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

Oliver Boxell

A self-paced reading experiment examined the processing of long-distance wh-dependencies. Some wh-dependencies were D(iscourse)-linked (i.e. involved a lexically-specified wh-phrase like which horse) and others involved a non D-linked wh-phrase like who. Both kinds of wh-phrase show evidence of being represented at available intermediate positions within the long-distance dependencies. It is argued that this is inconsistent with Pesetsky (1987) whose hypothesis claimed that D-linked dependencies are formed via binding, and as such are predicted to be able to bypass intermediate reactivations. At points in the time-course where the potential for the violation of dependency-formation constraints becomes apparent, a D-linked antecedent seems to benefit the parse. This motivated the idea that the reduction in constraint-sensitivity which D-linking seems to elicit could be the result of time-locked “stabiliser” effects during moments of complex processing. Finally, the data may indicate a behavioural difference between intermediate and verbal antecedent reactivation during processing.

1. Introduction: The descriptive and theoretical background
1.1. The mechanics of wh-dependencies

Transformational syntactic frameworks like Principles and Parameters (Chomsky 1981) and Minimalist Program (Chomsky 1995) use combinations of basic phrase-structure rules and transformations to represent the sentences of human language. The surface-structure of certain clauses is formed by moving a wh-word (e.g. who) from a base-generated (or deep-structure) position. An empty-category trace or copy of the moved constituent remains in that base-generated position such that the surface and deep structure positions are linked, forming a so-called syntactic-dependency or “chain”. In particular, the following clause-types can feature wh-movement (with correspondingly-numbered examples below): (1) wh-questions, (2) relative clauses and (3) complement clauses.

(1) Who did Mary love who?

Proceedings of ConSOLE XIX, 2012, 151-175
http://www.sole.leidenuniv.nl
© Oliver Boxell
(2) This is the girl who John loved who.

(3) The King wondered who he should buy a castle for who.

Non-transformational frameworks also propose wh-dependency formations. For example, Head-driven Phrase Structure Grammar (Pollard & Sag 1994) posits a “slash-feature” which represents the fronted wh-phrase and is passed from a head to its dependent(s).

Furthermore, wh-dependency formation appears to occur successive-cyclically. That is, by way of obeying locality constraints, it occurs in small step-wise operations. If a dependency forms across more than one clause, it is thought to unfold using within-clause steps. As it does so, Chomskyan syntacticians would posit the existence of intermediate empty category copies or traces of the antecedent at each intervening CP boundary as in (4).

(4) The manager wondered [CP who [C the secretary claimed [CP who [C that the new salesman had pleased who]]]].

Most non-transformational syntacticians also posit intermediate representations of a wh-antecedent when its dependency spans more than one clause. In particular, there is broad agreement that clause-boundaries (i.e. intervening CP structures within dependencies) host intermediate antecedent representations. Successive-cyclic dependency-formation has accumulated a variety of observational evidence, see e.g., Thornton (1995), McCloskey (2000, 2002) and Felser (2004).

1.2. Constraints on wh-dependency formation

“Islands” (Ross 1967) are constraints on wh-dependency formation which define domains within which a fronted wh-phrase may not be grammatically associated. The wh-island works in harmony with the locality constraints which drive successive-cyclicity, stating that if one wh-phrase is occupying an intermediate CP site within a long-distance dependency, any other wh-phrases with base-generated positions below that intermediate CP cannot engage in any wh-dependency formation since successive-cyclicity could not be accommodated. This is illustrated in (5a).

(5) (a) * [CP Whom [C did the doctor wonder [CP what [C he should send what to whom]]]]?

(b) The doctor wondered what he should send to whom.

In (5a), the wh-island is caused by what appearing at the intermediate CP, making it impossible for whom to set-up a dependency in a way which obeys the locality constraints which drive successive-cyclic dependency-formation. The result is therefore ungrammatical. We can be sure it is the whom-dependency which causes this ungrammaticality because when
whom is left *in-situ* (i.e. it does not participate in any *wh*-dependency formation) the result is (5b) which is grammatical.

### 1.3. An amendment to the mechanics of *wh*-dependencies: discourse-linking

It seems that *which*-N(oun) forms (e.g. *which man* and *which book*) have some quite unique properties when contrasted with basic *who* or *what* forms. Pesetsky (1987) terms *which*-N forms “Discourse-linked” (hereafter D-linked) and simple *who* type forms as being non D-linked. D-linked forms are so-called because they have lexical specificity through the N they carry, which limits the set of possible referents for the *wh*-expression to those which may be extrapolated from the discourse-context in which the construction is located. It seems that when the *wh*-phrase is D-linked, the grammaticality of an island violation is ameliorated.

Table 1 gives non D-linked examples of some islands and their D-linked counterparts. Note that the non D-linked island violations are ungrammatical, while the grammaticality of the same violations is significantly ameliorated where the *wh*-expression is D-linked.

<table>
<thead>
<tr>
<th>Island constraint</th>
<th>Non D-linked version</th>
<th>D-linked version</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Wh</em>-island</td>
<td><em>Whom did the doctor wonder what he should send to?</em></td>
<td>Which lady did the doctor wonder what he should send to?</td>
</tr>
<tr>
<td>Adjunct island</td>
<td><em>When did you buy a cake from Mr. Bun on?</em></td>
<td>Which Tuesday did you buy a cake from Mr. Bun on?</td>
</tr>
<tr>
<td>Conjunct island</td>
<td><em>What did you eat some cakes and drink?</em></td>
<td>Which coke did you eat some cakes and drink?</td>
</tr>
</tbody>
</table>

**Table 1: (Non) D-linked examples of island violations**

Pesetsky (1987) offers a formal account for why D-linked *wh*-dependencies might not be (as) obedient to island constraints. Baker (1970) proposed that *wh*-dependencies may be formed by *wh*-phrases (or rather some “*Q* operator” in the CP of relevant clause-types) taking scope over underlying constituents via unselective binding. Note that dependencies formed by binding are not thought to unfold successive-cyclically and so extend across any intervening CP structures in one single operation:

\[
(6) \quad [\text{CP Q, Which lady did the doctor wonder [CP what he should send to which lady]]}]
\]

Pesetsky’s claim specifies that *D-linked wh-phrases only* may (optionally) bind their underlying positions in this way. Since islands are conceived as constraints on *wh*-movement, the decreased sensitivity of D-linked dependencies to them may be explained by the fact that they are not formed via movement in the first place. More specifically, since *wh*-islands are essentially created by a violation of successive-cyclicity, which is a feature of *wh*-movement, it makes sense that D-linking would not be sensitive to such a constraint if its dependency formation results from binding which need not unfold in multiple steps. Non D-linked dependencies are sensitive to island constraints, and Pesetsky therefore assumes they are indeed formed via movement. In a bid to maintain framework-neutrality, I re-stipulate Pesetsky’s hypothesis:
D-linked *wh*-dependencies are set-up (in any formal framework) via a single **one-step binding-type** operation. Non D-linked *wh*-dependencies may be treated via ordinary *(multiple-step) movement-type or slash-pathway type* analyses (when extending across multiple clauses).

### 1.4. The prediction generated by the theory

Pesetsky (1987) offers a formal account based on mechanical differences between D-linked and non D-linked *wh*-dependencies in which the former bind and the latter move. This account arises out of the behavioural differences between the two distinct dependency types. Namely, D-linked dependencies are less sensitive to certain constraints which are associated with movement relative to non D-linked dependencies. The question now becomes whether or not Pesetsky’s exact formalism is the correct account of these behavioural differences. Since theoretical syntactic-formalisms attempt to represent grammatical knowledge, which in turn are thought to influence online instantiations of syntax during processing, there should be some psychological-reality to Pesetsky’s hypothesis if it is indeed an accurate account. The current research will adopt a method from the psycholinguistic sentence-processing literature which has been used with success to find evidence of intermediate representations of an antecedent in long-distance (multiple-clause) *wh*-dependencies. The current study includes a D-linking manipulation, leading to a straightforward prediction if Pesetsky (1987) is correct: Psycholinguistic evidence of intermediate representations of the antecedent in long-distance *wh*-dependency formation should be restricted to non D-linked *wh*-dependencies since only these involve multiple-steps. No such evidence should be found for D-linked dependencies since they are formed in a single (binding) step.

