On Optionality: A Brief History and a Case Study

Howard Lasnik
University of Maryland

I. Some History

(1) In classic transformational grammar, optionality vs. obligatoriness was coded on particular transformations, as an arbitrary specification.

(2) Excerpt from Chomsky (1957):

12. *Passive* – optional:
Structural analysis: $NP - Aux - V - NP$
Structural change: $X_1 - X_2 - X_3 - X_4 \rightarrow X_4 - X_2 + be + en - X_3 - by + X_1$ (34)

13. T^{ob}_{sep} – obligatory:
Structural analysis: $\begin{Bmatrix} X - V_1 - Prt - Pronoun \\ X - V_2 - Comp - NP \end{Bmatrix}$ (86) (92)
Structural change: $X_1 - X_2 - X_3 - X_4 \rightarrow X_1 - X_2 - X_4 - X_3$

14. T^{op}_{sep} – optional:
Structural analysis: $X - V_1 - Prt - NP$ (85)
Structural change: same as 13

15. *Number Transformation* – obligatory
Structural analysis: $X - C - Y$
Structural change: $C \rightarrow \begin{Bmatrix} S \text{ in the context } NP_{sing} - \\ \emptyset \text{ in other contexts} \\ past \text{ in any context} \end{Bmatrix}$ (29i)

16. T_{not} – optional
Structural analysis: $\begin{Bmatrix} NP - C - V \ldots \\ NP - C + M - \ldots \\ NP - C + have - \ldots \\ NP - C + be - \ldots \end{Bmatrix}$ (37)
Structural change: $X_1 - X_2 - X_3 \rightarrow X_1 - X_2 + n't - X_3$

17. T_A – optional:
Structural analysis: same as 16 (cf. (45)–(47))
Structural change: $X_1 - X_2 - X_3 \rightarrow X_1 - X_2 + A - X_3$

18. T_q – optional:
Structural analysis: same as 16 (cf. (41)–(43))
Structural change: $X_1 - X_2 - X_3 \rightarrow X_2 - X_1 - X_3$

(3) Chomsky (1956) had an interesting comment on this aspect of the theory:

(4) "The grammar can also be simplified very greatly if we ... distinguish between obligatory rules which must be applied when we reach them in the sequence and optional rules which may or may not be applied." [p.118]

(5) The grammar qua grammar surely isn't simpler if we allow this arbitrary 2-way specification. What Chomsky presumably had in mind was that given that the phenomena he was investigating displayed apparent optionality and obligatoriness, to describe those phenomena, the simplest thing to do is what he did.

(6) Chomsky (1965) reconsidered this simplicity issue:

(7) "... several important points have gradually emerged which suggest that a somewhat more restricted and conceptually simpler theory of transformations may be adequate.

First, it has been shown that many of the optional singulary transformations of Chomsky (1955, 1957, 1962) must be reformulated as obligatory transformations, whose applicability to a string is determined by presence or absence of a certain marker in the string." [p.132]

(8) Chomsky specifically mentioned two of the transformations in (2):

(9) "This was pointed out by Lees (1960a) for the negation transformation, and by Klima (personal communication) for the question transformation, at about the same time."

(10) This idea was developed by Katz and Postal (1964) with respect to the question of how syntax interfaces with semantics. Chomsky (1965) says:

(11) "Katz and Postal (1964) have extended these observations and formulated them in terms of a general principle, namely that the only contribution of transformations to semantic interpretation is that they interrelate Phrase-markers (i.e., combine semantic interpretations of already interpreted Phrase-markers in a fixed way). It follows, then, that transformations cannot introduce meaning-bearing elements ..."

(12) This formed the basis for the Standard Theory, based on the postulation of deep structure, and that level as the sole locus of semantic interpretation.

(13) Katz and Postal had argued as follows:

(14) "... there are ... many cases in the literature of syntactic facts characterized by optional singulary transformations where the output P-marker must have a semantic interpretation quite different from that of the input P-marker. Among these are the question transformation, the imperative transformation, the wh- attachment transformation, etc." [p.32]

(15) The concern of Katz and Postal was how to simplify the semantic interpretive rules. They discussed several possibilities, ultimately suggesting the following:

(16) "... no correctly formulated singulary transformation has an output with a semantic interpretation distinct from its input and that those transformations in the literature which violate this claim are incorrect ..."

