The Tough Construction is neither Extraction nor Control
Author(s): Yoon-Suk Chung
Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetics and Phonological Universals (1998), pp. 44-56

The Tough Construction is neither Extraction nor Control

Yoon-Suk Chung
UC Berkeley

1. INTRODUCTION. The English *tough* construction (TC), illustrated in (1), is a structure in which one of a small class of adjectives like *tough*, *easy*, etc.(= *tough* predicates) has a subject which is understood as the so-called missing argument of an infinitival complement.

(1) a. John is easy to please.
b. Mary would be difficult for anyone to talk to.

The TC has been identified in generative grammar either as a species of extraction, or as a species of control. In this paper, I will argue that the TC is neither extraction nor control, but a third type of construction that shares one salient grammatical property with these other two constructions.

First, I will review two accounts of the TC, the extraction account and the control account, and point out some problems with these accounts. I will then suggest a new construction grammar-based account of the TC.

2. TWO ACCOUNTS. The essential problem presented by the TC is how to account for the missing argument embedded under a *tough* predicate, the so-called *tough* gap.[1]

For the past four decades, this question has been answered in terms of already existing analyses of English syntactic phenomena. That is, the TC has sometimes been treated as a kind of extraction construction which otherwise include things like *wh*-questions, and topicalization constructions, etc. as illustrated in (2).

(2) a. Who do you think John admired?
b. Mary, I think John admired.

Alternatively, it has been treated as a kind of control construction, which otherwise includes equi and raising phenomena, as in (3).

(3) a. John tried to leave.
b. John seems to have left.

2.1. EXTRACTION ACCOUNT. The Extraction account basically says that the *tough* gap is created when an argument is extracted or displaced from an otherwise realized embedded argument position.

The extraction account started with Chomsky (1977), and has become the dominant view. For example, in GB, Chomsky (1986) suggests a D-structure (4a) for sentences like (1a):

(4) a. John is easy $[_{CP}\ [_{IP}$ PRO to please OP]
b. John$_i$ is easy $[_{CP}$ OP$_i$ $[_{IP}$ PRO to please t_i]

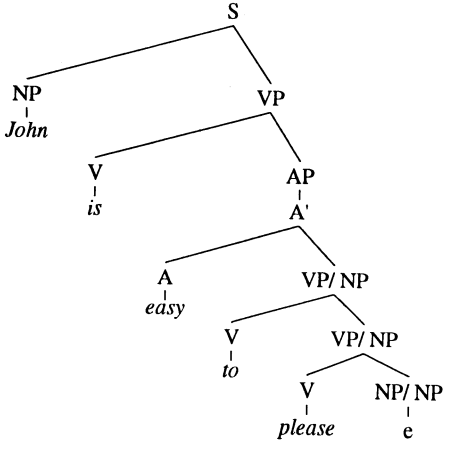
In (4b), a trace is created in the missing argument position as a result of movement of a null operator (OP) to the clause-initial, non-argument position of the embedded

infinitival phrase. This operator is later coindexed with the matrix subject of the TC (= the *tough* subject).

