

## 6 On the Existence (and Distribution) of Sentential Subjects

William D. Davies and Stanley Dubinsky

### 6.1 Introduction

This chapter principally concerns the status and distribution of sentential subjects, as in (1), and to a lesser extent, the relation of (1) to the extraposition construction shown in (2).<sup>1</sup>

(1) That John left early disappointed us.

(2) It disappointed us that John left early.

We will show that some long-standing assumptions about the distribution of nonnominal subjects (going back to Koster 1978 and Stowell 1981) are incorrect, and that the distribution of these subjects is broader than previous analyses have acknowledged. We suggest that previous conclusions about this class of data were driven by facts whose explanation likely lies outside the realm of syntactic structure. While much of the debate about the status of sentential subjects occurred twenty to thirty years ago, the issue is still quite relevant, inasmuch as a number of current syntax textbooks still present what turn out to be questionable analyses of these structures (see Culicover 1997; Haegeman and Guéron 1999; Lasnik 1999, 1995; and Radford 1997). This is, to some degree, surprising, especially given the fact that the Case-theoretic explanations for the facts have declined in importance with the paradigm shift from Principles and Parameters (i.e., GB) to the Minimalist Program. What follows is an attempt to lay out previously established facts as well as some new ones, to call attention to long-standing misconceptions about sentential subjects, to distill the correct generalization for subject positions in English, and to point to potential sources of new data on the topic.

### 6.2 The Background

Rosenbaum 1967 presented one of the earliest accounts of (1) and (2) in the transformational literature. There, all sentential subjects are generated as complements to

a subject pronoun *it* such that (1) and (2) have the same underlying structure, given in (3).

(3) [<sub>NP</sub> it [<sub>S</sub> that John left early]] disappointed us

In (1), according to Rosenbaum, the pronoun deletes, leaving the S in subject position dominated by a headless NP projection. In (2), the S extraposes to the right, leaving the pronoun *it* stranded in subject position (Delahunty 1981 and Iwakura 1978 propose similar analyses). Emonds 1970 (and 1972) noted several ways in which sentential subjects do not behave like ordinary NP subjects, and pointed out a number of problems with Rosenbaum's extraposition approach. Among other things, he notes that extraposition is obligatory (i) in subordinate clauses (4)–(5) and (ii) in the context of subject-aux inversion regardless of whether there is a questioned NP (6)–(7), and that (iii) clausal subjects are ungrammatical in the presence of other topicalized elements (8).

- (4) a. \*I think that that John left early disappointed them.  
b. I think that it disappointed them that John left early.
- (5) a. \*Although that the house is empty may depress you, it pleases me. (Koster 1978, (2b))  
b. Although it may depress you that the house is empty, it pleases me.
- (6) a. \*Did that John showed up please you? (Koster 1978, (3a))  
b. Did it please you that John showed up?
- (7) a. \*Who did that John left early disappoint?  
b. Who did it disappoint that John left early?
- (8) a. \*Such things, that he reads so much doesn't prove. (Koster 1978, (5b))  
b. Such things, it doesn't prove. (Koster 1978, (4b))

Concerned as he was with the coherence of the phrase structure component of the grammar and the capacity of overgeneration presented by rewrite rules of the form  $XP \rightarrow YP$  (e.g.,  $NP \rightarrow S$ ), Emonds proposed that sentential subjects are base-generated in an adjunct position and then "intraposed" into subject position by a rule of Subject Replacement (Brame 1976 proposes a similar analysis).

Emonds's 1976 analysis reverses the direction of movement he proposed earlier. Reacting perhaps to Postal's (1974) assertion that the intraposition analysis does not properly represent the semantics of the sentence at D-structure, Emonds replaces intraposition with extraposition and claims that sentential subjects such as in (1) are topicalized at S-structure. Observing as he did earlier that there is a correlation in the distribution of topics and sentential subjects, he states that "nongerund clauses will appear only in extraposition and topicalization NP positions" (p. 127). That is, where topics cannot appear, sentential subjects also cannot appear. His proposal for the D-structure of (1) and (2) is thus quite similar to that of Rosenbaum, except that

the former is claimed to have a null pronoun in the base. For Emonds, sentential subjects (and complements) are generated under an NP node with an empty N head, as shown in (9).

- (9) a. [<sub>NP</sub>  $\emptyset$  [<sub>S</sub> that John left early]] disappointed us  
 b. I hate [<sub>NP</sub>  $\emptyset$  [<sub>S</sub> that they left early]]

These sentential arguments must be moved out of their base position in order for the sentence to be grammatical. Extraposition (i.e., rightward movement) results in an expletive *it* in the argument position, as in (10). Topicalization, as in (11), results in erasure of the empty head.

- (10) a. [<sub>NP</sub> it] disappointed them [<sub>S</sub> that John left early]  
 b. I hate [<sub>NP</sub> it] [<sub>S</sub> that they left early]
- (11) a. [<sub>S</sub> that John left early] [<sub>NP</sub> ] disappointed them  
 b. [<sub>S</sub> that they left early] I hate [<sub>NP</sub> ]

Movement of the sentential subject out of its D-structure position is, in either case, obligatory and the ungrammaticality of (7a), for example, is due in Emonds's account to the sentential subject remaining in the subject position over which the auxiliary must move.