### 2. The experimental record

#### 2.1. Processing *wh*-dependencies

The psychological reality of *wh*-dependency formation has been well ascertained from collections of data which demonstrate, in real-time, the mental reactivation of fronted *wh*-antecedents at the underlying position with which they have association. Much of this work has been done using the self-paced reading method (Just et al. 1982) in which experimental participants read through a sentence word-by-word on a computer screen, using a pacing button to call up each successive segment. The computer records the reading-time for each segment, the length of which is associated with the difficulty of the processing-work for the corresponding item(s).

Crain & Fodor (1985) used this method to compare the processing of a *wh*-question with that of a “yes-no” question, as illustrated in (14).

(8) (a) *Wh*-question: What did the little girls ask us to sing those French songs for?
(b) Yes/no question: Did the little girls ask us to sing those French songs for any reason?

The segment following the first subcategorising verb encountered (*ask*) was occupied by an object (*us*) in both sentence-types. A so-called filled-gap effect was elicited for the *wh*-question at this site, but not for the “yes-no” question. This is where the parser appears to posit the object site as a gap which may be filled by a reactivation of the fronted *wh*-filler but then has to revise this analysis on discovering that the site is already filled. Such reanalysis
increases the demand on the processor, leading to the increased reading time known as the filled-gap effect. This effect was not observed for the “yes-no” control since there is no wh-dependency to be set-up in this condition.

Gaps (and thus filled-gap effects) seem, then, to occur in an argument position. They may correspond to the empty-categories (copies or traces) of Chomskyan grammar where a phonetically-null representation of the antecedent occupies its underlying argument position. Many other studies provide similar real-time processing evidence for antecedent-reactivation in wh-movement (e.g. Swinney et al. 1988; Frazier & Clifton 1989; Nicol & Swinney 1989; Nicol, Fodor & Swinney 1994; Chen et al. 2002 and others). Such evidence could support The Active Filler Hypothesis (Clifton & Frazier 1989) which suggests the antecedent of a dependency triggers an on-line search-strategy which posits potential gap-sites until one is found whereupon the antecedent is reactivated, filling the gap. However, many of the studies use a head-initial language, namely English, to demonstrate the reconstruction of the fronted constituent at an underlying position. Since the subcategorising verb precedes the hypothetical empty-categories in most sentences tested, there is an alternative explanation for these data. This is the so-called Direct Association Hypothesis (Pickering & Barry 1991) which assumes that a fronted antecedent is directly associated with the subcategoriser as part of that subcategoriser’s argument structure without relying on an actual gap. Thus, reactivation effects observed at or directly after the subcategoriser may result from this direct-association. In support of this, Boland et al. (1995) found that implausible antecedent-subcategoriser associations were identified at the verb position itself, suggesting the dependency had been completed on the verb rather than being delayed until a subsequent gap.

While it seems evidence of reactivating a fronted-constituent may be accounted for with either the Active Filler or Direct Association hypotheses, some effort has been made to dissociate the two. Nicol (1993) and Roberts et al. (2007) found evidence of antecedent reactivation at the second-object position in a sentence with a ditransitive verb. The effect observed was therefore not immediately adjacent to the subcategoriser, but was where a gap or empty-category would be posited later in the sentence. Additionally, studies on verb-final languages have shown reactivation effects at posited gap-sites which occur before the subcategoriser is encountered (e.g. German: Clahsen & Featherston 1999; Japanese: Nakano, Felser & Clahsen 2002). These data are inconsistent with Direct Association and are predicted by the Active Filler Hypothesis. In the event, however, it may be that further research is required to tease apart the different theories of dependency-formation during on-line processing.

2.2. Successive-cyclicity in real-time

While the reactivation of fronted constituents at their base-generated or verbal positions has been firmly established in the literature, little attention has been given to identifying intermediate activations of an antecedent during long-distance dependency parsing. One early attempt was Frazier & Clifton (1989) who compared real-time dependency formation across a single clause with that of across two clauses. Reading times for the latter were indicative of a significant increase in complexity, which the authors have suggested may be attributable to the costs of intermediate antecedent reactivation at the intervening clause boundary (i.e. successive-cyclicity). More recently, Gibson & Warren (1999, 2004) measured reading-times for wh-dependencies with an intervening CP structure and for ones with an intervening structure which was not a CP using self-paced reading. The materials are illustrated in (15):
(9) (a) The manager [CP who the consultant claimed [CP who that the new proposal had pleased who]] will hire five workers tomorrow.

(b) The manager [CP who the consultant’s claim about the new proposal had pleased who] will hire five workers tomorrow.

The rationale was that only (9a) accommodates an intermediate representation of the antecedent. The relative clause contains the verb claimed which selects for a complement clause, i.e. CP. This critically provides the position for an intermediate representation of the antecedent. The verb is nominalised in (9b) to make the construct-genitive argument the consultant’s claim about the new proposal. Consequently, an intermediate representation of the antecedent is not accommodated. The materials also included constructions without wh-dependencies, but which were otherwise the same.

Reading times were found to be significantly quicker at the critical verb pleased for dependencies which had an intervening CP (9a) relative to those without an intervening CP (9b). No such effects were found for the minimal-pair counterpart conditions with no wh-dependencies. Additionally, the reading times of the complementiser that were longer in the condition involving a wh-dependency relative to its minimal pair without one. Since this is the region where the intermediate representation of the antecedent is predicted to become activated, this effect may be attributable to the processing cost of such activation. However, this effect did not reach significance. Overall, these results seem to support the predictions which the presence of a mediating representation of the wh-phrase would make. Namely, intermediate reactivation of the antecedent facilitates its integration at the subcategoriser by boosting its prominence, and that the intermediate reactivation itself carries with it an instantiation processing cost.

This finding - and the whole notion of successive-cyclicity for that matter - is consistent with the Dependency Locality Theory (Gibson 1998, 2000). This theory stipulates that sentence processing complexity can be generated by any intervening elements (particularly newly introduced discourse-entities) within a dependency since they can reduce the activation of the dependency’s antecedent. Reactivating the antecedent at intermediate positions within long-distance dependencies therefore reduces the amount of intervening material and likewise the processing complexity for the dependency overall. Note also that ending a clause seems to involve “wrap-up” processes which may be costly for the processor (e.g. Kluender & Kutas 1993), while starting a new one entails making a new set of argument-structure predictions which need to be held in working-memory (Gibson 2000). Intermediate representations of an antecedent at a clause-boundary would, then, feasibly serve to alleviate the processing difficulty of an already complex structure. Furthermore, it would make intuitive sense to assume that breaking a long-distance dependency into smaller subsections with intermediate reactivations would be less taxing on working-memory than holding the first activation of the antecedent continuously until the subcategoriser is reached.

The findings of Gibson & Warren (1999, 2004) have been replicated in a study by Marinis et al. (2005). These researchers compared the native and non-native processing of similar materials, also using self-paced reading. Native speakers behaved exactly as they had in the Gibson & Warren study, with long-distance wh-dependencies which include an intervening CP facilitating a faster reading-time at the subcategoriser compared to dependencies of the same length but without an intervening CP. Non-native participants with a range of language backgrounds did not show evidence of representing the antecedent at
intermediate positions, whether or not their native language was thought to include successive-cyclicity (compare Dekydtspotter & Miller 2009). The lack of intermediate representations of the antecedent did not impede non-native speakers’ ability to accurately comprehend the sentences tested, supporting the idea that second-language processing instantiates less detailed syntax, possibly relying instead on other kinds of information (e.g. discourse, semantics and pragmatics) for comprehension (see Clahsen & Felser 2006a, 2006b).