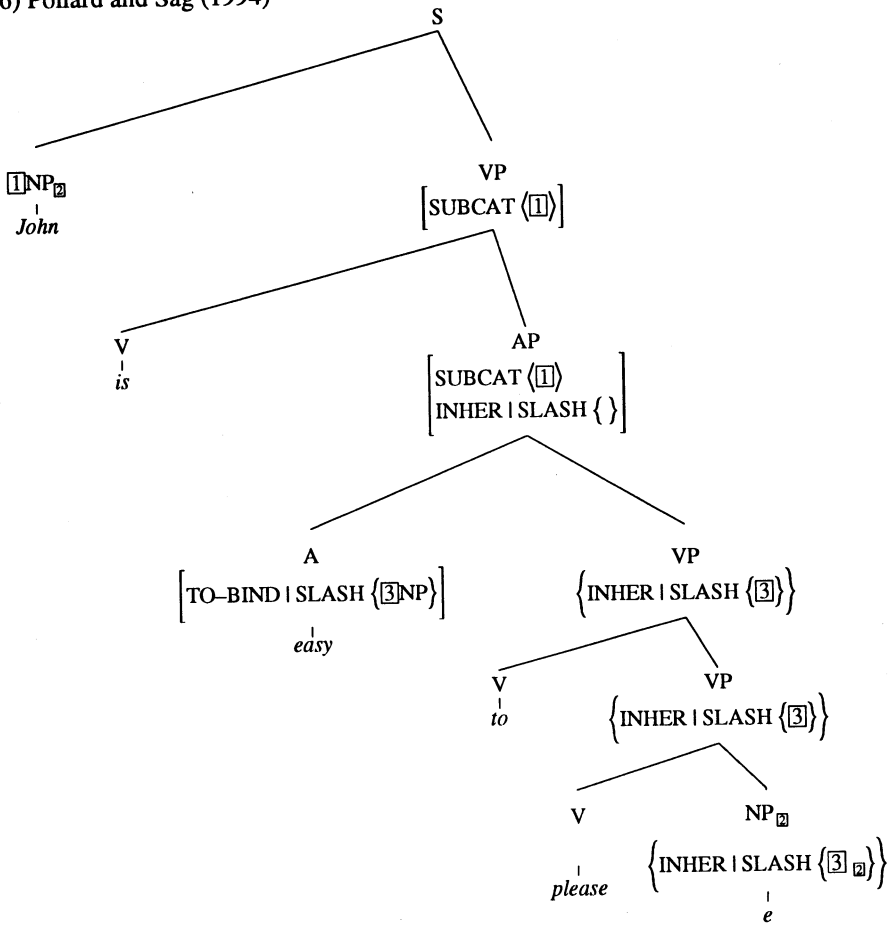
Basically, the same idea has been implemented in monostratal frameworks. In standard GPSG or HPSG such as Gazdar et al. (1985), and Pollard and Sag (1994), the *tough* gap arises as the value of the gap feature SLASH propagates from the gap position via extraction path to the position where it is cashed out, or receives its value.

The trees shown in (5) and (6) should make this clear. In (5) of the Gazdar et al. (1985) treatment, the SLASH value is cashed out by the *tough* subject at the node of a matrix sentence. In (6) from Pollard and Sag (1994), the SLASH value is cashed out by the *tough* predicate at the node of an AP. The lexical entry *easy* ensures the coindexation of the trace with the *tough* subject in the lexicon.

(5) Gazdar et al. (1985)



(6) Pollard and Sag (1994)



2.2. ARGUMENTS FOR THE EXTRACTION ACCOUNT. There are three main arguments for the Extraction account. First, like extraction, seen in (7a), the TC may involve long distance dependencies. See (7b).

(7) a. Which gadget did you persuade people to buy?
b. That gadget would be difficult to persuade people to buy.

In (7), the dependencies between the missing argument positions and their antecedents occur across more than one clause boundary.[2]

Second, like extraction, the TC is subject to some island constraints.[3] Consider the examples in (8)-(9).

(8) a. *Who do you imagine the likelihood of entertaining?[4]
b. *John would be difficult to imagine the likelihood of entertaining.

(9) a. *Who do you imagine a person who entertains?
b. *John would be difficult to imagine a person who entertains.

