Island constraints have formed a central component of grammatical theory since the groundbreaking work of Ross (1967) developing a proposal in Chomsky (1964). The papers collected in this volume address island constraints from a variety of theoretical linguistic and psycholinguistic perspectives. They result from a conference held at the University of Ottawa during the fall of 1989, at which specialists from several fields met to engage in interdisciplinary dialogue over the nature and manifestations of island constraints. In this brief introduction, we will attempt to summarize the intellectual setting that serves as backdrop to the investigations reported on here. To do so both fully and briefly presents a difficult challenge. But we think the inevitable risk of misrepresentation and offense is worth the effort. For it is only by making results accessible outside their domains of specialization that we can hope to foster the cross-disciplinary dialogue that we believe so essential to the growth of our knowledge and understanding of the mechanisms of language.

I. ISLANDS AND GRAMMATICAL THEORY

Syntactic accounts

Many languages, including English, display syntactic constructions in which a phrase in non-argument position is associated with a vacant argument position elsewhere in the sentence for purposes of semantic interpretation. Well studied English examples include matrix and em-
bedded *wh* questions, topicalization, and relative clauses, as illustrated in the respective sentences below.

(1) (a) In which room did you decide to put it _?
    (b) I never realized how happy he makes her _.
    (c) John, I don't think I like _ very much.
    (d) This is the man who I told you about _.

In (a), the *wh* PP is an understood complement of *put*, in (b) the *wh* AP is an understood complement of *make*, in (c) the sentence initial NP is the understood object of *like*, and in (d) the relative pronoun is understood as the object of *about*. In all these cases, the dislocated phrase (the antecedent) is in a non-thematic non-argument position, and is understood to fill a thematic function elsewhere in the sentence (signalled by ' _ ' in these examples). (In general, an argument position expresses a thematic role designated by the verb.) Moreover, the thematic (argument) position it is interpretively linked to is not filled with any lexical material — it is a gap.

A central feature of the constructions in (1) is that the relation of antecedent to gap is apparently unbounded, in the sense that there is no theoretical limit (though there is certainly a practical one) to the amount of material that may intervene. In (2) for example, the dislocated sentence initial phrase is related to a gap across at least three levels of clausal embedding.

(2) John, I think you told me (that) Bill said (that) Mary claimed (that) Sally is planning to recommend _ for a job.

But the lack of bounding in such cases is only apparent. This is perhaps best appreciated when such cases are contrasted with comparable examples in which there is no bounding effect whatsoever. For instance, consider the *as for* example in (3a), and the syntactically and interpretively equivalent example of topicalization in (3b).

(3) (a) As for John, I have a good friend who likes him.
    (b) *John, I have a good friend who likes _.

The contexts in which this bounding of antecedent-gap relations surfaces have been characterized since Ross (1967) in terms of 'island
constraints', which serve to block the association of antecedent to gap under specific syntactic conditions. Of the restrictions identified by Ross and others, the ones we will discuss here are the Complex NP Constraint, exemplified with a relative clause in (3b) and with a nominal complement in (4a), the Subject and wh Island Conditions (Chomsky, 1973) in (4b, c) respectively, and the Adjunct Island Condition (see Huang, 1982's Condition on Extracion Domain), illustrated in (4d, e).

(4) (a) *John, Mary made the claim that Sally plans to recommend _
for a job.
John, Mary claimed that Sally plans to recommend _ for a job.
As for John, Mary heard the rumor that Sally intends to
marry him.

(b) *John, an article about _ just appeared in the newspaper.
As for John, an article about him just appeared in the
newspaper.

(c) *Bill, I wonder who likes _.
As for Bill, I wonder who likes him.

(d) *The heat, we left early because of _.
As for the heat, we left early because of it.

(e) *The money, I lied so that I could keep _.
As for the money, I lied so that I could keep it.