Similar materials to Gibson & Warren (1999, 2004) were again appropriated by Finn, Kim & Piñango (2010). In this experiment, native English speakers did a cross-modal lexical decision task. Unlike the previous research, no evidence is found for intermediate reactivations of the antecedent. The authors argue that these data show that intermediate “gaps” have a different real-time status from those at argument positions, or even that they are evidence against the existence of psychologically-real intermediate representations of the antecedent. However, note that this latter conclusion is based on a null-result. Not finding evidence of a given phenomenon does not entail that it can never be found or that it doesn’t exist. Indeed, cross-modal priming relies on semantic information at the reactivation site, which is not necessarily present at intermediate positions since they do not involve many of the semantic characteristics of argument positions (e.g. thematic assignment and semantic integration with the subcategoriser). Thus, intermediate representations may be purely structural, and if they are, cross-modal priming effects would not be predicted. If intermediate reactivation is purely structural, the evidence reviewed in this section may resolve a problem identified by Fodor (1995) in which reactivation effects (at the verb) could rely purely on correctly linking a dislocated constituent with its verb semantically, making it difficult to isolate whether or not structural reactivation is actually psychologically real.

2.3. Discourse-linking

Although the status of successive-cyclicity in D-linked dependency formation and the time-course of D-linking amelioration effects have, to my knowledge, never been examined before, other aspects of its processing have been. Radó (1998) reports that, in a self-paced reading experiment, wh-questions were read faster when the antecedent was D-linked and associated with the subject position of a complement clause compared to when the antecedent was non D-linked and/or associated with another position. Radó suggests that D-linked antecedents are associated with the subject position more readily than non D-linked ones since the subject position is the default topic position. Since the D-linked discourse entities could be integrated into the conceptual structure early in the time-course, they could form natural associations with parts of the sentence associated with such key conceptual roles.

Similarly, Frazier & Clifton (2002) examined the possibility that the processor posits a discourse-referent for a D-linked antecedent early in the parse, by at least the end of the clause in which its surface-position is located. They hypothesised that if this were true, a D-linked antecedent should be considered by the parser as a referent for a personal pronoun in a following subordinate clause more frequently than a non D-linked alternative. (10) illustrates the materials tested:

\[(10) \quad \{\text{Which guy/Who}\} \text{ did Sam send a rifle to } \text{which guy/who [when he was threatened]}?\]
The results did indeed show that he in the adverbial clause co-refers to which guy more frequently than to Sam, but this pattern was not found when the wh-phrase used was who. The authors argued that there is no grammatical explanation for this, and so the results would appear to be a feature of the processing of D-linked antecedents which differs from the processing of non D-linked ones. Namely, D-linked wh-phrases seem to be reactivated on in the conceptual structure, giving them a certain discourse-prominence which makes them particularly salient antecedents for definite pronouns occurring in subsequent clauses. An alternative explanation for these data, however, is that there is a simple feature (e.g. gender) matching between the D-linked antecedent and the pronoun without the need to invoke the early discourse-instantiation hypothesis. Diaconescu & Goodluck (2004) conducted a replication of this study with Romanian speakers, Romanian being a language which includes null-pronouns. The null-pronouns were found to attract a D-linked antecedent interpretation just as the overt ones had in Frazier & Clifton (2002), with no possibility of feature-matching explaining the results.

While the above studies seem to show that lexical specificity bolsters reactivation for D-linked antecedents at certain conceptually-important positions, evidence from Shapiro et al. (1999) seems to suggest that reactivating D-linked antecedents per se taxes the processor more than non D-linked reactivation. A cross-modal priming task revealed that while priming effects were observed for non D-linked antecedents at the subcategoriser or gap-site, they were not present for D-linked ones until a significant period after that. This delay in D-linked antecedent reactivation may be attributable to the additional specified lexical information which must be reconstructed at the underlying position. Furthermore, Shapiro (2000) and de Vincenzi (1996) have further demonstrated that accessing D-linked information during wh-dependency formation is costly for the processor. Piñango et al. (2001), Piñango & Burkhardt (2005) and Schumacher et al. (2010) also show that this is true for dependencies which link reflexive pronouns with their antecedents, whereby non co-argument reflexives (which are considered D-linked) represent a heavier processing burden than co-argument reflexives (which are considered non D-linked).

Further, work by Donkers et al. (unpublished) suggests that the notion that D-linked reactivation may be more burdensome than non D-linked equivalents may be refined. They demonstrate that non D-linked and hypernym D-linked (i.e. D-linking with more generic superordinate terms as in which person) dependencies caused less processing-difficulty than hyponym D-linking (i.e. D-linking with more specific terms as in which nurse). The authors suggest the processing difficulty for the more specific D-linking is caused by increased restriction of the set of possible referents for the antecedent. Indeed, Goodluck et al. (2008) found a similar patterning for the overall comprehension of hypernym D-linking, and they suggested that more specific D-linking may stimulate (more detailed) visualisation of the referent which causes increased processing complexity for the dependency of which it is a part. They tested Broca’s aphasic patients and subsequently found similar patterns in child processing as well (see Goodluck 2010).

In sum, it seems there are two main observations to be drawn from the literature on the processing of D-linked wh-dependencies: (1) a D-linked antecedent is more readily reactivated with parts of a sentence with key conceptual roles, presumably due to its early-established discourse-prominence; (2) there seems to be a hierarchy of dependency formation difficulty in which non D-linked ones are the easiest, hypernym D-linked ones are harder and hyponym D-linked ones are most difficult.
3. The present study

The present study is an attempt to examine the psychological reality of the claim made by Pesetsky (1987) that non D-linked wh-dependencies are formed via conventional means (e.g. transformational movement or slash-pathways) while D-linked ones can be formed by unselective binding.

3.1. Materials and predictions

3.1.1. Self-paced reading task

The main task is based on that of Gibson & Warren (1999, 2004). The premise of the design relies on the fact that antecedent reactivation at the subcategoriser or subsequent gap-site should be easier to do when the antecedent has been recently activated via an intermediate representation relative to when there is no such intermediate representation. This is because the intermediate representation boosts the antecedent’s mental activation. In a self-paced reading context, this should manifest itself in faster reading times at the critical verb (or gap) region where there has been a local (intermediate) activation of the antecedent, and longer reading times at that region when there has not been.

Each participant saw a total of 79 sentences: 3 of these were practice items, 20 were critical items and 56 were fillers. The fillers were included in a bid to prevent participants from acquiring any default response strategies, while the practice items were used to ensure participants fully understood the task before experimental measurements were taken. The critical items were created with the conditions [+/- D-linking] and [+/- An intervening CP structure]:

(11) (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.

The present materials only include wh-dependencies (+/- D-linked) are foregoes the use of control conditions with no wh-dependencies. Several previous studies (e.g. Gibson & Warren 1999, 2004 and Marinis et al. 2005) have already established that the effects being searched for at the critical subcategoriser are specific to constructions which include wh-dependency formation, suggesting that other differences between the conditions (e.g. differences in regions preceding the critical region) cannot explain the predicted results. The materials include a wh-filler which heads a complement clause which is selected for by the subcategorisation frame of the main verb (wondered). This wh-filler is either D-linked (which
Oliver Boxell

160

gentleman) or not (who). The ensuing dependency then proceeds until the critical subcategoriser (pleased) where the antecedent will be reactivated. The intervening structure, however, either includes an intermediate CP by virtue of another verb which selects for a complement clause (claimed), or it does not include an intermediate CP by virtue of that verb having been nominalised (the secretary’s claim about the new salesman). The conditions with the intermediate CP should be able to accommodate an intermediate reactivation of the wh-phrase, while the ones without should not have the syntactic space to do so according to most formal theories. Pesetsky (1987) claimed that D-linked dependencies can be formed in a way which should bypass such intermediate reactivation. Therefore, it is predicted that reading-times at the critical subcategoriser (pleased) should be facilitated by a local representation of the antecedent only for the non D-linked dependencies. Additionally, increased reading-times at the intervening CP might only be predicted for non D-linked antecedents since only these should be reactivating there, incurring a processing cost in so doing.