(17) Katz and Postal then dicussed a range of transformational phenomena that had been treated as involving optional transformations: negation, interrogatives, and imperatives. They argued, instead, that in initial P-markers, there are negation, imperative, and interrogative markers that have semantic import. These then trigger obliagtory transformational operations (Subject-Aux inversion ((17) in (2) above), for instance). They claimed that not only would this simplify the theory of semantic interpretation, it would provide more adequate syntactic descriptions.

(18) The discussion of passive by Katz and Postal is interesting. They began with a pair of sentences discussed by Chomsky (1957), which appear to conflict with their theory:

(19) a. everyone in the room knows two languages
b. two languages are known by everyone in the room

(20) They first rejected Chomsky's claim that (19)a,b differ in meaning, a claim incompatible with their theory of semantic interpretation.

(21) "Although the facts are far from clear, the active [(19)a] seems to be open to the same interpretation attributed to the passive [(19)b], and conversely, the passive is open to the same interpretation attributed to the active. Both [(19)a] and [(19)b] can mean either 'everyone in the room knows the same two particular languages, Persian and Hottentot' or 'everyone in the room knows two languages, different for different people'. Thus it seems that both actives and passives containing quantifiers and pronouns are ambiguous in the same way and so are full paraphrases of each other." [p.72]

(22) However, they went on to claim that even if the active and passive did differ in meaning, that would still not undermine their theory:

(23) "But even if the meanings of examples like [(19)a] and [(19)b] are different, the argument that some transformations affect meaning does not hold. This argument must also assume that such examples are transformationally related, i. e., that the passives are derived from the application of a transformation to the P-marker underlying the corresponding active form."

(24) They claimed that actives and passives do not arise from a single PM via an optional passive T. Rather, the initial PM of passives has (among other things) a passive morpheme, which triggers the remainder of the changes.

(25) On the face of it, this is quite parallel to the negative, imperative, and question arguments. But there is, I think, a crucial difference:

(26) For those situations, the meaning difference can be pinned entirely on the designated morpheme. But it is not nearly so straightforward how the passive morpheme might yield scope reversal.

(27) There have been other sorts of arguments against the optional/obligatory distinction for rules. The most prominent concerns learnability.

(28) Arbitrary marking of optional or obligatory on rules results in what I call quantitative and qualitative acquisition problems.

(29) Imagine a classic transformational component of a grammar of the sort developed in the 1950's and early 1960's. Such a grammar had a large number N of transformations. As noted before, in that classic model, each T could be marked optional or obligatory. So even with everything else fixed, there are 2^N grammars to be considered by the learner, a quantitative problem.

(30) There is also a 'qualitative' problem. Asuming any T could be optional or obligatory, there are four situations for the learner, two of which cause no difficulty at all. If the child's hypothesis is that the given rule is optional, and the rule actaully is optional in the target grammar, nothing more needs to be done. Similarly for obligatory. The remaining two situations involve mismatches.

(31) When the hypothesis is *obligatory* and the traget is *optional*, the learning problem is relatively trivial. Positive evidence, in the form of sentences whose derivation involved non-application of the rule in question, should induce the requisite hypothesis change.

(32) The other mismatch is more difficult: where the hypothesis is *optional* and the target is *obligatory*. Here, all else equal, the target language is a proper subset of the hypothesized language, so there will be no relevant positive evidence. Under the widespread assumption that negative evidence is not sytematically available to learners, this situation is untenable.

(33) Braine (1971) was an early influential examination of the general subset problem. "Information about what is not a sentence would appear to be necessary in order for the learner to reject hypothetical grammars and grammatical rules which are "overinclusive" (i.e. which generate all the acceptable strings, and which err only because they also generate unacceptable strings). Since such grammars generate all the good sentences to which the learner is exposed, how can he discover that they are wrong unless his input data contains information about nonsentences?" [p.157]

(34) There are three immediately apparent ways to address this problem:
a. All rules are obligatory (an extension of the Katz-Postal conjecture). Apparent transformational optionality is a consequence of an optionally chosen marker in the base.
b. All rules are optional. Apparent obligatoriness is a consequence of a surface filter that would be violated if the rule didn't apply.
c. Rules can be obligatory or optional, but obligatory is the unmarked case, chosen in the absence of counter-evidence.

(35) Given the apparent optionality of a number of processes, a. was not extensively pursued in the middle stages of the development of transformational generative grammar. But versions of b. and c. definitely were.