Chomsky (1973) proposes to reduce a number of these restrictions to a single grammatical principle, Subjacency, governing the association of dislocated phrases and their respective gaps. Subjacency imposes a locality restriction on this association to the effect that a dislocated phrase may be related to a gap across no more than a single 'bounding node', which may be defined for English as any instance of NP or S that dominates the gap and not the antecedent. Chomsky proposes a further principle, that the COMP position of a clause may mediate between a higher antecedent and a lower gap only if the COMP is neither occupied by a wh phrase or wh complementizer nor is used to mediate between some other antecedent and gap. Assuming further that potential violations may be saved only through the provision of available mediating COMP positions, these principles together subsume the
effects illustrated in (3) and (4a–c) and some others as well. (See Chomsky, 1977, 1981 for discussion.) We return below to the fact that the Adjunct Island effect in (4d, e) must be independently stipulated on this account (see Huang, 1982). The application of Subjacency is illustrated in the abbreviated surface syntactic representation in (5) below.

(5) \[ \text{John} \{ \text{I have} [\_{NP} \text{a good friend who} [\_S \text{likes } e]]] \]

In (5), the relationship between the dislocated phrase John and the gap e across the multiple instances of NP and S in the representation is blocked by Subjacency.

In modern syntactic theories island constraints have received variable attention. We will now very briefly consider developments in the three principal theoretical frameworks, Lexical Functional Grammar (LFG), Generalized Phrase Structure Grammar (GPSG), and Government and Binding theory (GB). Since island constraints have received the least attention in LFG, it is perhaps easiest to begin there. Early work in LFG adopted a conventional structural view, adapting existing phrase structure accounts in the c-structure (constituent structure) representations of that theory (see Kaplan and Bresnan, 1982; Zaenen, 1980, 1983). A revised view is provided in Kaplan and Zaenen (1989), where it is argued that island constraints are functional in nature, and so more naturally expressed as conditions on f-structure (functional structure) rather than c-structure representations. Kaplan and Zaenen propose a formal device for expressing functional uncertainty in long distance dependencies, annotating f-structures containing the discourse functions TOPIC and FOCUS (see Bresnan and Mehombo, 1987). Kaplan and Zaenen’s f-structure analysis is taken to emphasize the essentially functional nature of the Adjunct Island Condition. Further work is needed in LFG to see whether other (perhaps all) island effects find such natural expression in f-structure.

Island constraints have received somewhat greater attention in GPSG, but the view that the various island effects witnessed earlier are best characterized through a unified analysis (Subjacency in Chomsky’s account) is not shared by proponents of this theoretical paradigm. GPSG offers phrase structure grammar accounts of syntax and the properties of grammatical constructions. Feature information is shared between mothers and daughters in local (context free) tree configurations, subject to general principles and conditions of well-formedness.
Dislocated phrases are related to non-local gaps through the use of a feature [SLASH] which appears on all and only those nodes that lie on the path between the dislocated phrase and the gap. Gazdar (1981) argues that the categorial distinction between [SLASH] indicated nodes and others provides a natural account of the Coordinate Structure Constraint of Ross (1967) (*I wonder what Bill ate and left for school,) and of the Across the Board (ATB) exceptions to it (*I wonder what Bill ate and Sally cooked.) Gazdar suggests phrase structure grammar accounts for a range of other effects properly seen as island effects in this view, including the Left Branch Condition (*Whose did you buy car?) and the that-1 effect (*Who do you think that left?). The [SLASH] feature plays a role again in the Gazdar et al. (1985) (henceforth GKPS) account of subject (4b) and adjunct (4d, e) islands. GKPS propose to treat this feature as both a head and a foot feature, and as such subject to the distributional requirements of both the Head Feature Convention (HFC) and the Foot Feature Principle (FFP). The FFP requires that foot features which appear on a daughter must also appear on the mother, while the HFC requires head features that appear on the mother to also appear on the head. Thus configurations in which a [SLASH] feature appears on a mother and on a non-head daughter will be barred unless the slash feature also appears on the head. This prevents locating a gap within a subject (if VP is the head of S) or within an adjunct, unless licensed by a parasitic association with a gap within the head (compare (4b, d, e) and . . . the people who pictures of embarrassed, . . . the articles that I filed without reading, . . . the people that I hired because I liked). Accounts of island effects in GPSG have hinged largely on the interaction of specific head and foot features with general principles of feature instantiation that have independent motivation in the theory.

