.

An evaluation of The Binding Hypothesis conducted by testing its psychological reality

Materials

- Based on Gibson and Warren (1999; 2004) and Marinis et al. (2005)
- Non-cumulative self-paced reading study with 4 conditions containing 2 critical manipulations (+/- D-linked antecedents; +/- Intermediate CP structure)

(a) - D-LINKING; + INTERMEDIATE REPRESENTATION

The manager wondered **who** the secretary claimed [**CP** ~~who~~ that] the new salesman had ~~pleased~~ ~~who~~ in the meeting.

(b) - D-LINKING; - INTERMEDIATE REPRESENTATION

The manager wondered **who** the secretary's claim about the new salesman had ~~pleased~~ ~~who~~ in the meeting.

(c) +D-LINKING; + INTERMEDIATE REPRESENTATION

The manager wondered **which gentleman** the secretary claimed [**CP** ~~which gentleman~~ that] the new salesman had ~~pleased~~ ~~which gentleman~~ in the meeting.

(d) +D-LINKING; -INTERMEDIATE REPRESENTATION

The manager wondered **which gentleman** the secretary's claim about the new salesman had ~~pleased~~ ~~which gentleman~~ in the meeting.

Predictions for these materials

Reading times at the subcategorising verb

- Facilitation (as the antecedent is “locally primed”) by the presence of an intermediate representation
- Such facilitation should also be restricted to non D-linked dependencies according to The Binding Hypothesis

Reading times at *that/about*:

- Increased reading time at intervening CP (*that*) as antecedent reactivation is processed compared to *about*
- Such findings should be restricted to non D-linked dependencies according to The Binding Hypothesis

Participants

- 40 native speakers of English (mean age 23 years; range 18-25, 23 female)
- All right-handed
- 37 were monolingual
- All reported that they had never been diagnosed with any language or general cognitive disorders
- All participants' working memory was screened using a standardised reading-span test (Daneman and Carpenter 1980). There were no statistically significant differences in participants' performances.

Reading times

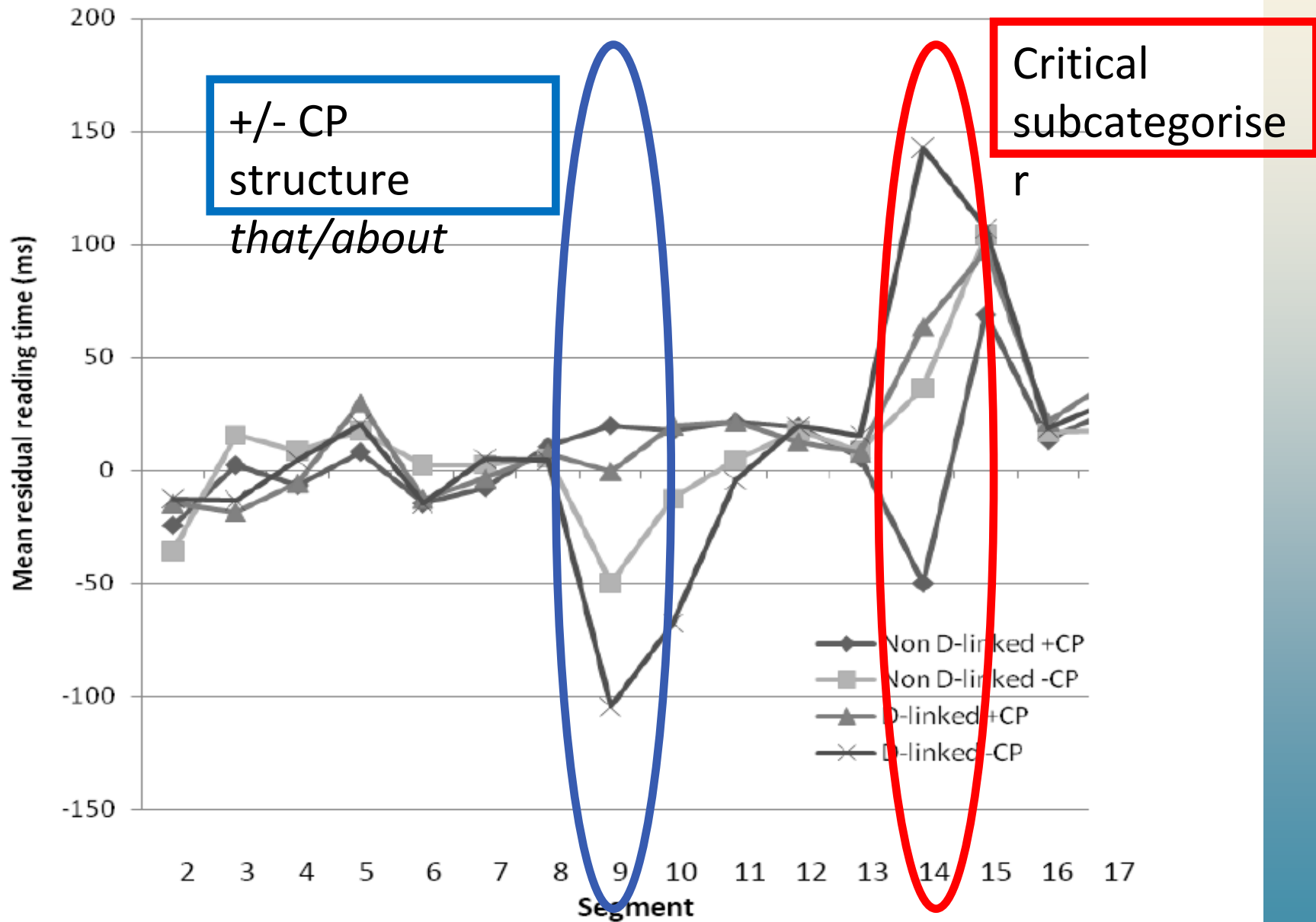
Residual reading times were calculated by taking participants' actual reading time for each segment and deducting their predicted reading time (as made by tailor-made regression equations for each participant).