In a bid to prevent participants from positing a reactivation of the antecedent as the direct object of the verb in the higher clause (i.e. claimed) in the [+Intermediate representation] conditions, all the verbs used in this position were biased for taking a sentential complement according to Garnsey, Pearlmutter, Myers & Lotocky (1997). The materials used in this study were actually adapted from Marinis et al. (2005) who additionally tested a further six verbs for such biases and these were also used.

Marinis et al. (2005) note that the materials in Gibson & Warren (1999) contain an asymmetry between the conditions with wh-dependencies compared to those without since one contained more words than the other and was thus potentially a more complex construction to parse. The critical conditions in this experiment have wh-dependencies with 8 word segments intervening between the wh-filler and critical subcategoriser. As such, the fronted wh-phrase was always presented as a single-segment in the word-by-word self-paced reading task, whether or not it was D-linked (and thus contained two words) or was non D-linked (and thus contained one word). All other segments for the critical items were presented as single words. To ensure that multi-word D-linked segments did not become especially salient to participants, the filler items included 16 sentences with non D-linked multi-word segments, and 16 others which included D-linked antecedents which were split over two single segments. After each critical item a “yes/no” comprehension question was asked to ensure participants had fully comprehended the sentence. Among the 56 fillers there were examples of D-linked and non D-linked wh-dependencies (with and without intervening CP structures) without comprehension questions, and other sentence-types with them, to ensure their presence did not alert participants to the significance of wh-dependencies, types of intervening structure and D-linking. The frequency of fronted wh-phrases and critical subcategorisers have been controlled by ensuring that all had fewer than 5000 “hits” on the British National Corpus. Finally, note that all of the materials are tensed since it has been shown (e.g. Boxell 2009, unpublished) that untensed clauses may ameliorate wh-dependency formation in contexts which may be otherwise difficult to process, including long-distances such as these materials use.

Four presentation lists were created which contained one version of each critical sentence (i.e. one of the four critical conditions above). One (and only one) version of each of the twenty critical sentences was therefore present in each list. The four experimental conditions were consequently equally represented in each list such that every participant saw the same number of items for each condition. Each of the four lists was read by an equal number of participants.
3.1.2. Plausibility-norming questionnaire

Whilst every effort was made to ensure that, other than the critical manipulations, the critical sentences were identical across conditions, there remains an asymmetry. The subject of the critical subcategoriser differs between the [+Intermediate Representation] and [-Intermediate Representation] conditions. In (11a) and (11c) the subject which must be integrated with pleased is the new salesman while in (11b) and (11d) it is the secretary’s claim about the new salesman. It is possible that, across the materials, there may be instances where integrating one version of the subject may be easier than the other if it is more plausible to do so. This in itself should not pose much of an issue since the same plausibility discrepancy would exist for both D-linked and non D-linked materials. However, in the case of the D-linked sentences specifically, this difference in the verb’s subject may make the lexically-specified wh-phrase a more or less plausible object. For purely illustrative purposes, using (11c) and (11d), there may be some materials in which the equivalent of the new salesman pleasing a gentleman may be more plausible than a secretary’s claim about the new salesman doing so. See (14) in Section 3.4.1 for an experimental item for which such a difference was indeed found. Plausibility differences can have an effect on processing difficulty, and thus reading-time measurements (e.g. Gibson & Pearlmutter 1998). To control for this possible discrepancy, an offline questionnaire was disseminated to the participants after they had completed the online task. The critical items were simplifications of the materials, including the same intervening structural difference (i.e. +/- Intervening CP) but with no wh-dependencies. Thus, the lexical item which had been fronted as part of the D-linked wh-dependencies in the online materials was now presented as the direct object of the critical verb, but the discrepancy in the subject is still present as illustrated in (12):

(12)  (a) The secretary claimed that the new salesman had pleased the gentleman.
       (b) The secretary’s claim about the new salesman had pleased the gentleman.

Two lists were created of which each participant only saw one. One version or the other of each of the twenty critical sentences was present in each list such that each participant only saw one version of any given sentence and each condition was equally represented across the materials. Participants were asked to rate each sentence on a 1-10 scale, where 1 was the least well-formed and meaningful and 10 was the most. 40 fillers were also added to the questionnaire which consisted of a variety of sentences which had considerable range in the acceptability of their form and meaning.

3.1.3. Working-memory test

As stated in Section 2.2., it is possible that successive-cyclicality is of benefit to working-memory. Since the experimental hypothesis concerns testing for such successive-cyclicality in real-time, it makes sense to screen participants’ working-memory capabilities since the presence, absence or even reliance on intermediate representations of an antecedent could have been linked to such capabilities as opposed to any other linguistically-determined reasons.

The task was based on Daneman & Carpenter (1980). Participants were required to read through sentences written on cards one by one, turning the cards over as they did. Following each sentence, they had to decide whether it was grammatical or not by placing a tick or cross on an answer sheet. This ensured participants read each sentence in full and did
not focus purely on memorising the last word of each sentence (as they were also required to do). After a block of several sentences, participants would then be asked to recall the final word of each of the sentences just read. The blocks consisted of two, three, four and five sentences and were tested incrementally in this order. There were three blocks at each of these levels before moving on to the next. In short, the task examines participants’ ability to hold words in memory while processing further sentences.

3.2. Participants

40 native English speakers (mean age: 23; range: 18-25) were recruited from the student community of the University of Essex and surrounding area. 23 participants were female, all participants were right-handed, and all but 3 were monolingual. All participants reported that they had never been officially diagnosed with any language or general cognitive disorders or learning disabilities; neither did they suspect they had any such problems.

3.3. Procedure

The self-paced reading task was administered using a non-cumulative word-by-word set-up (Just et al. 1982) on the DMDX software package (Forster & Forster 2003). Participants used a pacing button to read through the materials segment-by-segment as quickly as they could, revealing the next segment and removing the previous one as they did. The experimental trials were segmented into parts as is illustrated in (13).