Within GB, Chomsky's general strategy of reducing island constraints to a single overarching generalization governing the association of dislocated phrases and their associated gaps in constructions of unbounded dependency has persisted in subsequent theoretical development of the principle of Subjacency. Chomsky (1986) proposes a version of Subjacency in which bounding nodes are supplanted by barriers. Oversimplifying his proposal considerably, let us suppose that a barrier for a given node α is a maximal projection dominating α that is not L-marked. L-marking is a relation based on the assignment of (internal) θ-roles. Assume that only the subcategorized complements to
a head are L-marked, and, with Lasnik and Saito (1989), that VP is L-marked by I(NFL). Then adjuncts, subjects, and the IP (= S) complement to C(OMP) remain non-L-marked, and so qualify as barriers. Now suppose that Subjacency is formulated to block unbounded dependencies across more than one barrier. The Subject and wh Island Conditions (4b, c) will follow because of the barrierhood of subjects and IP's. Moreover, the adjunct islands (4d, e) and the relative clause case (3b) of the Complex NP Constraint are accommodated as well, by the barrierhood of adjuncts. The nominal complement case of the Complex NP Constraint (4a), accommodated on the bounding nodes view of Subjacency, remains unaccommodated on this account. In general, these are not so blatantly ungrammatical as their relative clause counterparts. We might suppose that any clause contained within NP will constitute a weak barrier independent of L-marking (see Chomsky, 1986). Or, we might assume that the variation in island effects in nominal complementation is properly attributed to the ‘bridge’ phenomenon, as discussed below. In other words, extraction from the clausal complement to a noun is in principle possible, with variation in actual acceptability a function of bridging.

Cross-linguistic variation

The most widely known study of Subjacency outside English is that of Rizzi (1982). Rizzi observes that although there is evidence of the application of Subjacency in Italian in the case of complex NPs, Italian shows systematic differences from English in the realization of the wh island and subject conditions. Consider for instance the following examples from Rizzi (1982).

(6) (a) *Questo incarico, che non sapevo la novità che avrebbero affidato a te, . . .
**This task, that I didn't know the news that they would entrust to you, . . .

(b) Il solo incarico che non sapevi a chi avrebbero affidato è poi finito proprio a te.
**The only charge that you didn't know to whom they would entrust has been entrusted exactly to you.'

---

---
Using a bounding nodes version of Subjacency, Rizzi proposes that the distribution of island effects in Italian may be seen to follow from Subjacency if languages may be allowed to vary in the choice of bounding nodes for Subjacency. If the bounding nodes for Italian are S' and NP rather than S and NP as in English, then the array of judgments in (6) falls under Subjacency, as the reader may verify. We will refer to the parameter of variation identified by Rizzi as the Subjacency Parameter.2

The Subjacency Parameter may be invoked to account for the almost complete lack of extraction in some Slavic languages. In the case that a language virtually excludes unbounded dependencies into embedded clauses, it may be claimed that all three bounding nodes (S', S, and NP) are in effect. Of course, this leaves unexplained the possibility in Polish for example for unbounded dependencies into subjunctive clauses only (Lasnik and Saito, 1984: 279). Evidently, the parameter of variation must be made sensitive to mood as well as category, though the possibility for a purely categorial analysis has not been investigated in this area so far as we are aware.

There have been studies on Subjacency-style island effects in other languages as well. While we cannot survey the full range of these studies, we can point to some of the other variants in the cross-linguistic manifestation of Subjacency effects. One variant that follows on Chomsky's original formulation is discussed by Huang (1982).
Huang shows that in Chinese the Complex NP Constraint is respected in structures of extraction (relativization and topicalization), but the wh island Condition is not. The reason, Huang argues, is that wh questions in Chinese do not involve the overt displacement of the wh phrase from its thematically interpreted position. This leaves the COMP position of an interrogative clause in Chinese free to mediate between a dislocated higher constituent and the position of a gap within the clause in the successive cyclic fashion required by Subjacency. Huang concludes that the predictions of Subjacency are confirmed in Chinese.