(36) Alluding to problems of language acquisition,Lasnik and Kupin (1977) gave a restrictive formalization of phrase structure and transformational grammar, and proposed that "Transformations are not marked optional or obligatory. The certainty of application of a transformation is decided by general principles to be described in the definition of derivation.." [p.181]

(37) Lasnik and Kupin went on to formulate a particular surface filter, with the consequence that "no particular T becomes obligatory, but rather it is obligatory that something be done somewhere along the line ... otherwise, the derivation must be 'thrown out'. This seems to be the proper generalization. What is obligatory is not the means used, but the end achieved." [p.188]

(38) This kind of optionality plus filtering eventually became central to GB theorizing.

(39) Chomsky and Lasnik (1977) developed this and related ideas much further, with explicit reference to langugage acquisition. Along with Lasnik and Kupin, they proposed a highly restrictive theory of transformations, with all transformations optional.

(40) C&L noted the problem raised for the optional/obligatory distinction by the absence of negative data. Filters, claimed to be parts of UG, then bear much of the burden of determining well-formed surface structures.

(41) "Our hypothesis, then, is that the consequences of ... obligatoriness ... can he captured in terms of sulface filters, something that surely need not be the case in principle; and further, that these properties can be expressed in a natural way at this level." [p.433]

(42) Baker (1979) explored the issue raised by Braine, and proposed some solutions for specific cases, but indicated that there is no general solution available.

(43) Dell (1981) did propose a general solution, in the context of a discussion about our main topic here - learnability of optional vs. obligatory rules. While Dell was concerned specifically with phonological rules, his reasoning carries over to transformations to the extent that there too optional vs obligatory might be a choice for the learner. Dell considered the situation where

> "there are two grammars G and G'such that the language generated by G is a proper subset of the one generated by G'.Call these *grammars (generating languages) in an inclusion relationship*; GIRs for short.G' will be said to be "more inclusive" than G. I am considering learning situations where G is the grammar to be acquired but where, given the present state of linguistic theory, the overinclusive grammar G' would seem to be an equally good candidate for selection by the learners, due to the absence of "negative" information from the primary linguistic data ..." [p.33]

(44) This is exactly the subset problem outlined above. To address the problem, Dell proposed that "the language acquisition device should be constructed so as to meet the following general requirement"

(45) "Whenever there are two competing grammars generating languages of which one is a proper subset of the other, the learning strategy of the child is to select the less inclusive one." [p.34]

(46) Dell then says about [(45)]:

> "I assume [(45)] to be part of Universal Grammar. [(45)] involves evaluating two competing grammars by comparing, not these grammars themselves, but the extension of the languages that these grammars generate." [p.34]

(47) And, directly relevant to our concerns, a corollary of [(45)]:

(48) "A particular case of the GIR situation arises when one compares a grammar G containing an obligatory rule with the grammar G' obtained by making the rule optional. G' is more inclusive than G. As a first implementation of the general requirement [(45)], then, I propose that the language acquisition device incorporate the following learning strategy:

> Until encountering evidence to the contrary, always assume that a rule is obligatory." [p.35]

(49) Dell's GIR is strikingly similar to the Subset Principle of Berwick (1982) and Berwick (1985) [page references will be to the latter].

(50) Presenting as a goal identifiability in the limit from positive evidence, Berwick says

> "Let us say that when a family of languages meets this condition that it obeys the Subset Principle (or that the family possesses the Subset Property)." [p.236]

And

(51) "In the special case where one target language is properly contained within

another, the point of this condition is to ensure that the acquisition procedure al ways guesses a subset language if possible, that is, the smallest language that is also compatible with the positive evidence so far encountered ."

(52) Like Dell's GIR, Berwick's Subset Principle concerns the ordering of learners' hypotheses. But there is an interesting difference.

(53) For Dell, the GIR is part of the language acquisition device. For Berwick, the Subset Principle could be part of the language acquisition device, but need not be:

> "lt should ... be stressed that the hypothesis ordering need not be done by the acquisition device itself; rather, it might have been fixed over time by an outside oracle, e.g., natural selection ." [p.237, fn.11]

(54) Berwick also notes that if the Subset Principle is, in fact, part of the language acquisition device, an apparent problem arises:

> "... it might appear as though it would be hard to check the application of this principle. For recall that in general the determination of whether $L_i \subset L_j$ is undecidable, for context-free languages and beyond." [p.237]

(55) (The same apparent problem would arise for Dell's GIR.) However, Berwick argues that the problem is only apparent:

> "In practice, the calculation is much easier. For example, suppose we limit the data that is used for acquisition to degree 2 , that is, only sentences with at most two embedded sentences. This is the constraint invoked by Wexler and Culicover in their learnability model Wexler and Culicover (1980). Now the subset computation is decidable."