(13) + | The | manager | wondered | {who/which gentleman} | the | {secretary/secretary’s} |
  1 2 3 4 5 6 7
{claimed/claim} | {that/about} | the | new | salesman | had | pleased | in | the | meeting.
  8 9 10 11 12 13 14 15 16 17

The computer recorded the reading time for each segment, on the premise that higher reading times for a segment (or its spill-over region thereafter) correspond to higher processing difficulty for that segment. Following every critical item and half of the fillers, a “yes-no” comprehension question was asked to ensure the preceding sentences were being processed and thus comprehended correctly. Participants responded to these questions using separate “yes-no” buttons. Half of the questions had a correct “yes” answer while the other half had a “no” one, the distribution of which was randomised across materials. At the start of the self-paced reading task, on-screen instructions and three practice items were presented to ensure participants had correctly understood the task. Following this, participants were instructed to press the spacebar to begin the actual trials. They could ask the researcher task-clarification questions before doing so. They were offered a break every twenty items to prevent lethargy from affecting performance, pressing the spacebar again when they were ready to recommence trials. The stimuli were presented on a grey screen using black Arial (24pt) font for the sentences and Times New Roman (24pt) for the comprehension questions. Text was presented in the centre of the screen, and each item was preceded by a cross (“+”) in the centre of the screen to focus participants’ eyes and attention. Following the self-paced reading task, participants next completed the working-memory task and finished with the off-line questionnaire. A short break was offered between each task, with sessions lasting no longer than one hour overall.
3.4. Results

3.4.1. Working-memory and plausibility norming

All participants scored between 83.33% - 100% (mean: 94.32%, SD: 1.84%) for correctly recalling memorised words in the working-memory task. As such, all participants are considered to have a high-span working memory. Furthermore, the (percentage) median of participants’ scores was used to split the 40 participants into two groups: the participants with scores below the median (i.e. the lower scores) were one group, while the ones above the median (i.e. the higher scores) were the other. A 2x4 ANOVA comparing the two groups’ average performances at each of the four block levels was carried out, showing no significant differences between the groups in their ability to recall target words: $F(1, 38) = .995, p > .05$

In sum, we can be reasonably confident that there are no significant variations in working-memory capacity among the participants tested. Since intermediate representations are likely to benefit working-memory, one may argue that if bypassing them is possible in any parsers at all, it is more likely to be in ones with access to a high-span working-memory. As such, the chance of finding supportive evidence for Pesetsky’s (1987) thesis is most likely maximised in participants of this type.

Regarding the plausibility-norming questionnaire, each of the 20 items were analysed individually. The participants’ ratings for the [+/- Intermediate Representation] conditions for each item on the 1-10 scale were individually compared by $t$-test. Only the comparison for item 12 was statistically significant: $t(39) = 21.299, p < .01$. This item is reproduced here as (14), and ratings for (14a) averaged 9 while for (14b) they averaged 3. This suggests that integrating charity as object of the verb attacked is more plausible when the subject of that verb is the TV journalist compared to when it is the minister’s statement about the TV journalist.

(14)  
(a) The minister stated that the TV journalist had attacked the charity.  
(b) The minister’s statement about the TV journalist had attacked the charity.

Since this plausibility difference could explain any reading-time or comprehension-question response accuracy differences between the D-linked structures with and without an intermediate CP, all data for item 12 were excluded from any further analysis. In sum, we may conclude that plausibility differences for integrating a fronted D-linked wh-phrase with its subcategoriser where the subject of that subcategoriser differs as per the [+/- Intermediate Representation] manipulation do not account for any observed differences in the remaining data.

3.4.2. Comprehension-question response accuracy

The mean percentages of correctly answered comprehension questions are given in Table 2 by experimental condition.

<table>
<thead>
<tr>
<th>Intermediate CP structure</th>
<th>Non D-linked Wh-dependencies</th>
<th>D-Linked Wh-dependencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>94.43%</td>
<td>89.98%</td>
</tr>
<tr>
<td>-</td>
<td>91.11%</td>
<td>92.16%</td>
</tr>
</tbody>
</table>

*Table 2: Average correctly answered comprehension-questions (%)*
2x2 participants and items ANOVAs comparing +/- D-linking conditions and +/- Intermediate CP structures did not yield any significant differences. Thus, participants seem to be equally good at answering these questions irrespective of the experimental manipulations. This is indicative of equally accurate overall comprehension of the sentences across the materials.

3.4.3. Reading times

The following data analysis is based on participants’ residual reading times. These times are the participants’ actual reading-times minus their predicted reading time as calculated by a regression equation. Such equations used all of a particular participant’s data for words of specified lengths to estimate (all-other-things-being-equal) the reading time for a word of a particular length for that particular participant. Thus, residual reading times are relativised along two dimensions: (a) for the participant in question; (b) for the word-length of each segment. Therefore, differences in generic participant reading-speed on the one hand, and word-lengths across the materials on the other, are neutralised by the adjustments which the use of residual reading times provide.

The two particular segments which are of importance in attempting to identify the presence or absence of an intermediate representation of an antecedent are nine and fourteen. Recall the materials were segmented as indicated in (13), repeated here as (15). Segment nine is that in which either that or about is located. Since that indicates the commencement of the CP complement of claimed it is thought that, if an intermediate representation is activated at the CP boundary, the processing costs involved in doing so should result in an increased reading time here. This would not be expected for about since this does not correspond to a CP boundary.

(15) The | manager | wondered | {who/which gentleman} | the | {secretary/secretary’s} |
    {claimed/claim} | {that/about} | the | new | salesman | had | pleased | in | the |
    meeting.

Segment fourteen is the verb with which the fronted wh-phrase is associated, in the above example this is pleased. The fronted wh-phrase should be reactivated at this subcategoriser by way of completing its dependency. However, the conditions in which such wh-dependencies have an intervening CP have a posited intermediate activation of the antecedent which the ones without the intervening CP do not. It is predicted that a local representation of the antecedent at the intervening CP should speed up the subsequent reactivation at the subcategoriser since the antecedent would already be somewhat “active” in the mind. Thus, reading times at the subcategoriser, as the reactivation is instantiating, should be faster for the conditions with an intervening CP relative to the ones without. Of course, with respect to the predictions for both segments nine and fourteen, the hypothesis must be refined with respect to Pesetsky (1987). If these data are to be supportive of this theory, in which D-linked wh-dependencies do not involve intermediate representations of the antecedent, evidence of such intermediate representations should be restricted to non D-linked wh-dependencies.

Only reading time data for sentences for which participants correctly answered the comprehension question were included in the following analysis. Residual reading times greater than 1000ms were also removed from the data included in the analysis, affecting 0.28% of the data. Recall also that all data pertaining to item twelve were excluded from analysis on the grounds set out above in Section 3.4.1. affecting an additional 5% of the data.
Figure 1 displays the mean residual reading times (across all 40 participants) for all the remaining data for each segment.

As the graph in Figure 1 demonstrates, average residual reading times for non-critical segments seem numerically similar for all conditions. The variation among the reading times for each condition seems to be mainly located in segments nine and fourteen (i.e. the critical segments) and the regions immediately following them (i.e. segments ten and fifteen). As such, the following statistical analysis focuses on these segments.

In segment nine, *that* in the sentences with the intervening CP was read slower than *about* in the sentences without the intervening CP. As such, the reading time difference between the conditions with the intervening CP compared to those without the intervening CP was significant according to the participants analysis, $F_1(1, 38) = 21.433$, $p < .001$, although it fell short of significance in the items analysis, $F_2(1, 17) = .638$, $p > .05$. Additionally, segment nine is read faster when the antecedent of the *wh*-dependency is D-linked relative to the condition in which it is not. This is true for both the sentence types with and without an intervening CP, and yields an overall significance of D-linking for segment nine, $F_1(1, 38) = 445.763$, $p < .001$; $F_2(1, 17) = 14.661$, $p = .005$.

According to the participants analysis there is also a significant interaction between +/− intervening CP structures and +/−D-linking, $F_1(1, 38) = 50.736$, $p < .001$, but such interaction is not significant in the items analysis, $F_2(1, 17) = 1.511$, $p > .05$. Planned comparisons reveal the faster reading time at segment nine for the D-linked constructions with an intervening CP is significantly different from the non D-linked constructions with an intervening CP in the participants analysis, $t_{1}(38) = 10.603$, $p < .001$, and marginally significant in the items analysis, $t_{2}(17) = 2.082$, $p = .076$. Similarly, the comparison of the +/− D-linked constructions without an intervening CP yields a significant difference at segment nine, with the D-linked one again being read faster than the non D-linked one, $t_{1}(38) = 14.427$, $p < .001$; $t_{2}(17) = 2.470$, $p < .05$.