The success of Subjacency in Chinese and Italian contrasts with its apparent failure in Scandinavian languages. Violations of the Complex NP Constraint have been documented in Swedish, Danish, and Norwegian. Violations of the wh island condition appear in Danish, Swedish, Norwegian and Icelandic. Examples from Danish (from Erteschik, 1973) appear below.

(7) (a) Det hus kender jeg en mand som har købt. (that house know I a man that has bought)

(b) Dette er bogen som jeg ikke ved hvem han har givet. (this is the book which I not know who he has given)

For a survey, discussion and analysis of relevant examples see Erteschik (1973), Allwood (1976), Taraldsen (1979), Maling and Zaenen (1982), Engdahl (1986) and the very useful collection of papers in Engdahl and Ejerhed (1982). To our knowledge, no fully adequate structural analysis of these phenomena has ever been proposed. Many researchers in this area have adopted a pragmatic account of the attested variation, adapting the proposals of Erteschik (1973) and Allwood (1976), as we discuss below. Others have adopted a mixed approach, seeking a syntactic solution for some constraints and a pragmatic solution for others, as for example Maling and Zaenen (1982), Engdahl (1986).

We note, however, that the bulk of examples in the literature on Scandinavian languages representing violations of the Complex NP Constraint involve relativization or questioning of the subject in the offending complex NP. Indeed, Allwood (1976: 9) and Engdahl (1986: 139) observe that relativization of some argument other than the subject leads to a decrease in acceptability in Swedish. Given the existence of general subject/non-subject asymmetries in many Scan-
dinavian languages in the distribution of pronouns and complementizers under dislocation of the subject, it is not unreasonable to wonder whether a structural analysis consistent with Subjacency is in fact possible for these cases, though this has never been attempted as far as we know. Consider also that example (7b) is not sufficient to show the inapplicability of the wh island constraint in Danish, given the possible variation Rizzi documents for Italian (see (6)). More complex cases are needed to determine the possible bounding nodes for these languages, whatever the source of variation may be due to. Again, to our knowledge, no comprehensive investigation has been reported that seeks to determine the range of variation in wh island violations in Scandinavian languages in terms of the Subjacency Parameter. But these remain outstanding questions, and the Scandinavian case against Subjacency remains essentially unchallenged by proponents of a syntactic approach to island effects.

Semantic and lexical variation

We now discuss two apparently non-structural factors that influence the potential for unbounded dependency relations, the definiteness and bridging effects. With regard to the former, a semantic factor that appears to influence the acceptability of violations of the Complex NP Constraint is the definiteness and/or specificity of the head of the offending relative clause. For instance, consider the following examples from English in (8)-(9) and Swedish in (10).

(8) (a)  ??He's someone who I've never met anyone that likes.
(b)  *He's someone who I've never met the person that likes.

(9) (a)  ??Robin's car, I can't find anyone to fix.
(b)  *Robin's car, I can't find the right mechanic to fix.

(10) (a)  Skaldjur är det många som inte tål.
  shellfish are there several who not stand
  'shellfish, there are several who can't eat'
(b)  ??Skaldjur känner jag mannen som inte tål.
  shellfish know I the man who not stand
  'shellfish I know the man who can't eat'
(Engdahl, 1986: 138–9)
The relatively low tolerance of definite/specific NP’s towards extraction even in cases where extraction is otherwise permitted is not well understood syntactically. This has paved the way for an account of the variation which makes appeal to the discourse function of definite/specific NP’s as well as to possible pragmatic factors, as for example in Allwood (1976), Andersson (1982), or Engdahl (1986). These authors also stress the contextual effects on variation, in that a given example presented out of context is much improved when embedded in an accommodating context. While a structural account might well be envisaged for the definiteness effect, the effect of context is not so readily addressed in structural terms.