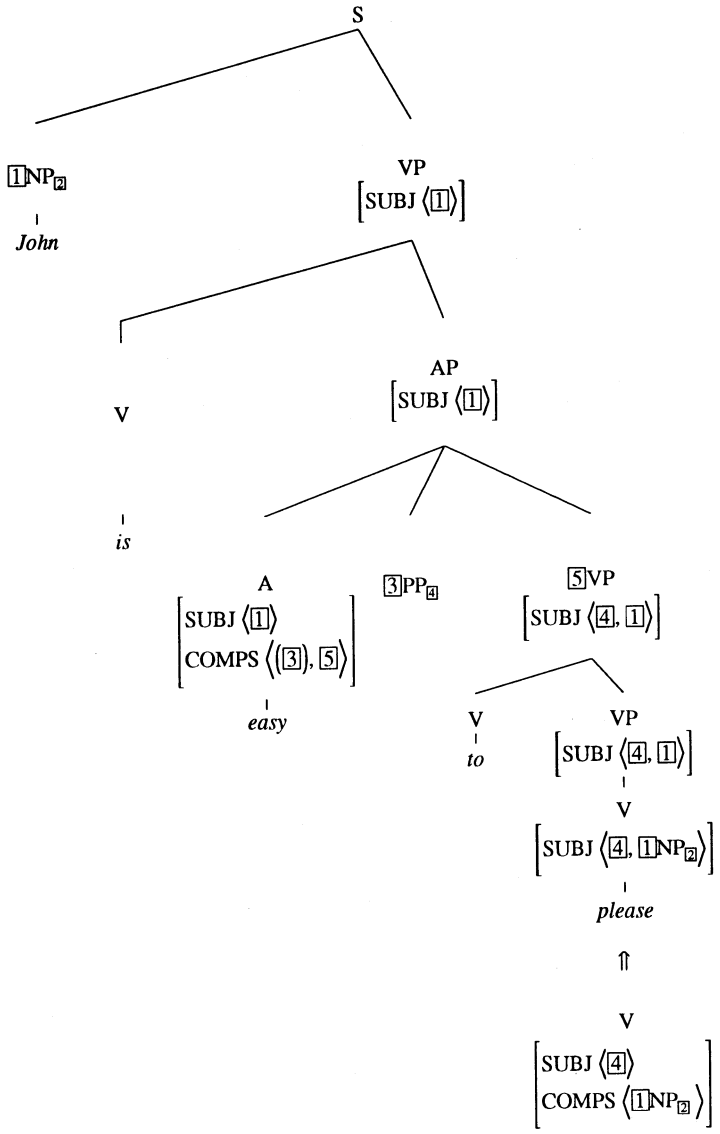
(8b), like (8a), is ungrammatical since it violates the Complex NP Constraint. The same is true with (9).

Finally, like extraction in (10a), the TC allows parasitic gaps, like the one in (10b).

(10) a. Which paper do you think you filed without reading?
b. This paper is easy to file without reading.

2.3. CONTROL ACCOUNT. The Control account, on the other hand, says that the *tough* gap is a 'controlled' argument in the way that the understood subject in an equi or raising construction is controlled by an argument of a higher predicate. Control-type accounts include the *tough* deletion account, which is essentially the same as the Equi NP deletion account, and a movement analysis which parallels the traditional treatment of raising. Control accounts also include some less familiar analyses in monostratal frameworks like Grover (1995). (11) is a representation for a structure (1a) on this account.

(11) Grover (1995)



(12) Grover's lexical rule

$$\left[\text{SYNSEM} \mid \text{LOC} \left[\text{CAT} \left[\begin{array}{l}\text{SUBJ}\ \langle ... \rangle \\ \text{COMPS}\ \langle ... , \boxed{1}\text{NP}, ... \rangle\end{array}\right]\right]\right]$$

$$\Rightarrow \left[\text{SYNSEM} \mid \text{LOC} \left[\text{CAT} \left[\begin{array}{l}\text{SUBJ}\ \langle ... , \boxed{1} \rangle \\ \text{COMPS}\ \langle ... \rangle\end{array}\right]\right]\right]$$

For Grover's HPSG analysis, the *tough* gap is created when the missing argument is promoted, by a lexical rule (12), from the COMPS list to the SUBJ list of the embedded verb's subcategorization structure, seen at the bottom of the diagram (11). The promoted subject argument, $\boxed{1}\text{NP}_{\boxed{2}}$ in (11), is controlled by the *tough* subject in accordance with general principles that govern the relationship between the subject of the controlled VP and its controlling subject.

2.4. ARGUMENTS FOR THE CONTROL ACCOUNT. There are two arguments for the Control account. First, the 'controller' of the *tough* gap occurs in an argument position, i.e. specifier of IP, not a non-argument position, i.e. specifier of CP. So, if your theory forces you to choose between control and extraction accounts, this fact makes the TC look more like a control phenomenon.

Second, the TC is only licensed by specific lexical predicates such as *tough*, *easy*, etc., just as control constructions are only licensed by certain predicates.

(13) a. This door is impossible/*possible to open.
b. John tried/*screamed to open the door. (*agentive reading)

Just as a predicate like *screamed* in (13b) does not participate in the equi construction, a predicate like *possible* in (13a) may not participate in the TC.[5]

2.5. PROBLEMS WITH THE EXTRACTION AND THE CONTROL ACCOUNTS. Proponents of the extraction account emphasize the extraction properties, and de-emphasize the control properties, dismissing them as irrelevant or trivial. Proponents of the control account, on the other hand, play down the extraction properties, and emphasize the control properties of the TC. Thus, each account is unable to account for some portion of the data.