II. A Case Study

(56) I believe <u>her</u> to have convinced Bill

(57) As is well known, in 'deep' respects, the underlined NP in (56) behaves like the subject of the lower predicate, while in 'surface' respects, including anaphora and morphological case, it behaves like the object of the matrix verb.

(58) The morphological case of the subject of the infinitive in English is an objective case most typically associated with a direct object. And, for English, there is good evidence that the matrix verb, for example <u>believe</u> in (56), is responsible for that objective case. Overwhelmingly, the English Accusative-Infinitive construction occurs only as the complement of an otherwise transitive verb which is independently capable of licensing case on its complement. When an English transitive verb is made passive, it loses that capability:

(59) I believe him

(60) *It is believed him
cf. He is believed

(61) It is believed that she convinced Bill

(62) The English Accusative-Infinitive construction patterns with (60) rather than with (61):

(63) *It is believed her to have convinced Bill

(64) For English, it was standardly assumed in early generative grammar that the downstairs thematic subject becomes an upstairs object via transformation(s).

(65) Chomsky (1955) had a process combining the higher and lower predicates and turning the underlying downstairs subject into the object of this complex predicate.

(66) Rosenbaum (1967), on the other hand, had the downstairs predicate extraposing out of the embedded infinitive, leaving the complement subject as the sole occupant of the complement, thus, in essence, making it the object of the matrix verb. Postal (1974) offered a technical variant of the same basic idea, but with the complement subject actually raising to matrix object position.

(67) "Three traditional arguments for higher object status" of the accusative subject in English [Postal 1974]. (These assume that the relations involved implicate clause-mates.)

(68) a. Jack believed Joan to be famous
b. Joan was believed to be famous by Jack

(69) a. *Jack$_i$ believed him$_i$ to be immoral
b. Jack$_i$ believed himself$_i$ to be immoral

(70) They believed each other to be honest

(71) A new alternative, Chomsky (1973): The relations in (68)-(70) don't demand clause-mates. Rather, they just require that the two related elements not be separated by a finite clause boundary (the Tensed Sentence Condition). I sometimes call these 'boundary strength arguments'.

(72) But there are other phenomena [Postal (1974), Lasnik and Saito (1991)] that indicate that the accusative subject is at least as high in the structure as elements of the matrix clause ('height arguments').

(73) ?The DA proved [the defendants to be guilty] during each other's trials

(74) ?The DA accused the defendants during each other's trials

(75) ?*The DA proved [that the defendants were guilty] during each other's trials

(76) ?The DA proved [none of the defendants to be guilty] during any of the trials

(77) The DA accused none of the defendants during any of the trials

(78) ?*The DA proved [that none of the defendants were guilty] during any of the trials

(79) The students solved three problems each

(80) *Three students each solved the problems (i.e., on the reading 'The problems were solved by three students each')

(81) *The students proved that three formulas each were theorems (i.e., on the reading 'Each of the students proved that three formulas were theorems')

(82) ?The students proved three formulas each to be theorems

(83) Jones proved the prisoners guilty with one accusation each

(84) Jones proved the defendants to be guilty with one accusation each

(85) Jones prosecuted the defendants with one accusation each

(86) ??Jones proved that the defendants were guilty with one accusation each

(87) Chomsky (1991) proposed raising, but not to object position, rather to Spec of an agreement projection above the VP, Agr_O,(where, he proposed, objects also raise) in the LF component (because of concerns about the word order; overt raising would incorrectly, for English, place the object or ECM subject to the left of the verb).

(88) But Lasnik and Saito argued that at least some of the phenomena they explored implicate overt raising. For instance, covert operations never seem to affect anaphoric binding possibilities, yet, as we have seen, the ECM subject can bind an item unequivocally in the matrix clause. Lasnik and Saito left this as a mystery.

Further

(89) Sometimes the raising is audible, as in these examples from Postal (1974):

(90) I figured out [it was more than 300 miles from here to Tulsa]

(91) I figured it out [*t* to be more than 300 miles from here to Tulsa]

(92) Or these from Kayne. ((Though not Kayne's analysis. Johnson (1991) and Koizumi (1993) did offer a raising account.))