Another non-structural factor that strongly influences the possibilities for relations of unbounded dependency is the bridging phenomenon of Erteschik (1973). She observes that while some verbs tolerate relations of unbounded dependence into their sentential complements, others do not. For instance, compare (11) and (12).³

(11) What did she think/say that Bill ate?

(12) *What did she whisper/exclaim/comment that Bill ate?

Erteschik observes that the bridging phenomenon is not restricted to verbal heads but holds also for adjectival heads.

(13) (a) What is it unlikely/certain that Bill ate?

(b) *What is it questionable/tragic/interesting that Bill ate?

To this, we add the observation that the bridging effect extends to nouns as well.

(14) (a) What were you appalled by the ??notion/*fact that she stole?

(b) What did you hear a ?rumor/*comment that she stole?

Fodor (this volume) considers still other examples, including prepositions and complementizers, that illustrate the very general cross-categorial nature of this bridging phenomenon. Anticipating this result, we conclude that any head may be a bridge, and we will correspondingly refer to the effect as the ‘bridge head’ effect. A bridge head is a head that permits extraction from its complement.
There are other examples that would also appear amenable to a bridging analysis, though they have not been previously considered to be. Consider the English (15) and the parallel cases in (16) from Scandinavian (Maling and Zaanen, 1982: 234).

(15) Who did you write/*destroy a book about?

(16) (a) Vem skrev/*förstörde Pelle en bok om? (Swedish)
(b) Hvem skrev/*ødela Pelle en bok om? (Norwegian)
(c) Umræv skrifaði/*eyðilagði Palli bök? (Icelandic)

‘Who did Pelle write/*destroy a book about?’

If write is understood to be a bridge verb in (15) and not destroy, then apparently every head that intervenes between the gap and the dislocated phrase must be a bridge head. This is illustrated also in the following example from Swedish (Engdahl, 1986: 139). (See also Allwood, 1976.)

(17) (a) Centerpartiet känner jag en man som röstar på.
the Center party know I a man who votes for

(b) ??Centerpartiet ringer jag upp en man som röstar på.
the Center party call I up a man who votes for

These examples confirm our earlier conclusion that extraction is possible only from the complement to a bridge head.

Syntactic theoreticians have generally taken the bridge head effect to be a necessary condition on extraction, independent of syntactic constraints such as Subjacency. Indeed, initial reflection suggests that bridging by itself is not sufficient to characterize the full range of island effects. It does not appear to capture the invariant adjunct, subject and wh island effects. On the other hand, bridging might accommodate even these invariant effects if unbounded dependencies are possible only into the non-wh subcategorized complement to a bridge head, taking neither subjects nor adjuncts to be subcategorized. In this view, all island effects of the type discussed here might be reduced to lexical properties, assuming the absence of unbounded dependencies to be the norm. This result would be consistent with Borer and Wexler’s (1987) proposal to limit parametric variation to the lexicon, but would appear
unable to account for a systemic parameter of the sort defended by Rizzi as discussed above. While appealing, therefore, it is not a view we will endorse here.

A final point in connection with bridging is Erteschik's (1973) proposal to characterize the class of bridge heads in terms of the pragmatic notion of 'dominance'. A constituent $\alpha$ is dominant if the speaker intends to direct the hearer's attention to the intension of $\alpha$. Then the island constraints follow from the general principle that extraction is possible only from a dominant constituent. In these terms, non-bridge heads attract dominance away from their complements, so blocking extraction; bridge heads don't. Erteschik-Shir's approach is adapted in much other work, such as Allwood (1976) and Kuno (1987), and has been very influential in studies of island effects in the Scandinavian languages, as noted above.