Furthermore, there are other properties of the TC which one or the other or both accounts cannot explain. First, unlike extraction, the TC is not always an island for extraction. Multiple gaps are sometimes permitted in the TC, like (14a). This property is problematic for an extraction account since canonical extraction constructions *are* islands to extraction, as seen in (14b).

(14) a. Which violin is this sonata easy to play _ on _?[6]
b. * Which violin did you ask which sonata Mary played _ on _?

Second, (15) shows that the TC does not allow nominative gaps.

(15) a. *John is difficult for me to believe _ went to Chicago.
b. *Mary is tough to believe _ likes John.

This property is problematic for an extraction account since, as seen in (16), embedded subject gaps *are* possible with canonical extraction constructions.

(16) a. Who do you believe _ likes John?
b. Mary, I think _ likes John.

Third, as seen in (17a), the case marking in the TC is different between the gap and its antecedent: the antecedent is nominative, but the gap is in the accusative position. This property is problematic for an extraction account since in extraction constructions the case assigned in the missing argument position is carried over to the antecedent as in (17b).

(17) a. She $_{NOM}$ is easy for me to please $_{-ACC}$
b. Whom $_{ACC}$ do you think you pleased $_{-ACC}$?

Fourth, unlike control, the 'controllee' of the TC must be a non-subject of the non-finite verb phrase, as shown by (15), repeated here as (18).

(18) a. *John is difficult for me to believe _ went to Chicago.
b. *Mary is tough to believe likes _ John.

This is in contrast with control where the controllee is typically a subject.

Finally, there are problems for both the extraction and the control accounts. Depending on the speaker, a *tough* gap may or may not be associated with the *tough* subject across a finite clause boundary. For example, while Hukari and Levine (1991), and Grover (1995) judge the examples in (19) to be ungrammatical, Bresnan (1971), and Pollard and Sag (1994) accept them.

(19) a. %John is difficult for me to believe Mary loves.
b. %Mary is easy for me to think John cheated.

The extraction approach cannot account for why extraction from finite clauses is bad for some speakers since extraction is perfectly good with vanilla extraction constructions, like the ones in (20).

(20) a. Who do you want me to believe Mary loves?
b. John, I really want you to believe Mary loves.

The control approach, on the other hand, cannot account for why extraction from finite clauses is good for the other speakers since the controllee of control predicates typically occurs within a non-finite VP complement, and in (19), the VP complement is finite.

Despite these apparent problems, the relegation of the TC to either extraction or control is inevitable in most major syntactic frameworks, since they lack any

means to capture the common ground shared by extraction and control. In HPSG (Pollard and Sag 1994, Grover 1995), control is handled in the lexicon, whereas extraction is treated in the syntax, with a strict line drawn between the two components. In GB (Chomsky 1977, 1981, 1986, Cinque 1990), both control and extraction are handled in the syntax, but by different modules. The occurrence and interpretation of the understood subject of a non-finite VP complement, control, is basically regulated by the PRO theorem or NP-movement, and extraction is regulated by subjacency constraints, together with some surface filters. Since control and extraction are regulated by independent modules, a unified account of a phenomenon which exhibits both of the properties is not feasible.

3. NEW PROPOSAL

3.1. SIMILARITIES AND DIFFERENCES BETWEEN EXTRACTION, CONTROL, AND TOUGH CONSTRUCTION. Table 1 represents the common ground between the TC and extraction, on the one hand, and control on the other hand. From this point, I will adopt Construction Grammar terminology for extraction and control. We call extraction the Left Isolate Construction, and we call control the Coinstantiation Construction.

	Left Isolate	Coinstantiation	Tough
Non-local argument construal	+	+	+
Long distance dependency	+	-	+
Single valence structure	-	+	+

Table 1. Relationship between the Left Isolate, Coinstantiation and Tough Construction

The first row of Table 1 shows that all three constructions involve non-local argument construal. In other words, they involve specification of an argument's semantics by an argument that is not realized syntactically within a local structure. So, in the resulting construct, some argument is construed as instantiating a valence requirement of a predicate that occurs in a syntactic domain not governed by that predicate. The second row of Table 1 indicates that the TC shares with the Left Isolate Construction the property, absent in the Coinstantiation Construction, that the dependency may be a long-distance one. The third row indicates that the TC shares with the Coinstantiation Construction the property, absent in the Left Isolate Construction, that the entire construction occurs within a single valence structure. That is, in the TC, there is no left-isolate sister as a distinct syntactic position, as there is in the Left Isolate Construction.