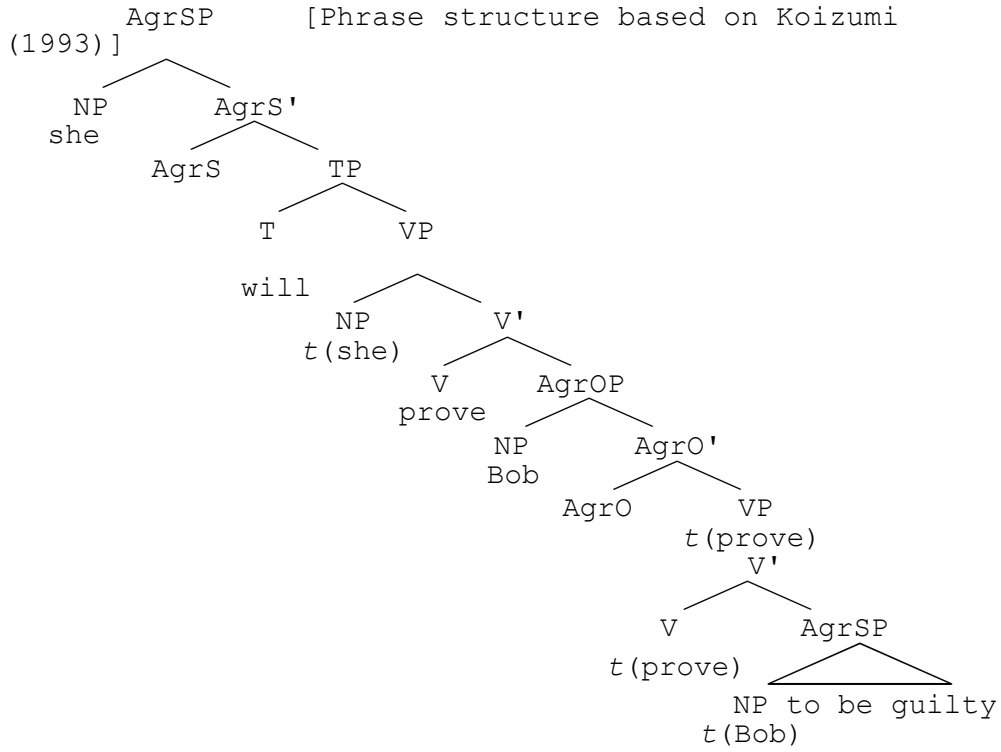
(93) They're trying to make John out to be a liar Kayne (1985, p.113), Johnson (1991)

(94) cf. They're trying to make out that John is a liar

(95) Koizumi proposed a way that the raising could be overt while still producing the correct word order, his 'split VP hypothesis', with the NP raising and the main V raising still higher.

(96) She will prove Bob to be guilty

(97)