Koster 1978, building on Emonds's observations, proposes that sentential subject clauses such as in (1) occupy a "satellite" position outside the main clause, and are therefore outside the domain of the subject-aux inversion rule at all levels of structure. Koster attempts to derive this from PS rules, stating that subject position is defined as an NP position and that his proposal to have sentential subjects occupy a "satellite" position from the outset does away with dubious (exocentric) rules such as NP  $\rightarrow$  S. Under his analysis, (5a) is ungrammatical because the "satellite" clause cannot be subordinate to the complementizer *although*. In (6a) and (7a), the auxiliary *did* has moved across the "satellite" position rather than the subject position as in (6b) and (7b). In (8a), *such things* and *that he reads so much* compete for the same (topicalized) position.

Stowell (1981) also adopts the position that exocentric phrase structure (PS) rules are automatically ruled out by any theory that incorporates principles of X-bar syntax, and goes on to propose a theory-internal principle (in the Principles and Parameters/GB framework) to motivate the analysis proposed in Koster 1978. While the prohibition of exocentric PS rules can, by itself, preclude a sentential subject immediately dominated by NP, it does not explain why the canonical subject position appears to exclude Ss themselves.

In keeping with the GB position of explaining the distribution of arguments (partly) in terms of Case assignment principles, Stowell (1981) proposes the Case-Resistance Principle (CRP), given in (12).

(12) *Case-Resistance Principle (CRP)*

Case may not be assigned to a category bearing a Case-assigning feature.

According to Stowell, because [+Tense] is a Case-assigning feature that is responsible for nominative Case, tensed sentential subjects may not themselves be assigned Case. The CRP thus predicts that the tensed subject in (13a) must occupy some position other than the canonical subject position, while the infinitival subject in (13b) may occur in this position.

- (13) a. [that John left early] [e disappointed us]  
 b. [[for John to leave early] would disappoint us]

The table in (14) summarizes the several analyses of sentential subjects (SSs) presented here.

(14)

	Rosenbaum 1967	Emonds 1970	Emonds 1976	Koster 1978
SSs are D-structure subjects	yes	no	yes	no
SSs are S-structure subjects	yes	yes	no	no

### 6.3 It Is Not a Case of Case

While sporadic challenges for particular languages have appeared from time to time (e.g., Chung 1991), Stowell's CRP has remained conventional wisdom, and is adopted in some recent texts (Culicover 1997, 51–52; Radford 1997, 297). Presumably, the CRP should also exclude VPs and PPs from Case-bearing positions, since both V and P are categories that bear an accusative Case-assigning feature. We would thus expect VPs, PPs, and tensed CPs all to avoid the Case-marked subject position, in contrast with NPs and APs.<sup>2</sup> Thus, according to Stowell, neither the clausal subject in (1) nor the PP subjects in (15) occupy the same position as the NP subject in (2).

- (15) a. On the porch would be a good place to leave the toys.  
 b. On the porch reminds me of a good place to leave the toys.

A casual examination of the distribution of PP subjects shows it to be patently false. In contrast with (6a), (16) is perfect.<sup>3</sup>

- (16) a. Would on the porch be a good place to leave the toys?  
 b. Does on the porch remind you of a good place to leave the toys?

Additionally, VPs and IPs, both of which are [–N] categories, can also occur in these constructions as in (17) and (18).<sup>4</sup>

- (17) a. Why does eat pizza seem to be all he wants to do?  
       b. Does wait and see sound like what they were willing to do?
- (18) a. Would to leave early really reflect poorly on us all?  
       b. When does to let sleeping dogs lie strike you as the best course of action?

Thus, Stowell's proposed CRP cannot be maintained. Further, a careful examination of clausal subjects indicates that the distribution of these too is wider than what Emonds and Koster have claimed. The examples in (19) are not nearly as bad as (6a) and (7a).

- (19) a. To whom is that pigs can fly most surprising?  
       b. Is that I am done with this homework really amazing? (Naturally occurring datum)

Crosslinguistic considerations also call into question any appeal to Case resistance as an adequate explanation of the English data. In his careful study of sentential arguments in a wide variety of languages, Dryer (1980) demonstrates the crosslinguistic tendency for both sentential subjects and sentential objects to occur in clause-final position and a secondary tendency (especially in SOV languages) for these elements to occur in clause-initial position. On the face of it, then, it might appear that such data are consistent with a Case-resistance explanation. However, these are merely robust tendencies. Dryer provides copious evidence that clause-internal sentential complements are dispreferred but definitely not disallowed. For example, in SOV languages, sentential objects frequently occur in the canonical object position (that is, sentence internally, following the subject and preceding the verb). While there are some SOV languages in which sentential objects obligatorily follow the verb—Dryer cites Persian and Turkish as examples of these—many more SOV languages (e.g., Choctaw and Japanese) do allow clause-internal sentential objects. Under the simplest set of assumptions, these sentential objects occur in canonical object position, a Case position. Given that very few languages absolutely require their sentential objects to occur in clause-final (or clause-initial) position, crosslinguistic data provide yet another challenge to Case resistance.