The same model of statistical analyses was used for segment ten. Since it immediately follows segment nine, the numerical differences found here may include statistically significant “spill-over” effects from what has just occurred before it. Indeed, the slower reading times for the [+Intervening CP] condition continues to be statistically significant from the faster [−Intervening CP] ones in the participants but not the items analysis, $F_1(1, 38) = 5.620$, $p < .05$; $F_2(1, 17) = .121$, $p > .05$. As with segment nine, the D-linked conditions are read faster than the non D-linked ones when the intervening construction does not include a CP but this effect drops away when it does. As such, the effect of D-linking is only significant for the participants analysis, $F_1(1, 38) = 6.735$, $p < .05$; $F_2(1, 17) = .892$, $p > .05$. In this segment, there was no significant interaction between +/− intervening CP structures and D-linking, $F_1(1, 38) = .989$, $p > .05$; $F_2(1, 17) = 1.019$, $p > .05$. Planned comparisons reveal that the faster reading time for the D-linked condition with an intervening CP relative to the non D-linked condition with an intervening CP retains significance in the participants analysis but not the items analysis, $t(38) = 3.216$, $p < .05$; $t(17) = -.673$, $p > .05$. The faster reading time for D-linked constructions relative to non D-linked ones without intervening CP structures retains its significance from segment nine, $t(38) = 12.435$, $p < .001$; $t(17) = 2.006$, $p < .05$. In sum, the statistical effects found in segment ten follow a similar broad pattern as segment nine but with some dropping off as the numerical differences among the conditions reduces. By the time segment eleven is reached almost all of the numerical and statistical differences between the experimental conditions have disappeared.
The main critical segment is fourteen. This is the subcategoriser (e.g. pleased) with which the fronted wh-constituent is associated, and thus should yield a reactivation thereof. Recall that a localised representation of the antecedent at an intermediate CP is thought to ease this reactivation process, making for shorter reading times at the subcategoriser relative to conditions where no such intermediate CP is accommodated (see Gibson & Warren 1999, 2004; Marinis et al. 2005). Overall, constructions with an intermediate CP were read significantly faster at this segment compared to those without, $F_1(1, 38) = 39.471, p < .001; F_2(1, 17) = 14.819, p = .005$. There was also a significant difference between the reading times of those subcategorisers which terminate a D-linked dependency relative to those which are non D-linked, with D-linked ones being read slower than non D-linked ones, $F_1(1, 38) = 432.214, p < .001; F_2(1, 17) = 144.757, p < .001$. The interaction between the intervening structure effect and the D-linking effect was, however, not significant, $F_1(1, 38) = .588, p > .05; F_2(1, 17) = .221, p > .05$.

Planned comparisons compared reading times at the critical subcategoriser for structures with an intervening CP and those without an intervening CP, firstly when those structures were not D-linked and secondly when they were D-linked. As expected, following Gibson & Warren (1999, 2004) and Marinis et al.’s (2005) findings, the shorter reading time of the subcategoriser when preceded by an intermediate CP structure compared to when preceded by a structure without an intermediate CP was indeed significant when those structures were not D-linked, $t_1(39) = -21.326, p < .001; t_2(18) = -10.432, p < .001.$
Particularly pertinent with respect to Pesetsky’s (1987) hypothesis is whether this difference also holds where the constructions were D-linked. Once again the reading times for the subcategoriser were significantly shorter when preceded by an intermediate CP structure relative to when preceded by a structure without an intermediate CP, $t_1(39) = -9.303, p < .001$; $t_2(18) = -6.108, p < .005$.

As Figure 1 shows, many of the conditions “peak” in their reading time lengths in segment fifteen, with the only exception being the condition with a dependency which is D-linked without an intervening CP which took the longest to process at segment fourteen. This may suggest that at least for the other conditions, for which reading times were quicker at the subcategoriser, some of the antecedent reactivation processing may spill-over into segment fifteen, creating the increase in reading time seen there. The main effect of the difference between intervening structure (+/- CP) can be found in the participants analysis only in this segment, $F_1(1, 38) = 4.190, p < .05$, as is also the case for the main effect of D-linking, $F_1(1, 38) = 7.767, p < .01$. It would seem that this residual statistical significance which spills-over from segment fourteen is driven by the non D-linked structure with an intermediate CP continuing to be read faster than the segment fifteen of any of the other conditions. Indeed, the only significant result from the planned comparisons for this segment was the participants analysis comparing the non D-linked structures with and without an intermediate CP, $t_1(39) = -2.645, p = .012$. Segment fifteen reading times generally, then, appear to have already converged, losing most of the distinctive effects they may have had previously. Of course, the main point to draw from segment fifteen and furthermore segments sixteen and seventeen (as seen in Figure 1) is that the effects seen at the critical subcategoriser in segment fourteen rapidly disappear and residual reading times for the subsequent segments return to a “normal” pace, wherein the differences between the experimental conditions are not significant.

4. Discussion and conclusions

4.1. Pesetsky’s hypothesis

The data collected for segment fourteen (i.e. the critical subcategoriser) indicate that both non D-linked and D-linked dependency formation includes intermediate representations of the antecedent at CP clause boundaries. This can be interpreted from the significantly faster reading times at the subcategoriser for conditions with an intermediate CP compared to those without. The intermediate CP accommodates an intermediate representation of the antecedent (i.e. most formalisms agree such structures have the “syntactic space” for such representations) while structures without the intermediate CP do not. The rationale was that localised activation of an antecedent (at the intermediate CP) will boost its prominence in the parse, facilitating faster reactivation of that same antecedent at the subcategoriser. This would not be possible for constructions which do not include intermediate representations, and would yield slower reading times at the subcategoriser.

That this difference is found as much for the D-linked as the non D-linked dependencies is particularly problematic for Pesetsky (1987). Recall that this hypothesis states that the ameliorative effect of D-linking on island constraint violations is attributable to the fact that it should be possible for D-linked dependencies to form in a binding-type operation rather than a Chomskyan “movement” (or non-transformational “slash-pathway”) style of dependency formation. This is because constraints of this type, it was reasoned, specifically apply to “movement”. This experiment examined long-distance dependency formation because, in this environment, the nature of binding-type and the nature of movement and slash-pathway type dependencies differ in a particularly crucial way: the former unfolds in a
single operation which can span multiple clauses, while the latter are predicted to require intermediate reactivation at clause-boundaries by way of unfolding successive-cyclically. Thus, if Pesetsky (1987) were correct, D-linked long-distance dependency formation should not be affected by whether or not the intervening structure of that dependency accommodates an intermediate representation. The significantly faster reading times for D-linked antecedent reactivation at the subcategoriser when the preceding structure contained an intermediate CP relative to when it did not, however, suggests D-linked wh-dependencies are affected by intervening structure in the same way non D-linked dependencies are. In turn, this indicates D-linked dependencies are not formed by single-step binding but instead follow the non D-linked dependencies (which show the same difference) in using a multiple-step movement or slash-pathway form of dependency in which antecedent reactivation at the subcategoriser can benefit from localised representations at an intervening CP.

Recall that the data from segment nine may also bear on Pesetsky’s hypothesis in as much as the instantiation of an intermediate representation in this segment - at least in the conditions which accommodate it - may yield longer reading times than those where such instantiations are not possible. Indeed, constructions with intermediate CP structures were read significantly slower than those without. However, this difference must be treated carefully. Segment nine in the [+Intermediate CP] conditions was that while in [-Intermediate CP] it was about. Thus, the difference observed here could result from the fact that they are different words from different word classes with different frequencies and other properties. Furthermore, and perhaps more significantly, that marks the start of a new clause while about does not. As many studies (e.g. Kluender & Kutas 1993) have shown, “wrap-up” procedures at the end of one clause can incur additional processing costs, while the predictions the parser may make at the start of a new clause about the structure to come may also consume processing resources (see Gibson 2000). Any or all of these factors could explain the increased reading time for the conditions with an intermediate CP, without necessarily invoking the issue of intermediate representation instantiation costs at all.