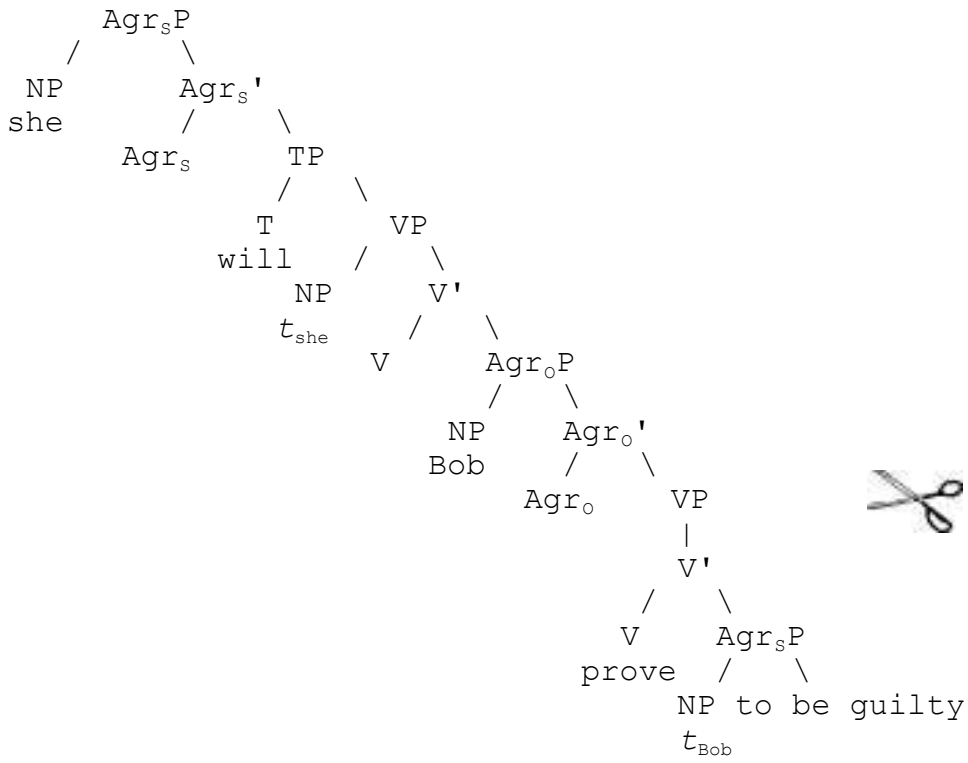


(98) An additional argument for overt raising of an ECM subject (or, for that matter, a matrix object); Pseudogapping as VP ellipsis (Jayaseelan (1990)), with the remnant having raised to Spec of Agr_O (Lasnik (1995)).

(99) Mary hired John, and Susan will ~~hire~~ Bill

(100) He proved Jones (to be) guilty and she will ~~prove~~ Bob ~~(to be) guilty~~

(101)



(102) One nice feature of a raising analysis (that of Chomsky (1991) or Koizumi (1993)) is that structural Case is always licensed in the same configuration: Spec of a functional head, a point made by Chomsky.

(103) BUT there is evidence that the raising into the higher clause does not invariably take place, **at least for some speakers**. That is, while for most speakers the raising is obliagtory, for others it is optional.

(104) %They're trying to make out John to be a liar Kayne (1985), Johnson (1991)
(105) They're trying to make John out to be a liar

(106) A curious surface constraint noted by Lasnik (1972) and further discussed by Postal (1974) and used by him to argue for raising, and for its obligatoriness, leads to the same conclusion:

(107) *Not*-initial NPs occur only in (derived) subject position. Postal (1974, p.95)
(108) Not many gorillas have learned to tap-dance
(109) ?*Joe kissed not many models
(110) Not many Albanians have been interviewed by Sevareid All from Postal (1974)

(111) %Harry proved not many of those formulas to be theorems
cf.
(112) Harry proved that not many of those formulas were theorems

(113) Postal used the claimed badness of (111) to argue for obligatory raising. However, for some speakers, the example isn't so bad, suggesting that the raising need not have taken place.

(114) The contrast emerges even more clearly in the *make out* infinitival construction:
(115) %They made out not many articles to have been published
(Needless to say, this one isn't good for those who don't get *make out* NP *to* ... word order in the first place.)
(116) *They made not many articles out to have been published (bad for all)

(117) An observation about scope that Zubizarreta (1982) attributes to Chomsky, and that is discussed again by Chomsky (1995), provides further evidence for the optionality of object shift with ECM subjects for some speakers:
(118) a. (it seems that) everyone isn't there yet
b. everyone seems [t not to be there yet]
(119) Chomsky (p.327) argues as follows: "Negation can have wide scope over the Q in [(118)a]... but not in [(118)b]", concluding that "...reconstruction in the A-chain does not take place, so it appears."
(120) Thus, a universal quantifier in subject position can be understood in the scope of clausal negation; but not if that quantifier has undergone raising.

(121) I believe everyone not to be there yet [Based on Chomsky (1995)]
(122) For some speakers, Chomsky among them, this can have 'everyone' with scope below 'not', just as in the situation of "I believe that everyone isn't there yet". But many speakers don't allow this interpretation.
(123) What happens when the word order indicates whether or not raising has taken place?
(124) The mathematician made every even number out not to be the sum of two primes
[Only has the crazy reading that the mathematician was pretending that no even number is the sum of two primes.]
(125) The mathematician made out every even number not to be the sum of two primes
[For speakers who get this word order in the first place, in addition to the crazy reading, also has the sane reading that the mathematician was pretending that Goldbach's conjecture is false.]

cf.
(126) The mathematician made out that every even number isn't the sum of two primes
Lasnik (1999), Lasnik (2001)

(127) Thus, we have more evidence that, at least for some speakers, the raising in ECM constructions is optional, while for others it is obligatory.

One final argument for optionality of raising:

(128) *John$_i$ injured him$_i$
(129) *John$_i$ believes him$_i$ to be a genius

(130) *Mary injured him$_i$ and John$_i$ did too

(131) %Mary believes him_i to be a genius and $John_i$ does too

(132) Bizarrely, a PF process, deletion, looks like it is repairing a Condition B violation in the ECM situation, at least for some speakers.