Chamorro provides empirical evidence that sentential arguments can occur in Case positions. Chung (1991) argues explicitly that CP subjects and objects in Chamorro take case and that the Case Resistance Principle cannot be operable in this language; she limits her argument to Chamorro. Evidence for this comes from a system of what she calls “*wh*-agreement,” in which the morphology on a verb reflects the grammatical function of a *wh*-phrase that has moved across it. In (20), the subject is questioned and the *um* affix marks *wh*-agreement on the root *istotba*. In (21), the direct object

is questioned, which results in a nominalization of the clause (indicated by the *ña* possessive suffix); here, the affix *in* marks the *wh*-agreement.

- (20) Hafa<sub>i</sub> *umistotba* hao *t<sub>i</sub>*? (Chung 1991, (15b))  
 what Infl.disturb.Prog you  
 [Wh.nomin]  
 ‘What is disturbing you?’
- (21) Hayi<sub>i</sub> *chiniku-ña* si Dolores *t<sub>i</sub>*? (Chung 1991, (17b))  
 what kiss-Agr the D  
 [Wh.obj]  
 ‘Who did Dolores kiss?’

In simplex clauses, Chung formulates this in terms of the verb agreeing in Case with a clausemate that is a *wh*-trace.

In long-distance questions, the verb of the clause in which the extracted element originates takes agreement reflective of the grammatical function of the extracted element, but higher verbs take agreement that reflects the grammatical function of the clause from which the element is extracted. Thus, in (22), the embedded clause is nominalized because its direct object has been extracted. However, the higher verb takes the *um* affix: this is because the clause from which the element is extracted is the subject.

- (22) Hafa<sub>i</sub> *umistotba* hao [ni malago’-ña i lahi-mu *t<sub>i</sub>*?  
 what Infl.disturb you Comp want-Agr the son-your  
 [Wh.nomin] [Wh.obj]  
 ‘What does it disturb you that your son wants?’ (Chung 1982, (45a))

This contrasts with the situation when a subject has been extracted from a sentential subject, as in (23), where the verbs of both clauses take the *um* affix.

- (23) Hayi na lahi<sub>i</sub> *umistotba* si Jose [ni bumisita *t<sub>i</sub>*  
 who L boy Infl.disturb the J Comp Infl.visit  
 [Wh.nomin] [Wh.nomin]  
 haga-ña]? (Chung 1991, (23a))  
 the daughter-his  
 ‘What boy does it disturb Jose visited his daughter?’

As Chung argues, inasmuch as *wh*-Agreement is a manifestation of Case agreement and CPs trigger *wh*-Agreement, we must conclude that CPs in argument positions bear Case.

Levin and Massam (1986) also argue for the assignment of Case to CPs in Niuean. They note that some Niuean verbs with sentential complements take absolutive subjects while others take ergative subjects. Among the verbs with ergative subjects are

*iola* ‘know’, *manatu* ‘think, wonder’, *talahuaua* ‘say’, and others.<sup>5</sup> In (24), the matrix subject *kau kaihā* ‘thieves’ takes the ergative case particle *he*.

- (24) Ne kitia **he kau kaihā** kua mate tuai e molī he fale.  
 PST see ERG group thief PERF die PERF ABS lamp in house  
 ‘The thieves saw that the lamp in the house had gone out.’

Levin and Massam explicitly argue against analyzing the clausal complements as NPs on the basis of extraction facts. They also propose that sentential subjects of predicates such as *hangahanga* ‘appear’, *lata* ‘be right’, *mitaki* ‘good’, and others occur in subject position, accounting for the impossibility of Raising-to-Subject with these predicates, as opposed to predicates such as *maeke* ‘be possible’ and others that allow Raising. Levin and Massam conclude that these facts from Niuean provide evidence against Safir’s (1982) hypothesis that sentential complements cannot be assigned Case.

#### 6.4 Alternative Explanations

Two of the arguments brought forth by Emonds and Koster against the existence of sentential subjects include structures that are crosslinguistically marked. These are the arguments based on the degraded nature of sentential subjects that occur inverted with auxiliaries in questions (6a) and (7a), and sentential subjects that follow a topicalized NP (8a).<sup>6</sup>

- (6) a. \*Did that John showed up please you? (Koster 1978, (3a))  
 (7) a. \*Who did that John left early disappoint?  
 (8) a. \*Such things, that he reads so much doesn’t prove. (Koster 1978, (5b))

In each of these structures, the sentential subject occurs internal to the clause—that is, neither in clause-initial nor clause-final position. Recall that Dryer (1980) demonstrates that the preferred position of sentential arguments (both subjects and complements) is clause-final position. He further demonstrates that, despite the fact that clause-internal sentential arguments can occur (especially objects in SOV languages), this position is greatly dispreferred crosslinguistically, with some languages (e.g., Persian and Turkish) seemingly banning them entirely. This fact thus adds a confounding factor to the data marshaled as evidence against sentential subjects (and sentential arguments in Case-marked positions as a whole).<sup>7</sup> But as demonstrated above, the evidence against the existence of sentential subjects is questionable.