3.2. THE THREE CONSTRUCTIONS. Now I will demonstrate how the similarities and differences between the three constructions are captured in a construction grammar approach.

Before discussing each of these constructions, two preliminary notes are in order. First, the descriptions of the constructions we will look at employ feature structures consisting of attributes and their appropriate values. Second, the descriptions of the constructions will be fleshed out just enough to illustrate the similarities and the differences between them.

3.2.1. THE LEFT ISOLATE CONSTRUCTION. The Left Isolate Construction (= extraction) is given in Figure 1.

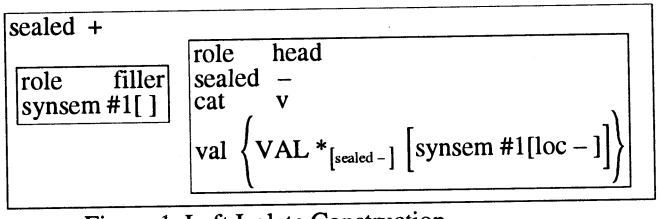


Figure 1. Left Isolate Construction

Figure 1 says that the Left Isolate Construction consists of two daughters. It is a verbal structure since the right head daughter is verbal, as indicated by 'role head', 'cat, v'. This is achieved by an independent Head Principle in Construction Grammar. Note that since the right daughter does not mention any information about specific verbal forms, it can be either finite or non-finite.

If we move to the valence structure at the bottom of the right daughter, we see some notation in the valence set. For the moment, ignore the operator Kleene-starred VAL with the attached subscript feature structure [sealed -]. The remaining valence value, represented as 'synsem #1[ ]', when read in conjunction with the same synsem value in the left daughter, has the effect of saying that the left daughter fills or satisfies a valence requirement of a predicator that occurs in the right daughter. Note that what is unified is syntactic and semantic information, so that the case marking effects we saw in (17) are maintained between an 'extracted' element and the corresponding valence requirement. The 'loc -' value in the synsem of the right daughter represents the fact that the valence requirement is not locally instantiated, that is, it is extracted. Note also that no restriction is imposed on the extracted element so that the extracted element can satisfy any grammatical function in the valence of the right daughter.

Recall that there might in principle be an unrestricted number of other intervening predicators between the left-isolated valence element, and the predicator whose valence it satisfies. The operator VAL* indicates this.

The formal expression VAL* [synsem #1] has the effect of saying that there is a valence element, which unifies with the left-isolated daughter, which is arbitrarily multiply embedded in the right daughter.

The feature structure 'sealed +' in the external structure of the Left Isolate Construction captures the fact that the Left Isolate Construction is an island, that nothing can be extracted or left-isolated from it, as we saw in (8)-(9). This feature stucture is in constrast with the one 'sealed -' in the right daughter since any predicator in the right daughter may permit its argument to be left-isolated.

3.2.2. THE COINSTANTIATION CONSTRUCTION. A simplified version of the Coinstantiation Construction (= control) is given in Figure 2.

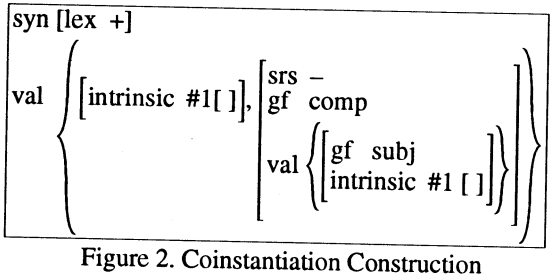


Figure 2. Coinstantiation Construction

The top line of the Coinstantiation Construction encodes the fact that the construction is licensed by a specific lexical item.

The lexical predicator heads a valence structure that has two valence elements: the first is an argument, and the second is a controlled subjectless non-finite verbal complement, as indicated by 'srs -' and 'gf comp'.