(133) Suppose Postal (1966), Postal (1974) was right (contra Chomsky (1973)) that the relevant structural configuration for such obviation is based on the notion clause-mate. (For related discussion, see Lasnik (2002) and Ausin (2001).)

(134) Weak pronouns must cliticize onto the verb, as proposed by Oehrle (1976)

(135) The detective brought him in

(136) *The detective brought in him Chomsky (1955)

(137) I gave it to Mary

(138) *I gave Mary it

(139) Suppose cliticization demands structural locality (in addition to linear locality). And suppose that in (129) him stays in the lower clause to evade a Condition B effect. The resulting failure to cliticize would cause a PF violation, but in (131) the failure is repaired by ellipsis, as the would-be clitic is gone.

(140) In (130), on the other hand, the pronoun and its antecedents are clause-mates all along, so deletion doesn't help. The cliticization requirement will invariably be satisfied but Condition B will invariably be violated.

(141) So how is Case licensed when the ECM subject doesn't raise? Possibly, as in LGB, by government across the infinitival clause boundary. Or, a more modern version of the idea, via long distance Agree.

(142) But there is a potentially nicer possibility.

(143) Davis (1984) rejected the classic LGB analysis, which was based on the idea that S, unlike $\bar{S}$, is not a maximal projection, so does not block government of the embedded subject by the matrix verb, allowing the latter to Case-mark the former. Thus, following "$\bar{S}$-deletion", a process triggered by a particular class of verbs, 'believe' governs and Case-marks 'John' in

(144) We believe [John to be intelligent]

(145) But, Davis noted, once clauses are fully incorporated into $\bar{X}$-theory, this becomes untenable: S is actually the maximal projection of Infl, so the needed government would not be available, and its failure to block government just a stipulation. As an alternative, Davis proposed that the Case feature of 'believe' percolates down to the head of its complement, non-finite Infl, thus providing the latter with accusative Case-marking ability.

(146) Two notes on this:

a. Modern Minimalist feature inheritance is remarkably reminiscent of Davis's approach

b. In the 'Accusative-Infinitive' construction of Latin and Classical Greek, non-finite Infl was evidently directly responsible for the accusative Case on its specifier. Under Davis's proposal, these languages are actually quite similar to English, the only difference being that in English, non-finite Infl needs to inherit the Case-assigning feature while in Latin and Classical Greek, infinitival Infl has that feature intrinsically.

(147) Davis's idea is of benefit with respect to another observation of Kayne's (stated in my terms): With small clause complements, raising is obligatory for all speakers.
(148) They're trying to make John out a liar Kayne (1985, p.108-109)
(149) *?They're trying to make out John a liar

(150) This correlates well with the 'not'-initial NP constraint:
(151) Mary believes that not every politician is a liar
(152) Mary believes not every politician to be a liar [OK for some speakers]
(153) *Mary believes not every politician a liar [* for all]

(154) Why should this be? If, as I have been assuming/arguing, the issue is Case, it would be strange if the subject of a small clause could not be governed by the matrix verb while the subject of an infinitive could be. If one would anticipate a difference at all, it would be that a small clause boundary is weaker than any full clause boundary.
(155) On a Davis type approach, one would just have to say that whatever the head of a small clause is, it's not the kind of thing that can have (or inherit) a Case feature.

(156) Now, what is the obligatory/optional distinction for ECM infinitivals? One reasonable possibility is based on (155). For the speakers having obligatory raisng (the majority of those I have asked), even ECM infinitival Infl can't inherit a Case feature. For the minority with optional raising, ECM infinitival Infl can inherity a Case feature. There's plenty nore to be said about this, but since my time is up, it will have to wait for another occasion.

References

Ausin, Adolfo. 2001. *On A-Movement*. Doctoral dissertation, University of Connecticut, Storrs, CT.

Baker, Carl Lee. 1979. Syntactic theory and the projection problem. *Linguistic Inquiry* 10: 533-581.

Berwick, Robert. 1982. *Locality principles and the acquisition of syntactic knowledge*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, Mass.

Berwick, Robert. 1985. *The Acquisition of Syntactic Knowledge*. Cambridge, Mass.: MIT Press.

Braine, Martin. 1971. On two types of models of the internalization of grammars. In *The ontogenesis of grammar: A theoretical symposium*, ed. D. I. Slobin. New York: Academic Press.