The distribution of nonnominal subjects clearly precludes the Case-theoretic explanation that has been advanced for it, and leads to the conclusion that other, nonsyntactic, factors have significantly clouded the picture. Having shown that sentential (or clausal) subjects do exist, we will now consider why they are kind of weird and

so often avoided, and examine some of the nonsyntactic explanations for their distribution. Grosu and Thompson 1977 (G&T) reached similar conclusions with respect to then-current theoretical assumptions of syntactic theory, asserting that “neither Generative Semantics nor the Extended Standard Theory can provide an interesting (or even adequate) account of [the distribution of NP clauses]” (p. 139). They went on to suggest that the problems associated with this distribution “can be illuminated only by appealing to certain principles of language processing . . . even though no sufficiently detailed or precise theory of language processing is available at the moment” (p. 139). While Dryer (1980) showed the G&T solution to be unsuccessful overall, it is clear that they were on the right track. Some recent work appears to provide better explanations for the distribution of sentential subjects. In particular, we think that their distribution may be better explained in terms of prosody and phrasal weight (Erdmann 1988) or processing factors, such as the “integration cost component” of Gibson’s (1998) Syntactic Prediction Locality Theory (SPLT). It is also evident that at least some of the data types presented by Koster were “red herrings,” whose acceptability was severely degraded by “garden-path” processing difficulties.

#### 6.4.1 Weight

As Delahunty (1983, 383) suggests, the relative unacceptability of many of Koster’s examples may simply be a consequence of “the relative ‘weights’ and perhaps prosody” of the relevant constituents. Compare the unacceptable subject-auxiliary inversion example (25a) from Koster with the relatively acceptable structural analog (25b) from Delahunty.

- (25) a. \*Did that John showed up please you? (Koster 1978, (3a))  
       b. To what extent did that Fred failed to show up anger those of his devoted fans who had waited by the stage door since dawn of the previous day?  
           (Delahunty 1983, (11))

In Koster’s example the sentential subject is twice the length (in syllables) of the matrix predicate, while in Delahunty’s example, the six-word sentential subject is followed by an eighteen-word matrix predicate.

Erdmann (1988) confirms this. He extracted from linguistic corpora data showing the relative distribution of CP and IP (infinitival) in situ versus extraposed subjects. In these corpora, he found that the frequency of extraposition is extremely high with adjectival main-clause predicates (e.g., *is surprising*). In his data, 92.04 percent of CP subjects and 95.07 percent of IP subjects extraposed. This is not surprising given the tendencies that Dryer and G&T report. However, when the weight of the predicate VP was considered, it was found to play a significant role in determining when extraposition was avoided. With CP subjects in clause-initial subject position, the adjectival predicate was “light” 56 percent of the time and was “heavy” 44 percent of



the time. In contrast, when CP subjects were extraposed, the adjectival predicate was “light” 88 percent of the time and “heavy” 12 percent. Figures are comparable for IP subjects (though slightly less dramatic). What this suggests is that extraposition is decreasingly motivated as the VP over which it moves becomes heavier. In results garnered from a corpus-based study, Wasow (1997) reports that the effect of phrasal weight in inducing heavy NP shift is linked to the relative weight of the two constituents, the second phrase being approximately three times heavier than the first (by any relevant measure).

It is reasonable therefore to suspect that the unacceptability of Koster’s subject-auxiliary inversion example in (25a) has more to do with the sentential subject being twice as long as the following predicate (four words to two words), than with any violation of grammatical constraints.

#### 6.4.2 Memory and Parsing Impediments

Section 6.4.1 discussed the role of phrasal weight in determining the placement of these phrases primarily from the perspective of production. There is also some indication that both memory limitations and parsing ambiguity play a role, from a processing perspective, in rendering unacceptable center-embedded sentential subjects.

In both G&T 1977 and Dryer 1980, the dispreference for clause-internal CPs (such as those produced in English by CP-Aux inversion) is explicitly tied to Fodor, Bever, and Garrett’s (1974) explanation of the unacceptability of multiple center-embedding. Dryer’s comment on this is as follows: “If we assume that clauses are the fundamental units of sentence processing, and that material is emptied from short term memory at clause boundaries, clause-internal sentential NPs will interrupt the processing of the main clause” (p. 161).

Hawkins (1994) proposes his Early Immediate Constituent principle to account for word-order tendencies observed in both production and perception studies. Based on the reported data, he claims that “the human parser prefers linear orders that maximize the IC-to-non-IC ratios of constituent recognition domains” (p. 77). Where extraposed and nonextraposed sentential subjects are concerned, this means that extraposed sentential subjects are preferable and easier to process than nonextraposed sentential subjects.

- (26) a. [s [it] [pleased you]] [that John showed up]  
       b. [s [that John showed up] [pleased you]]

In (26a), the three-word string *it pleased you* forms an S dominating the immediate constituents *it* and *pleased you*. In (26b), the first three words *that John showed* do not form an IC.

Eye-tracking experiments conducted by Frazier and Rayner (1988) indicate that sentential subjects were harder to process than their extraposed sentential subject

analogs. For two experiments—one in which test sentences were given without a context and one in which they were contextualized—they report that reading times for the stimuli with sentential subjects were significantly longer than those with extraposed sentential subjects. This finding again appears to point to the importance of ‘weight’.