The unification index #1[ ] in the controlled complement, when read in conjunction with the same intrinsic value of the first valence element, indicates that the subject requirement of the complement has its intrinsic features — non-relational syntactic and semantic information — supplied by another valence element of the predicator. The fact that the controlling element is not realized as a separate constituent corresponds to the fact on other accounts that it will be satisfied by a constituent that appears in an argument, as opposed to a non-argument, position.

3.2.3. THE TOUGH CONSTRUCTION. A simplified version of the TC is given in Figure 3.

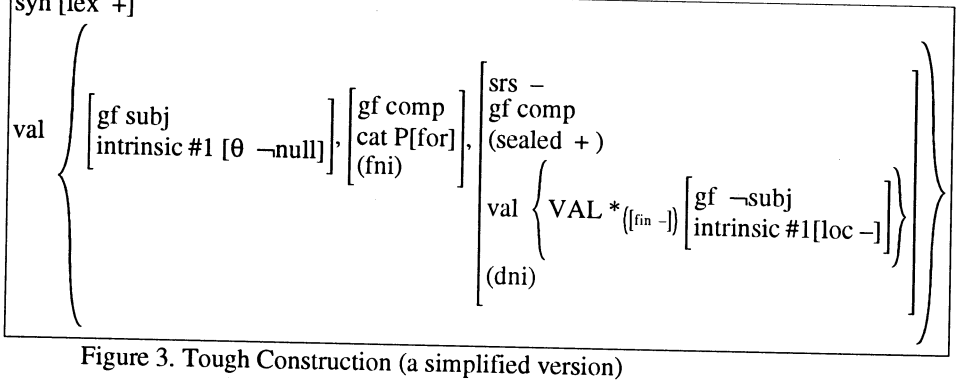


Figure 3. Tough Construction (a simplified version)

The top line says that, like Coinstantiation, the construction is licensed by a specific lexical item. The valence structure says that the predicate has three valence elements.

The first element encodes the *tough* subject. As indicated in footnote 2, the *tough* subject may be equi, but not raising, objects. This restriction is marked 'θ ¬null ' in the *tough* subject intrinsic valence.

The second and third valence elements represent the two oblique arguments, indicated as 'gf comp' that *tough* predicates take. The second element is a prepositional phrase headed by the preposition *for*. This is the 'for anyone' constituent that we saw in (1b). However, the syntactic realization of this prepositional phrase is not mandatory. When the prepositional phrase is not manifested, the missing constituent is interpreted as people in general or as the speaker or the addressee or some third party whose perspective is being employed in a discourse: free null instantiation (fni).

The third valence element is a subjectless infinitival complement, similar to the controlled complement in the Coinstantiation Construction. The constituent is not always an island, unlike the Left Isolate Construction, as we saw in (14a) and footnote 6. This restriction is indicated by '(sealed +)'. The constituent may also be absent. When it is absent, the address may be able to construe its content from the the context, and is marked as '(dni)' (definite null instantiation) at the bottom of the valence structure.

The third valence element has its own valence requirement. It resembles the valence requirement of the Left Isolate Construction's right daughter in that it allows for multiple embedding. It differs from the Left Isolate Construction in that only intrinsic information is shared with another valence element, namely, the *tough* subject. It differs from the Coinstantiation Construction in that the grammatical function of the unified element must be a non-subject within the complement so that the case markings between the *tough* subject and the *tough* gap must be different. As we saw in (19), the coindexation of the tough *subject* with the *tough* gap may or may not, depending on the dialects, be made across the a finite clause boundary that an intervening predicator hosts. This fact is captured by marking '(fin -)' on a valence structure that any intervening predicator has.

4. CONCLUSION. In this paper, we hope to have shown that the TC has its own syntactic properties that cannot belong to either the extraction construction or the control construction.

On the Construction Grammar account, the TC is subsumed neither under the Left Isolate Construction nor under the Coinstantiation Construction. The TC instead shares particular properties of both constructions, and satisfies its own well-formedness conditions. Such an account is not available to other approaches due to their general organizational principles.

If the analysis proposed is correct, it casts doubt on the basic assumption of the distinction between lexicon and syntax in GB or HPSG. The case study of the TC shows that such a widely accepted view cannot be maintained, and points to the desirability of a more flexible design of grammar such as that provided by Construction Grammar.