Residual reading times calculated in this way control for differences in participants' reading speed.

Only reading times for sentences for which participants correctly answered the comprehension question were included in the subsequent analysis (= 8.9% of the data excluded).

Reading times greater than 1000ms were also excluded (affecting 0.28% of data).

Mean residual reading times (ms)



A summary of the findings

At the critical subcategoriser (*pleased*)

→ **Main effect of D-linking:** D-linked conditions take longer to read than non D-linked ones, $F1(1, 38) = 432.214$, $p < .001$; $F2(1, 17) = 144.757$, $p < .001$

→ **Main effect of intervening structure:** An intermediate CP structure facilitates reading times for both D-linked and non D-linked antecedents, $F1(1, 38) = 432.214$, $p < .001$; $F2(1, 17) = 144.757$, $p < .001$ (also confirmed by planned comparisons)

→ **No interaction between D-linking and intervening structure**

At the +/-CP boundary *(that/about)*

- **Main effect of D-linking:** D-linked conditions are read faster than non D-linked ones for both sentences which are + and –an intermediate CP, $F1(1, 38) = 445.763, p < .001$; $F2(1, 17) = 14.661, p = .005$ (also confirmed by planned comparisons)
- **Main effect of intervening structure:** +CP conditions are read slower than –CP ones, $F1(1, 38) = 21.433, p < .001$ **BUT** $F2(1, 17) = .638, p > .05$
- **Interaction of D-linking and intervening structure:** $F1(1, 38) = 50.736, p < .001$ **BUT** $F2(1, 17) = 1.511, p > .05.$

Discussion

→ Faster reading times found at the critical subcategoriser where an intermediate CP structure is present relative to when one is not.

→ This was true for both non D-linked and D-linked dependencies.

→ The Binding Hypothesis predicts that only non D-linked antecedents should be reactivated at intermediate CP structures if D-linked antecedents bind their underlying position.

→ These data do not seem to support the predictions taken from The Binding Hypothesis.

→ The idea that different mechanisms instantiate D-linked dependencies from non D-linked ones can be ruled out.

→ +CP boundaries (*that/about*) were read slower than –CP boundaries, when D-linked and when non D-linked.

→ Maybe this is due to setting-up an intermediate representation at the CP?

→ Maybe this is due to clause boundaries entailing “wrap-up” procedures (Kluender and Kutas 1993)?

→ Maybe this is due to the parser’s predictions about the new clause to come, e.g. planning argument structures (see Gibson 2000)?

→ Maybe this is due to *that* and *about* having frequency differences, valency-type differences, word-class (and word access) differences?

D-linked condition reading times at the underlying subcategoriser were **slower** than non D-linked counterparts.

D-linked condition showed **faster** reading times at the clause boundary compared to non D-linked counterparts.

Slower reading times at the verb: probably due to reactivating lexical information (converging with Shapiro et al. 1999 and others).

What does this say about intermediate reactivation?

We might say intermediate antecedent reactivation is **purely structural** since these data may suggest the lexical-semantic content of an antecedent is selectively reactivated at the verb and not at intermediate positions.

(Implications for Copy Theory of Movement: “weaker” copies at intermediate chain positions?)

Two key questions left open:

If Pesetsky's (1987) account of the ameliorative properties of D-linking is unsupported, then how do we account for them?

Why are conditions with D-linked antecedents read faster at *that/about*, whether or not this was a CP clause-boundary?