However, some preliminary data, collected by Widmann (2005), suggests that the relative weight of sentential subjects and their predicates may play a less critical role in speakers’ acceptability judgments in certain contexts than the absolute length of the center-embedded sentential subject. In two experiments, each presenting a different center-embedding context, speakers judged most acceptable center-embedded subjects that were neither too long nor too short.

Widmann’s experiments tested subject-auxiliary inversion contexts, as in (27a), and subordinate-clause contexts, as in (27b).

- (27) a. Does that the parent wanted to come home cause any problem for the older children?  
       b. Although that the parent wanted to come home caused problems for the older children, it was not a terrible inconvenience.

In manipulating the length of the embedded sentential subject and the following predicate (four versus ten words; ten versus four words; seven versus seven words), experimental results indicated that speakers judged significantly more acceptable, embedded sentential subjects that were neither overly long nor overly short (i.e., seven-word sentential subjects were preferred over both four- and ten-word sentential subjects).

While further experiments are needed to confirm these results, an initial hypothesis put forward is that long sentential subjects (i.e., ten words) may tax short-term memory resources in the manner described by Dryer (above). Conversely, overly short center-embedded sentential subjects (i.e., four words) do not permit the parser sufficient time to recover from the initial misanalysis triggered by the first two words of (27a) and (27b). In reading *does that...* or *although that...*, it is hypothesized that the parser initially posits the following structure:

- (28) a. [<sub>CP</sub> does [<sub>IP</sub> [<sub>NP</sub> that...  
       b. [<sub>CP</sub> although [<sub>IP</sub> [<sub>NP</sub> that...]

The next two disambiguating words, *the parent*, force the parser to backtrack and reanalyze the structure as containing a center-embedded CP:

- (29) a. [<sub>CP</sub> does [<sub>CP</sub> that [<sub>IP</sub> [<sub>NP</sub> the parent...  
       b. [<sub>CP</sub> although [<sub>CP</sub> that [<sub>IP</sub> [<sub>NP</sub> the parent...]

In the four-word sentential subject context (e.g., *that the parent came*) there is not sufficient time following the disambiguating region for the parser to recover, before

it must begin to parse the rest of the sentence. Hence, the seven-word sentential subjects, which permitted such recovery but did not overly tax memory, were judged significantly more acceptable. These results, if confirmed by further experimentation, are important in that they indicate that (1) more is at play than just the relative weight of constituents, and (2) parsing factors affecting acceptability judgments might be distinct from factors determining preferences in production.

Here again, we are led to the conclusion that the unacceptability of sentences used by Koster as evidence against the “existence of sentential subjects” is more likely the result of psycholinguistic factors than of grammatical principles.

#### 6.4.3 Integration Costs

Another processing factor that can contribute to variation in the acceptability of clause-internal CPs is the integration of new referents into the discourse. As Gibson (1998, 16) points out, “Elements which cause . . . integration cost . . . are words introducing new discourse referents. . . . Doubly nested RC structures are easier to process when a first- or second-person pronoun (an indexical pronoun) is in the subject position of the most embedded clause, as compared with similar structures in which a proper name, a full NP or a pronoun with no referent is in the subject position of the most embedded.” This is presumably a factor in rendering the relative clause in (30a) less acceptable than the one in (30b).

- (30) a. ?The lawyer the banker irritated sued him.  
b. The lawyer I irritated sued me.

Similar (though weak) effects show up in sentences containing clause-internal CPs, where (31a) is somewhat worse than (31b).

- (31) a. Did that players slipped on the ice truly worry the coach?  
b. Did that I finished on time really impress the teacher?

Confirmation of such differences would require experimental data.

#### 6.5 Repercussions for Syntactic Theory

Thus far we have presented reasons to eschew a syntactic explanation for the distribution of sentential subjects, providing evidence against previous syntactic analyses of such and pointing to other, nonsyntactic factors that govern their distribution. One might ask at this point (if one has not already), why we (as syntacticians) are so exercised about this issue. There are two answers. First, we think that ignoring the sort of factors discussed in section 6.4 is bad for the pursuit of syntactic analysis generally, inevitably leading to very weak (or very wrong) arguments for an analysis. Second, the “received wisdom” of the CRP in syntactic circles has stood in the way

of improved accounts of subject properties in English (and other languages). We will take up these issues in turn.

### 6.5.1 Garden-Path Obstacles to Linguistic Inquiry

Not only do prosodic and processing factors provide a reasonable alternative to the Case-theoretic explanation of the distribution of sentential subjects, but they also can help to illuminate the extreme weakness of some of the syntactic arguments that have been put forward in favor of Case-theoretic explanations. For example, one of Koster's (1978) arguments against sentential subjects involved showing that they could not appear with a preceding topicalized phrase. Koster's illustration of this is given in (32).

(32) \*Such things, that he reads so much doesn't prove. (Koster 1978, (5b))

This is an undeniably unacceptable sentence. But it appears that the "ungrammaticality" of (32) is not a function of the incompatibility of sentential subjects and topics. Example (33a) is almost equally unacceptable, differing from (32) only in having a head noun dominating the sentential subject, and (33b) should be grammatical if the unacceptability of (32) were merely due to extraposition not having applied.