NOTES

1. This issue should not be confused with another independent issue of equi vs. raising which concerns the semantic relation between the *tough* predicate and the subject. See also Grover (1995).
2. Jacobson (1992) notes, however, that it is sometimes hard to get a deeply embedded gap in some *tough* sentences:

(i) ?John is hard for me to imagine Mary wanting to meet.

But such difficulties might be explained by the semantic oddness of the to-infinitive phrase, i.e., 'I imagine Mary wants to meet X'. More generally, Zwicky (1987) independently notes that the anomaly seems to be related to the fact that object-to-subject raised NPs, unlike equi NPs, usually do not participate in the TC. Contrast (i) and (ii) with (iii):

(ii) ?John is hard to expect to marry Leslie.
(iii) John is hard to persuade to marry Leslie.

I will not look into the details of the issue, since it does not have any bearing on the extraction vs. control issue. The observed differences between raising objects and equi objects will, however, be incorporated in our analysis in section 3.

3. Or the subjacency condition since it subsumes many, though not all, of the island constraints.

4. Grover (1995) notes, however, that some NPs including picture nouns behave somewhat differently:

(i) *Sandy is hard to sell some pictures of.
(ii) Who did you sell some pictures of?

(iii) *France would be impossible to meet the king of.
(iv) Which country did George meet the king of?

Since the examples involved share the same syntactic structure, the unexpected difference in grammaticality shown in (i, ii) and (iii, iv) might be due to some semantic factors, for which we offer no account here.

5. See Akatsuka (1979) for reasons why a predicate like *possible* cannot function as a *tough* predicate.

6. Thus there is a well-known difference in grammaticality judgements between (14a) and (i):

(i) *Which sonata is this violin easy to play_on_ ?

We do not, however, offer any accout of the difference here.

ACKNOWLEDGEMENTS

I am grateful to Chuck Fillmore, Andreas Kathol, especially Paul Kay, and David Peterson for their helpful contributions. Usual disclaimers apply.

REFERENCES

AKATSUKA, NORIKO. 1979. Why *Tough*-Movement is impossible with *possible*. Chicago Linguistic Society 15.1-8.

BRESNAN, JOAN. 1971. Sentence Stress and Syntactic Transformations. Language 47.257-281.

CHOMSKY, NOAM. 1977. On *wh*-movement. Formal syntax, ed. by P. Culicover, T. Wasow, and A. Akmajian, 71-132. New York: Academic Press.

CHOMSKY, NOAM. 1981. Lectures on Government and Binding. Dordrecht, Holland: Foris.

CHOMSKY, NOAM. 1986. Knowledge of Language: Its Nature, Origin and Use. New York: Praeger.
CHOMSKY, NOAM. 1995. A minimalist program for linguistic theory. The minimalist program 167-218. Cambridge, MA: MIT Press.
CINQUE, GUGLIELMO.1990. Types of A'-Dependencies. Cambridge, MA: MIT Press.
FILLMORE, CHARLES, and PAUL KAY. To appear. Construction Grammar. Stanford, CA: CSLI.
GAZDAR, GERALD; EWAN KLEIN; GEFFREY. K. PULLUM; and IVAN SAG. 1985. Generalized Phrase Structure Grammar. Cambridge, MA: Harvard University Press.
GROVER, CLAIRE. 1995. Rethinking Some Emtpy Categories: Missing Objects and Parasitic Gaps in HPSG. Edinburgh: University of Edinburgh dissertation.
HUKARI, THOMAS, and ROBERT LEVINE. 1991. On the Disunity of Unbounded Dependency Constructions. Natural Language & Linguistic Theory 9.97-114.
JACOBSON, PAULINE. 1992. The Lexical Entailment Theory of Control and the *Tough*-Construction. Lexical Matters, ed. by I. Sag and A. Szabolcsi, 269-299. Stanford: CSLI.
KAY, PAUL, and CHARLES FILLMORE. 1998. Grammatical constructions and linguistic generalizations: the *What's X doing Y?* construction. Language 74.
POLLARD, CARL, and IVAN SAG. 1994. Head-Driven Phrase Structure Grammar. Stanford, CA: CSLI.
ZWICKY, ARNOLD. 1987. Slashes in the passive. Linguistics 25.639-669.