In a bid to combat such confounds for interpreting segment nine data, planned comparisons examined the two intervening CP conditions (+/- D-linked) and the two conditions without an intervening CP (+/- D-linked). The non D-linked constructions with an intervening CP were read significantly slower than their D-linked counterparts. It is unlikely, however, that this is caused by setting-up an intermediate representation of the antecedent in the non D-linked condition while not doing so in the D-linked condition. This is because the comparison of the two conditions without an intervening CP also yielded a similarly significant difference, and neither of these latter conditions can accommodate an intermediate representation of the antecedent in any case. Thus, the difference must be explicable in some other way which does not involve intermediate representation instantiation at all. Rather, the effect of D-linking itself would seem to drive the differences between both pairs of conditions. An explanation is offered in Section 4.2.

In sum, it seems that the increased reading times found for constructions with an intervening CP cannot necessarily be attributed to setting-up an intermediate representation and the difference between the two conditions with intervening CPs cannot necessarily be attributed to one setting-up an intermediate representation while the other does not. Thus, while data were collected for segment nine with the intention of using them to comment on the presence or absence of intermediate antecedent representations, it would seem they cannot contribute to that particular discussion with sufficient clarity. Instead, we must rely on the discussion of segment fourteen above. On that basis, then, it would seem these data are not consistent with the predictions generated from Pesetsky (1987).
4.2 D-linking: A “processing stabiliser”

If the discussion in Section 4.1 is correct, then two key questions remain unanswered. Firstly, if Pesetsky’s (1987) account of the ameliorative effects of D-linking appears to be inconsistent with these data, then what does cause said amelioration of constraints like islands? Secondly, why is segment nine read significantly quicker during the computation of D-linked dependencies whether or not that segment constitutes an intervening CP structure? What follows is one possible unified answer for both of these questions.

As Figure 1 illustrates, there is an effect of D-linking in segment nine, both for dependencies with an intervening CP and for those without. The effect is driven by the fact that the segment is read faster when the antecedent of the dependency is D-linked. Why should this be? The answer may reside in the fact that segment nine, in all of the critical materials, could very well represent a moment of potential complexity. Imagine if, while the parser is still searching for a reactivation or gap-filling site for the wh-filler, one of the two following scenarios occurred: (1) on encountering the structure with an intermediate CP, the intermediate CP position was occupied by another wh-phrase; (2) on encountering the structure without an intermediate CP (which instead has a complex DP subject), the object of the preposition about becomes a vacant gap position. (22a) illustrates the first of these scenarios and (22b) the second:

(16) (a) *?The manager wondered which gentleman the secretary claimed [CP who had pleased which gentleman] in the meeting.
   (b) *?The manager wondered which gentleman the secretary’s claim about which gentleman had pleased in the meeting.

In short, both of the scenarios are island constraint violations – (16a) is a wh-island violation and (16b) is a complex subject island violation. Segment nine occurs at the moment for potential island creation since it is either that, marking the intermediate CP which could have created a wh-island had it been occupied, or about which could have created a subject island had its object position been vacant. Furthermore, since these constructions have an on-going wh-dependency in formation, segment nine actually represents a moment of potential island violation. An awareness of the potential for the parse to crash at this point would likely carry a processing cost. Additionally, because that marks a new clause boundary, recall that there seems to be inherent processing costs associated with ending one clause and beginning a new one, also making segment nine (in the relevant condition) a likely moment of processing difficulty.

We might posit an albeit speculative hypothesis in which the faster reading times for D-linked dependencies at the time-locked moment of parsing difficulty is linked to the amelioration effects of D-linking on those very same complexities, i.e. islands and clause-boundary processing. Indeed, just as electrophysiological responses can be observed for “wrap-up” processes at clause boundaries (Kluender & Kutas 1993), the parser seems aware

---

1 Of course, since these are D-linked dependencies, the violation of these island constraints may seem more acceptable than would otherwise have been the case.

2 Note that Boxell (2009) found an amelioration effect of D-linking for wh-dependencies which included a clause-boundary and were entirely grammatical relative to ones which were also grammatical but were only single-clauses in an offline questionnaire. Furthermore, it has been generally established in the literature that long-distance dependencies can involve increased processing costs relative to short distance ones (see, e.g., Phillips et al. 2005).
of island violation there as well. McKinnon & Osterhout (1996) found ERP effects at the moment in which a *wh*-island boundary was parsed while a *wh*-filler had not yet been reactivated at an underlying position. Since this real-time sensitivity to the syntactic complexities themselves is so acute, it might seem plausible to suppose that sensitivity to features which can ameliorate those same phenomena can be equally time-locked to the relevant moment of the parse.

As discussed in Section 2.3., D-linking clearly includes an increased degree of lexical specificity for a *wh*-dependency’s antecedent, which becomes prominent in the discourse-structure of a sentence early in the time-course of its processing (Radó 1998; Frazier & Clifton 2002; and Diaconescu & Goodluck 2004). It might be hypothesised that it is this property which “stabilises” the parse. I define a “stabiliser” as any property of language which adjusts the scope of the discourse-conceptual structure related to a parse which subsequently makes the successful completion of that parse, namely deducing its overall meaning, more likely. The parse may be thought of as possessing increased levels of “stability” when such properties occur since the processor may be able to rely more heavily on this discourse-conceptual structure as an information source to complete the parse, enabling it to be (at least to some extent) desensitised to a range of constraints and difficulties which may have otherwise been problematic, maybe even causing the parse to crash. That is, the early-instantiated pragmatic and/or semantic properties of the D-linked antecedent, and the resultant plausibility implications for the way in which the dependency may feasibly unfold could very well be a sufficient “stabiliser” to permit an avoidance or ignorance of the island constraint *as and when* its boundary is encountered. Having a D-linked antecedent means the possibilities for how the parse may feasibly proceed are restricted by the parts of the discourse-structure which are constructed early in the parse, and as such the main target of sentence comprehension processing, namely deducing the intended meaning of the input dependency (i.e. *who* did *what to whom*) is now simplified by virtue of having been partially completed. The parsing of the dependency can now only proceed in a way which is compatible with the plausibility cues which may be extrapolated from that partially completed discourse-structure. Thus, the deductive power of the parser for establishing how the dependency should be formed and ultimately what meaning should be extracted from it is boosted. Clearly, if meaning is transparently deducible from the structure, the relevance of any constraints which may have otherwise had an effect on it (e.g. islands) may be diminished.

In short, it could be this “stabilising” which yields the ameliorative effect of D-linking, and this may have an observable psychological reality at critical moments of complexity in the parse in which sensitivity to (potential) complexities is reduced, meaning parsing may continue to proceed in a timely fashion. This is what may be observable in segment nine. The data in Figure 1 may even lead us to assert that reading times can actually speed up at such a critical moment in the parse during D-linked dependency formation. Maybe this is a bid to literally “jump-over”, “speed-past” or “avoid” the location of a possible syntactic constraint violation with which the parser already knows it need not be concerned.

The data from segment fourteen may support the idea that the lexical properties of a D-linked antecedent are indeed transmitted across the entire *wh*-dependency. That is, these properties seem to be held in working-memory and are reconstructed at the underlying reactivation position at the subcategoriser. This would explain the main effect of D-linking found in segment fourteen, whereby the D-linked conditions were read more slowly than their non D-linked counterpart conditions. The lexical specificity of D-linked antecedents may cause increased reading times at the underlying subcategoriser since their semantic and
pragmatic properties would have to be reconstructed, and the plausibility of its integration with the subcategoriser and its subject would need evaluating in addition to which lexical access processes may also be needed. This is a finding which converges with that of Shapiro et al. (1999), who found delayed cross-modal priming effects for D-linked antecedents at underlying gap positions relative to non D-linked controls, among others (see Section 2.3.).