- (33) a. \*Such things, the fact that he reads so much doesn't prove.  
 b. \*Such things, it doesn't prove that he reads so much.

Since (33a) contains a complex NP subject rather than a sentential subject, Koster's analysis provides no account for its unacceptability, and since (33b) has an extraposed sentential subject, it would appear that the unacceptability of (32) is not due to the subject clause occupying any "satellite" position. We would instead argue that the unacceptability of (33a) and (33b) relates to factors that make it difficult to parse. The fact is that outside of some very well-rehearsed examples such as "Beans, I like," topicalized structures in English are usually slightly less acceptable than non-topicalized structures, because they are highly stylized sentences. (In fact, in our experience it is not unusual to find that students in an introductory syntax class initially judge topicalization structures as unacceptable.) So consider the pair in (34).

- (34) a. Few people debate such things.  
 b. Such things, few people debate.

The relative unacceptability of (33a) compared with (34b) might be due to the heavy NP in subject position in combination with the relatively marked topicalization order. For those who find (32) even less acceptable than (33a), we might add the problem of a garden-pathing. A common noun (such as *things*) with a following *that*-clause is generally parsed as a relative clause. The problem one confronts in parsing (32) is the fact that this initial parse will prove unsuccessful, and the processor will have to backtrack and revise the initial analysis (if that is even possible,

given the other confounds). In addition, the sentence in (32), as well as that in (33a), suffers from the problem of weight. The predicate *doesn't prove* is very light (in Erdmann's terms) when compared with the subject *(the fact) that he reads so much*. When these confounding factors are removed, topicalization combined with a sentential subject improves in acceptability, as seen in example (35). Here the relative-weight confound has been ameliorated, in that the predicate *bothers to no end* is slightly heavier than the subject *that John's a fool*.

- (35) a. Ted, that John's a fool bothers to no end, not Horatio.  
 b. The instructor, that John's a fool bothers to no end, not the TA.

While neither sentence in (35) is particularly acceptable, we find (35b) worse than (35a), and would conjecture that this may be due to the garden-path tendency to initially analyze the string *the instructor that* as an NP followed by a relative clause. Sentence (35a), in contrast, does not suffer from this effect.

### 6.5.2 Back Toward a Better Analysis of Sentential Subjects

The arguments of Koster 1978 and Stowell 1981 are inconsistent with an analysis of subjects such as was proposed in Lees 1960 and Rosenbaum 1967 and was adopted in many subsequent works (e.g., Chomsky 1973; Emonds 1976; and Delahunty 1983, to name a few). Both Lees and Rosenbaum (and those who followed) analyzed sentential subjects as having an NP node immediately dominating an S.<sup>8</sup> In recent papers (Davies and Dubinsky 1999, 2001), we have presented new evidence suggesting that Lees and Rosenbaum were closer to the mark than has since been thought, at least with respect to the syntax of sentential subjects. The arguments presented in our papers support the notion that a range of languages, including English, have a syntactic requirement that all subjects be DPs—importantly, this induces a DP node dominating non-NP subjects. Under this analysis, sentential subjects have the structure shown in (36).

- (36) [DP  $\emptyset$  [CP that Shelby lost it]] is quite apparent

Four types of arguments support this analysis. First, non-NP subjects undergo obligatory raising, a fact that Delahunty (1983) also cites in arguing for the NP-hood of sentential subjects. In (37) *that Shelby lost it* has raised to be matrix subject, something that would be quite anomalous if the clause were “resistant” to Case.

- (37) [CP that Shelby lost it]<sub>i</sub> appears [t<sub>i</sub> to be true]

Second, sentential subjects can trigger subject agreement on the verb, as observed by McCloskey (1991, 564) and shown in (38).

- (38) [CP [CP that the march should go ahead] and [CP that it should be canceled]]  
 have been argued by the same people at different times.

Third, a manner of licensing the quantificational adverb *equally* is via a plural NP or an NP with a mass noun as head. Conjoined CPs in subject position can license *equally*, as (39a) shows, while nonsubject CPs cannot (39b).

- (39) a. That he'll resign and that he'll stay in office seem at this point equally possible. (McCloskey 1991, 564)  
 b. Dale thought that Dana left and that Terry wouldn't come (\*equally).

Finally, sentential subjects can host emphatic reflexives, as in (40a), while sentential complements cannot (40b).

- (40) a. That there were twenty-five miles to go was itself enough to discourage Edwin.  
 b. Edwin hoped that there were less than twenty-five miles to go (\*itself).