**LF island effects**

The foregoing discussion has been concerned exclusively with surface (overt) island effects. There has been considerable discussion, however, of whether there is any evidence of non-overt (LF) island effects. In GB, LF is posited to be the level of representation at which all grammatical aspects of semantic interpretation are displayed. In particular, LF is a semantically disambiguated syntactic representation that encodes sentence level scope distinctions for operator/variable configurations that arise in natural language. Since May (1977), it has been widely assumed that LF derives from S-structure by application of Move $\alpha$, so that LF, like S-structure, has a phrase structure representation that is constrained by structurally based principles of association, such as the Empty Category Principle (ECP). (See Chomsky, 1981.) Although some earlier linguistic discussions of quantification proposed that quantifier scope is constrained by Subjacency (Rodman, 1976; May, 1977), it has been standardly assumed that Subjacency does not constrain LF applications of Move $\alpha$. This conclusion is motivated by consideration of examples such as (18), in which, assuming that the wh phrase what in situ at S-structure undergoes movement at LF to the highest occupied COMP position, Subjacency is not respected in this non-overt movement.
(18) (a) Who remembers who bought what?
(b) Who recognizes the person that bought what?
(c) Who left after Bill bought what?

In contrast, S-structure movement of what does obey Subjacency, as in
the otherwise parallel examples in (19).

(19) (a) *What does Bill remember who bought?
(b) *What does Bill recognize the person that bought?
(c) *What did John leave after Bill bought?

That non-overt (LF) movement does not respect Subjacency is indi-
cated also by the observation that the equivalents of examples (19) are
grammatical in languages without overt movement in wh questions,
such as Chinese and Japanese (see Huang, 1982; Lasnik and Saito,
1984).

Despite the contrast between (18) and (19), some have sought to
preserve the claim that Subjacency does hold of LF as well as S-
structure movement. Indeed, if such a claim can be maintained, it lends
support to the hypothesis that LF is a phrase structure representation
derived by application of Move a. One possibility is to suppose, with
Choe (1985) and Nishigauchi (1986), that LF derivations may access a
‘pied piping’ convention that gives rise to representations that respect
Subjacency. In this view, the LF movement of a wh in situ may pied
pipe the entire island containing the wh phrase; this movement would
be (trivially) consistent with Subjacency. It remains to be seen whether
the pied piping mechanism can be elaborated in a manner that does not
re-introduce a distinction between LF and S-structure movement, since
no similar pied piping operation is exhibited in S-structure movement
(see Lasnik and Saito, 1989 for discussion). An alternative approach is
that of Pesetsky (1987), who proposes that only some wh phrases
(those that are not D-linked (discourse linked)) undergo movement at
LF and only these phrases will display non-overt Subjacency effects.
Unfortunately, Pesetsky makes appeal to the same LF pied piping
mechanism to accommodate the contrast between (18) and (19), with
the attendant potential problems of the sort alluded to earlier. Without
a more complete and independently motivated account of the LF pied
piping mechanism, we conclude, given the contrasts with examples (19),
that Subjacency does not constrain the LF movement of such phrases as
*what in examples such as (18).
Finally, it must be mentioned that non-overt (LF) island effects are
evident in the interpretation of adjunct wh-in-situ as in the examples
below. 10

(20) (a) *Who remembers whether [Bill fixed the car how]?
(b) *Who knows the person that [fixed the car how]?
(c) *Who left after [Bill fixed the car how]?
These examples contrast with (18) too, in the non-overt movement of
*how in contrast to *what. Evidently, the LF movement of adjunct wh-in-
situ displays island effects of a sort very similar, if not identical, to
surface Subjacency effects. The standard treatment of such examples is
in terms of the ECP (see Huang, 1982; Lasnik and Saito, 1984; Chomsky, 1986). 11
This is again a strong argument in favour of the
hypothesis that LF is given a configurational representation, to the
extent that the ECP can be argued to have a configurational definition.

Conclusion

This brief discussion is intended to give a flavour of what is meant by
island effects and how these effects have been addressed in the theo-
retical literature. While there may not be full agreement concerning the
proper characterization of these effects, cataloguing them has provided
us with clear evidence of their existence and a substantial body of
knowledge concerning their properties. As the papers in this volume
will show, and as we will discuss briefly in the next section, the range of
our knowledge and theoretical understanding of island constraints has
given rise to limited but expanding investigation into questions regarding
their status in language acquisition and their deployment in pro-
cessing.