Time for some post-hoc “wild speculation”...

The Stabiliser Hypothesis (SH)

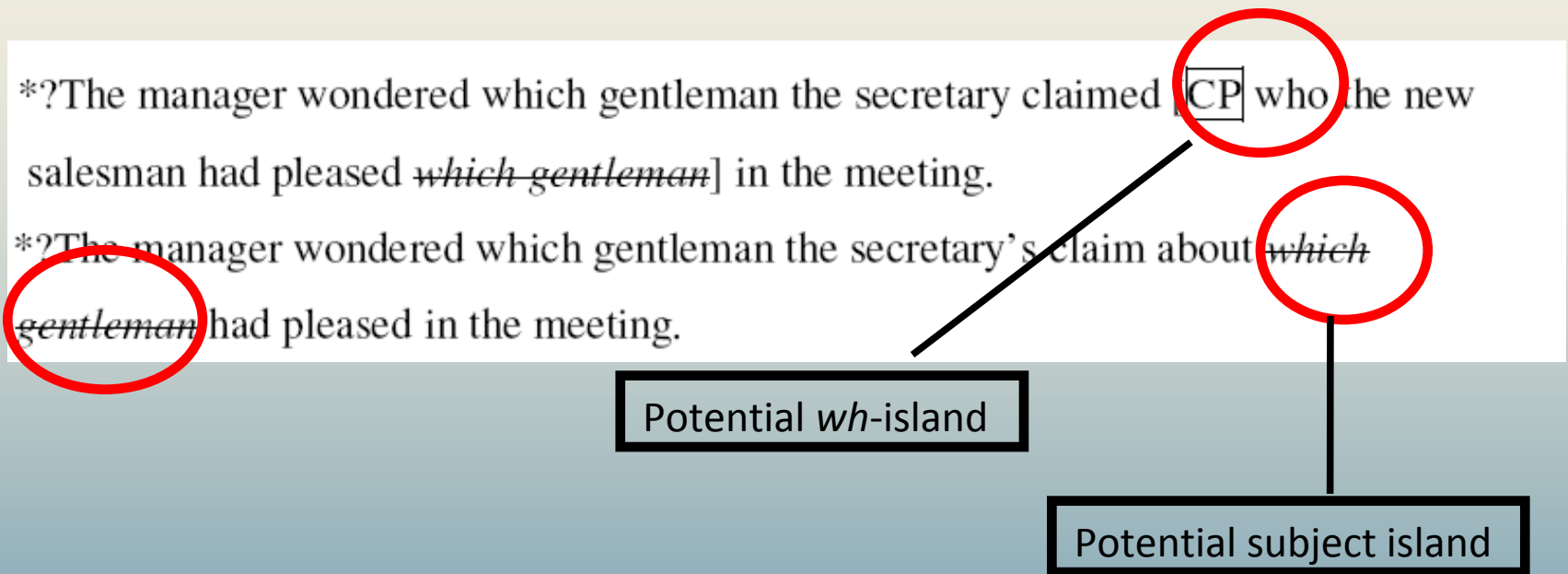
Recall it is widely accepted that a property of D-linking is early discourse/conceptual structure instantiation for the antecedent (Radó 1998; Frazier & Clifton 2002; Diaconescu & Goodluck 2004).

SH goes one step further.

The early-established conceptual structure D-linking provides “stabilises” the parse for when it encounters **time-locked moments of complexity or difficulty.**

I am going to argue that this “stabilising” is what can be seen at the *that/about* segment.

Segment 9 is a *potential* island violation in both the conditions with and without an intervening CP structure:



Maybe the parser is aware of a potential for the parse to crash at this moment in the time-course (namely, at the *that/about* segment).

After all, the parser is aware of real island boundaries at the exact moment they are parsed (McKinnon & Osterhout 1996). Maybe the parser prepares for *possible* islands in advance...

D-linking, then, could “stabilise” the parse, neutralising *potential* island boundaries.



Non D-linked (+CP)

D-linked (+CP)

Non D-linked
(-CP)

D-linked (-CP)

9

Maybe stabilisation can explain why D-linking can ameliorate or reduce sensitivity to real island violations as well?

Why might D-linking stabilise (potential) island violations? More of the conceptual/ discourse structure is built earlier in the parse. This means:
→ The range of possibilities for how the parse may plausibly proceed is restricted;
→ The target of comprehension processing (deducing the meaning of the input) is simplified.

The parser can “speed-past”, “ignore” or “work-through” potential constraints, falling-back on the discourse/conceptual-structure to figure out *who* did *what* to *whom* and complete the dependency.