However, this is not true only of sentential subjects: *all* non-NP subjects exhibit these properties. PP subjects and AP subjects also undergo obligatory raising (41), can trigger verb agreement (42), and license *equally* (43). Additionally, PPs can host emphatic reflexives, but only when they are subjects (44a).<sup>9</sup>

- (41) a. [<sub>PP</sub> under the bed]<sub>i</sub> appears [<sub>t<sub>i</sub></sub> to be a good place to hide]  
 b. [<sub>very tall</sub>]<sub>i</sub> appears [<sub>t<sub>i</sub></sub> to be just how he likes his bodyguards]  
 (42) a. [<sub>PP</sub> [<sub>PP</sub> under the bed] and [<sub>PP</sub> in the fireplace]] are not the best (combination of) places to leave your toys (Levine 1989, 1015)  
 b. [<sub>AP</sub> [<sub>AP</sub> very brawny] and [<sub>AP</sub> very studious]] are what Cindy aspires to be  
 (43) a. Under the bed and in the closet equally remind me of that game of hide-and-seek we played.  
 b. Very tall and quaintly studious equally bring to mind my sixth-grade science teacher.  
 (44) a. Under the bed and in the closet are themselves reasonable places to stash the cash.  
 b. We stashed the cash under the bed and in the closet (\*themselves).

As these data demonstrate, all non-NPs exhibit NP-like properties just when they are in subject position. This fact has been overlooked both in analyses that take sentential subjects to be subjects as well as in analyses that claim that they are not so. Thus the lexical category of a subject makes no difference to its distribution in English because all subjects are contained in DP projections (although, as argued in Davies and Dubinsky 2001, this is not the case for some other languages, such as Bulgarian).

Therefore, ironically perhaps, at one level we must concur with Koster's claim that "sentential subjects don't exist" in English—but not for the reasons Koster

advanced. Rather, sentential subjects do not exist in English because English admits only DP subjects. This fact lends support to the notion that the explanation for the idiosyncratic behavior of English sentential subjects lies outside the domain of syntactic structure, most likely in the domain of language processing

### Notes

For helpful comments, questions, and feedback on the content of this chapter, we thank Frederick Newmeyer, Maria Polinsky, Roumyana Slabakova, as well as audiences at the LSA Annual Meeting 2001, Northeastern Linguistics Society 2001, Southeastern Conference on Linguistics 2000, Eastern States Conference on Linguistics 1999, Chicago Linguistics Society 1998, Western Conference on Linguistics 1999, the University of Iowa, and the University of South Carolina, where portions of this chapter were presented.

1. The following abbreviations are used in the glosses, beginning with example (20): Wh.nomin = nominative *Wh*; Wh.obj = objective *Wh*; Infl = inflection; Prog = progressive; Agr = agreement; Comp = complementizer; PST = past, ERG = ergative, PERF = perfective, ABS = absolutive.

2. Safir (1983) suggests that PP and AP subjects are “honorary NPs” and thus raise, as in example (i), but does not distinguish between them in terms of their capacity to either assign or receive Case.

- (i) a. Angry/unwanted is a terrible way to feel.
- b. To the moon seems to be a good place to go.

3. Stowell (1981) notes that PPs may be subjects in copular constructions, and limits their occurrence to that. His data (p. 268):

- (i) a. [under the chair] is a nice place for the cat to sleep (Stowell 1981, (27))
- b. is [under the chair] a nice place to for the cat to sleep?
- (ii) a. \*[under the chair] pleased the cat (Stowell 1981, (28))
- b. \*did [under the chair] please the cat?
- (cf. Did it please the cat under the chair?)

He speculates that perhaps the copula might have a special property that allows nominative case to “be absorbed or deflected away from subject position” and must also stipulate that (iia) cannot be due to topicalization because for some reason “reconstruction is obligatory in LF for PP arguments” (p. 269) and the CRP holds at all levels of grammar.

Notice as well that this purported special property of the copula does not, as it should, rescue sentential subjects from the CRP. This is seen here in (iii).

- (iii) is [that the cat is under chair] ok with you?

Example (iii) is no better, and no worse, than other subject-aux inversion examples with other verbs and modals.

4. Of course, under Stowell’s account, it might be claimed that the CRP does not apply to nonfinite IPs because they are [–tense]. However, under more recent theories of phrase structure (e.g., Grimshaw 1991), IP is just an “extended projection” of VP and shares its lexical category features [+V, –N]. Additionally, we note that the sentences in (18) are awkward,

and the status of their acceptability may not differ significantly from comparable sentences with tensed CP subjects. Reasons for this are addressed below.

5. Among the predicates taking absolutive subjects are *amanaki* 'hope', *manako* 'want', *piko* 'believe', and others.

6. In the remainder of the chapter, we forgo indicating degrees of acceptability on any examples that we generate. While there are important differences in acceptability that we will point to, these differences are relative, not absolute. Data cited from other sources are reproduced with the acceptability judgment of the source.

7. This dispreference led to a number of syntactic proposals that would rule out internal sentential subjects in English such as Ross's (1973) Internal S Condition and Kuno's (1973) constraint that subject sentences can only appear in sentence-initial position.

8. Lees and Rosenbaum also analyzed sentential complements of verbs as having NP immediately dominating S. This was abandoned within Extended Standard Theory in Chomsky 1973 so that extraction from sentential complements would not violate subadjacency. The resulting asymmetrical analysis of sentential subjects and objects, as formulated in Chomsky 1973, is a view we argue to be essentially correct (Davies and Dubinsky 1999, 2001).

9. Because emphatic reflexives can only be used with phrases whose denotation is definite, AP subjects, which are properties, do not share this NP-like trait.