Chomsky, Noam. 1955. The logical structure of linguistic theory. Ms. .[Revised 1956 version published in part by Plenum, New York, 1975; University of Chicago Press, Chicago, 1985].

Chomsky, Noam. 1956. Three models for the description of language,. *IRE Transactions on Information Theory* IT-2: 113-124.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT press.

Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen Anderson and Paul Kiparsky, 232-286. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In *Principles*

and parameters in comparative grammar, ed. Robert Freidin, 417-454. Cambridge, Mass.: MIT Press. [Reprinted in Noam Chomsky, *The minimalist program*, 129-166. Cambridge, Mass.: MIT Press, 1995].

Chomsky, Noam. 1995. Categories and transformations. In *The minimalist program*, 219-394. Cambridge, Mass.: MIT Press.

Chomsky, Noam and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 11: 425-504. [Reprinted in Howard Lasnik, *Essays on restrictiveness and learnability*, 42-124. Dordrecht: Kluwer, 1990]

Davis, Lori. 1984. *Arguments and expletives*. Doctoral dissertation, University of Connecticut, Storrs.

Dell, Francois. 1981. On the learnability of optional phonological rules. *Linguistic Inquiry* 12: 31-37.

Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9: 577-636.

Katz, Jerrold J. and Paul Postal. 1964. *An integrated theory of linguistic descriptions*. Cambridge, Mass.: MIT Press.

Kayne, Richard. 1985. Principles of particle constructions. In *Grammatical representation*, ed. Jacqueline Guéron, Hans-Georg Obenauer, and Jean-Yves Pollock, 101-140. Dordrecht: Foris.

Koizumi, Masatoshi. 1993. Object agreement phrases and the split VP hypothesis. In *Papers on Case and Agreement I: MIT working papers in linguistics* 18, 99-148.

Lasnik, Howard. 1972. *Analyses of negation in English*. Doctoral dissertation, MIT, Cambridge, Mass.

Lasnik, Howard. 1995. A note on pseudogapping. In *Papers on minimalist syntax, MIT working papers in linguistics* 27, 143-163. [Reprinted in Howard Lasnik, *Minimalist analysis*, 151-174. Oxford: Blackwell, 1999]

Lasnik, Howard. 1999. Chains of arguments. In *Working minimalism*, ed. Samuel D. Epstein and Norbert Hornstein, 189-215. Cambridge, Mass.: MIT Press. [Reprinted in Howard Lasnik, *Minimalist investigations in linguistic theory*, 139-157. London: Routledge, 2003].

Lasnik, Howard. 2001. Subjects, objects, and the EPP. In *Objects and other subjects: Grammatical functions, functional categories, and configurationality*, ed. William D. Davies and Stanley Dubinsky, 103-121. Dordrecht: Kluwer Academic Publishers.

Lasnik, Howard. 2002. Clause-mate conditions revisited. *Glot International* 6: 94-96.

Lasnik, Howard and Joseph J. Kupin. 1977. A restrictive theory of transformational grammar. *Theoretical Linguistics* 4: 173-196. [Reprinted in Essays on restrictiveness and learnablity, Howard Lasnik, 17-41. Dordrecht: Kluwer, 1990]

Lasnik, Howard and Mamoru Saito. 1991. On the subject of infinitives. In *Papers from the 27th Regional Meeting of the Chicago Linguistic Society, Part I: The general session*, ed. Lise M. Dobrin, Lynn Nichols, and Rosa M. Rodriguez, 324-343. Chicago Linguistic Society, University of Chicago, Chicago, Ill. [Reprinted in Howard Lasnik, *Minimalist analysis*, 7-24. Oxford: Blackwell, 1999].

Oehrle, Richard. 1976. *The grammatical status of the English dative alternation*. Doctoral dissertation, MIT, Cambridge, Mass.

Postal, Paul M. 1966. A note on 'understood transitively'. *International Journal of American Linguistics* 32: 90-93.
Postal, Paul M. 1974. *On raising: One rule of English grammar and its theoretical implications*. Cambridge, Mass.: MIT Press.
Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. Cambridge, Mass.: MIT Press.
Wexler, Kenneth and Peter Culicover. 1980. *Formal principles of language acquisition*. Cambridge, Mass.: MIT Press.
Zubizarreta, Maria Luisa. 1982. *On the relationship of the lexicon to syntax*. Doctoral dissertation, MIT, Cambridge, Mass.