There is, however, a crucial caveat to this “stabiliser” hypothesis. As mentioned above, it is currently only a mere hypothetical speculation. Since it was not the hypothesis under investigation in the present study, there was no control in which there was a segment nine equivalent which was not a potential island or complexity. Such a condition would be a prerequisite to attempt to show that D-linking “stabiliser” effects are indeed time-locked specifically to moments where problems or difficulties like (potential) islands are encountered by the parser. In short, a great deal of further research is required in order to pursue the “stabiliser” hypothesis of D-linking. This first pass at such an idea, however, allows us to posit a feasible explanation of the main effect of D-linking found in segment nine on the one hand, and a feasible explanation for how and why D-linking has the ameliorative properties that it does, and even a time-course for those properties, on the other.

The “stabiliser” hypothesis, if correct, may suggest that the ameliorative properties of D-linking may be explained by its processing rather than by relying on a formal syntactic account like Pesetsky (1987) at all. However, it may worth noting that there are alternative formal accounts which may explain the properties of D-linking without resorting to (unselective) binding. Pesetsky (2000), for example, posits that D-linked wh-phrases may (optionally) use feature movement to satisfy locality constraints like Superiority (Chomsky 1973) or Attract Closest (Richards 1997) while the surface word string itself continues to give the appearance of a Superiority violation. See Shields (2008) for a similar formal account, and Van Craenenbroeck (2004, 2010) which suggests complex wh-phrases may have successive-cyclic empty-operator movement. In sum, it would seem that while the present paper suggests Pesetsky (1987) may not be an adequate theory of D-linking, further research is required to examine the other theories of both performance and competence which may be.

4.3 Different types of antecedent reactivation

There would seem to be a distinction between the two main effects of D-linking found in the present study. At segment nine, it has been suggested that the lexical properties of D-linked dependencies may stabilise the possible complexities found there as an explanation for the faster reading times for the D-linked conditions relative to their non D-linked counterparts. At segment fourteen, the need to reconstruct those lexical properties at the subcategoriser have been cited as an explanation for why reading times are slower for the D-linked conditions. If both of these interpretations are indeed correct, along with the indications found (in segment fourteen) that D-linked dependencies do include antecedent reactivation at the intermediate CP, then the type of reactivation found at the intermediate position is likely to be different from that found at the underlying subcategoriser. Full reactivation of a D-linked antecedent at its underlying verb - lexical specificity included - causes a slowing of reading time to below that of the non D-linked conditions. The D-linking effect found at segment nine, which is inclusive of, but not exclusive to, intermediate reactivations of the antecedent includes faster reading times of the D-linked conditions relative to the non D-linked ones. Thus, if full reactivation of the lexical specificity causes a slowing of the reactivation of the D-linked antecedent at the verb, it would seem that full lexical reactivation is unlikely to be occurring at the intermediate position. Any amelioration of (potential) complexities found at segment
nine which may be caused by the lexical specificity of a D-linked antecedent as suggested in my “stabiliser” hypothesis is presumably driven by the lexical information which is held in the background working-memory involved with forming any (D-linked) wh-dependency per sé, rather than being associated with a specific reconstruction of the antecedent at this position. This would be consistent with the fact that the segment nine D-linking effect is entirely independent of whether or not an intermediate representation of the antecedent is present, or indeed possible.

In sum, then, it would seem that antecedent reactivation at intermediate positions within a wh-dependency is distinct from that at the underlying subcategoriser position, particularly with respect to the reconstruction of lexical information. This finding potentially bears on several other debates in the sentence processing literature. As suggested in Section 2.2., the inability of Finn, Kim & Piñango (2010) to find evidence of intermediate representations of an antecedent using cross-modal priming could be because those intermediate gaps are purely structural, not involving the semantic content on which cross-modal priming depends. The dissociation between intermediate and subcategoriser gaps found in the present study with respect to lexical content could support the notion that the intermediate reactivation of an antecedent is indeed lacking in the lexical-semantic information which would drive a cross-modal priming effect. Similarly, in Section 2.2., it was suggested that if evidence could be found to show that intermediate reactivation is purely structural (and devoid of semantic information) it could help to resolve a problem identified by Fodor (1995). That is, since antecedent reactivation (at the verb) could rely purely on correctly linking a dislocated constituent with its verb semantically, it is difficult to isolate whether or not structural reactivation exists during real-time processing. While the pre-existing literature’s support for successive-cyclicity goes some way in isolating an occasion where antecedent reactivation is inherently less able to rely on the semantic content of the verb, the present study’s D-linking manipulation may isolate an actual behavioural difference between the nature of intermediate and verbal reactivation during parsing. Namely, the lexical-semantic content of an antecedent seems to be selectively reactivated at the verb and not at intermediate positions. Finally, online evidence of successive-cyclic dependency formation supports a strong version of the Active Filler Strategy (Clifton & Frazier 1989). Under this principle of on-line dependency formation, an active wh-word (or filler) triggers a search for the nearest potential reactivation site (or “gap”) which can be filled by a representation of the antecedent. Long-distance dependencies, then, can include intermediate gap-filling so as to divide up the formation into several shorter parts. Direct Association (Pickering & Barry 1991) posits a linking of the fronted antecedent directly with its subcategoriser and so it does not account for intermediate reactivations. Furthermore, Direct Association may rely more on the conceptual links between the antecedent and its underlying position, which would clearly not account for purely structural intermediate reactivation.

4.4. Concluding remarks

The present study set out with the intention of searching for the psychological reality of some predictions made by a formal account of D-linking, namely that its wh-dependencies can form via (unselective) binding (Pesetsky 1987). This would have explained why such dependencies are less sensitive to “movement” constraints like islands. However, at clause-boundaries, D-linked antecedents behaved like non D-linked ones and were reactivated. This was evidenced by the fact that antecedent-subcategoriser integration was faster in constructions with an
intervening CP, facilitating a mediating representation of the antecedent, compared to constructions without an intervening CP. Intermediate antecedent representations are thought to speed up subsequent subcategoriser reactivation since they boost the prominence of the argument in the parse, and reduce the working memory cost of the dependency formation. Since intermediate representations are a feature of movement-type dependency formation, but not of binding, it would seem that Pesetsky’s (1987) account is inconsistent with these findings. While other formal accounts of D-linking have been put forward (e.g. Pesetsky 2000), the segment nine data of this study has motivated the (speculative) hypothesis that the early-discourse instantiation which D-linking provides, and has often been alluded to as the source of its amelioration effects (Radó 1998; Frazier & Clifton 2002; Diaconescu & Goodluck 2004), may in fact have specific time-locked effects. That is, when the processor encounters a moment in which constraint violation is a possibility and/or other processing complexities may be present, a D-linked dependency is “stabilised” by virtue of having more of its overall discourse-conceptual structure already assembled. It will be interesting to see if such stabiliser effects are indeed unique to moments of potential complexity and if they can be linked to the amelioration of constraint violations. Finally, this study has provided potential evidence of the psycholinguistic distinctiveness of intermediate gaps from their subcategoriser counterparts in that it is likely that only the latter have lexical content.

Acknowledgements

I am sincerely grateful to Claudia Felser for her insightful advice and encouragement during the completion of this work, and for making the materials used by Marinis et al. (2005) available. I am also grateful for comments from members of the Psycholinguistics Research Group at the University of Essex and from audience members at the 19th Conference of the Student Organisation of Linguistics in Europe and the 4th Quantitative Investigations in Theoretical Linguistics conference. I would also like to thank Eefje Boef for her useful review. Naturally, I assume full responsibility for any errors.

Oliver Boxell
University of Essex
osboxe@essex.ac.uk

References