## References

- Brame, M. 1976. *Conjectures and refutations in syntax and semantics*. New York: North Holland.
- Chomsky, N. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart and Winston.
- Chung, S. 1982. Unbounded dependencies in Chamorro grammar. *Linguistic Inquiry* 13:39–77.
- Chung, S. 1991. Sentential subjects and proper government in Chamorro. In C. Georgopolous and R. Ishihara, eds., *Interdisciplinary approaches to language*, 75–99. Dordrecht: Kluwer.
- Culicover, P. 1997. *Principles and parameters: An introduction to syntactic theory*. Oxford: Oxford University Press.
- Davies, W. D., and S. Dubinsky. 1999. Sentential subjects as complex NPs: New reasons for an old account of subadjacency. *Papers from the Thirty-Fourth Regional Meeting, Chicago Linguistic Society*, 83–94. Chicago Linguistic Society, University of Chicago.
- Davies, W. D., and S. Dubinsky. 2001. Functional architecture and the distribution of subject properies. In W. D. Davies and S. Dubinsky, eds., *Objects and other subjects: Grammatical functions, functional categories, and configurationality*, 247–279. Dordrecht: Kluwer.
- Delahunty, G. 1981. *Topics in the syntax and semantics of English cleft sentences*. Bloomington: Indiana University Linguistics Club.
- Delahunty, G. 1983. But sentential subjects do exist. *Linguistic Analysis* 12:379–398.
- Dryer, M. 1980. The positional tendencies of sentential noun phrases in universal grammar. *Canadian Journal of Linguistics* 25:123–195.



- Emonds, J. 1970. Root and structure-preserving transformations. Doctoral dissertation, MIT.
- Emonds, J. 1972. A reformulation of certain syntactic transformations. In S. Peters, ed., *Goals of linguistic theory*, 21–62. Englewood Cliffs, NJ: Prentice Hall.
- Emonds, J. 1976. *A transformational approach to English syntax: Root, structure-preserving, and local transformations*. New York: Academic Press.
- Erdmann, P. 1988. On the principle of “weight” in English. In C. Duncan-Rose and T. Vennemann, eds., *On language*, 325–339. London: Routledge.
- Fodor, J., T. Bever, and M. Garrett. 1974. *The psychology of language: An introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill.
- Frazier, L., and K. Rayner. 1988. Parameterizing the language processing system: Left versus right branching within and across languages. In J. A. Hawkins, ed., *Explaining language universals*, 247–279. Oxford: Blackwell.
- Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68:1–76.
- Grimshaw, J. 1991. Extended projections. Unpublished manuscript, Rutgers University.
- Grosu, A., and S. Thompson. 1977. Constraints on the distribution of NP clauses. *Language* 53:104–151.
- Haegeman, L., and J. Guéron. 1999. *English grammar: A generative perspective*. Malden, MA: Blackwell.
- Hawkins, J. A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- Iwakura, K. 1978. On root transformations and the structure preserving hypothesis. *Linguistic Analysis* 4:321–364.
- Koster, J. 1978. Why subject sentences don’t exist. In S. Jay Keyser, ed., *Recent transformational studies in European languages*, 53–64. Cambridge, MA: MIT Press.
- Kuno, S. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.
- Lasnik, H. 1995. Last resort. In S. Haraguchi and M. Funaki, eds., *Minimalism and linguistic theory*, 1–32. **Numazu City: Numazu Linguistics Seminar.**
- Lasnik, H. 1999. *Minimalist analysis*. Malden, MA: Blackwell.
- Lees, R. 1960. *The grammar of English nominalizations*. The Hague: Mouton.
- Levin, J., and D. Massam. 1986. Classification of Niuean verbs: Notes on case. In P. Geraghty, L. Carrington, and S. A. Wurm, eds., *FOCAL I: Papers from the Fourth International Conference on Austronesian Linguistics*, 231–244. Canberra: Australian National University.
- Levine, R. 1989. On focus inversion: Syntactic valence and the role of a SUBCAT list. *Linguistics* 17:1013–1055.
- McCloskey, J. 1991. *There, it, and agreement*. *Linguistic Inquiry* 22:563–567.
- Postal, P. 1974. *On Raising*. Cambridge, MA: MIT Press.
- Radford, A. 1997. Syntactic theory and the structure of English: A minimalist approach. Cambridge: Cambridge University Press.
- Rosenbaum, P. 1967. *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.

- Ross, J. R. 1973. The same side filter. *Papers from the Ninth Regional Meeting, Chicago Linguistic Society*, 549–566. Chicago Linguistic Society, University of Chicago.
- Safir, K. 1982. Syntactic chains and the definiteness effect. Doctoral dissertation, MIT.
- Safir, K. 1983. On small clauses as constituents. *Linguistic Inquiry* 14:730–735.
- Stowell, T. 1981. Origins of phrase structure. Doctoral dissertation, MIT.
- Wasow, T. 1997. Remarks on grammatical weight. *Language Variation and Change* 9:81–105.
- Widmann, C. 2005. Factors at play in determining the acceptability of sentential subjects in English: The role of constituent relative weight. Unpublished manuscript, University of South Carolina.