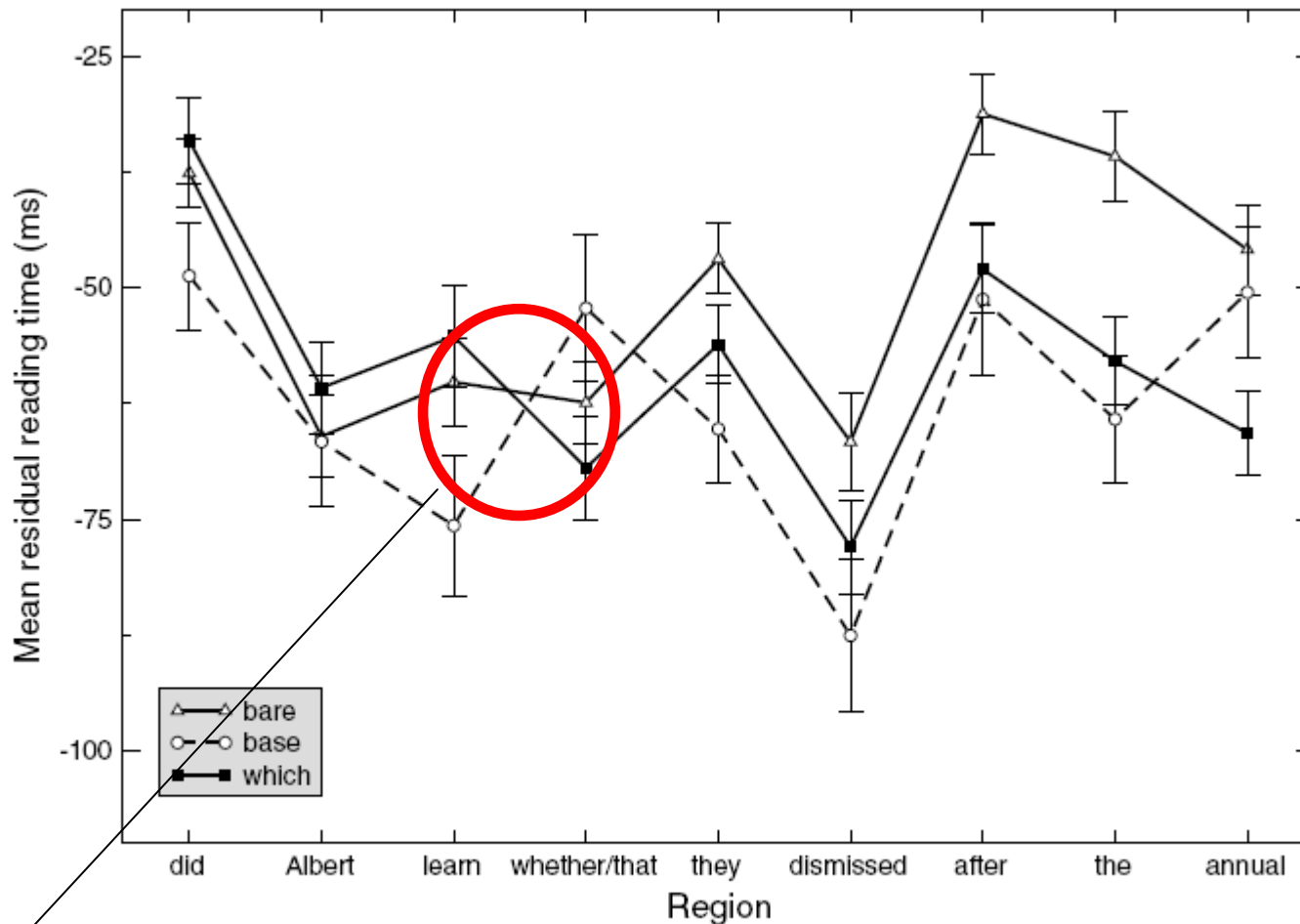


FIGURE 4. Mean residual reading times in experiment 2, ranging from first word after WH-phrase to three words after the subcategorizing verb. Error bars show (+/-) one standard error.

Looks like a stabiliser effect! But there is no (*which* without an island) control condition.

Hofmeister and Sag (2010)

My current research is investigating this further.

Conclusions

Support for The Binding Hypothesis as an explanation for the ameliorative effects of D-linking on constraints like islands has not been found.

The idea that D-linked dependencies are formed using a different mechanism (e.g. binding rather than movement) from non D-linked ones can be ruled out.

Intermediate antecedent reactivation does not seem to include the same level of lexical information that an underlying verb position does.

These data *may* motivate a new processing-based account of the amelioration of D-linking on constraints like islands.

Acknowledgements

Claudia Felser

The Psycholinguistics Research Group at
The University of Essex

The participants involved in the experimental work

Email address: osboxe@essex.ac.uk

(Reference list available by email)