

## 5. Modal predicates and argument structure

The modal operators we have considered so far have been of two semantic types: (i) type  $\langle st, t \rangle$ , which we assumed for modal auxiliaries, main modal verbs like **need to**, **have to**, and adjectives like **likely**<sup>1,2</sup>; and (ii) type  $\langle st, et \rangle$ , the type of attitude verbs like **believe** and **want**. The syntactic properties of the lexical items of each of these semantic types are diverse:

- (175) for semantic type  $\langle st, et \rangle$ :
- (a) verbs (adjectives) with **that**-clause and **for**-clause complements:  
V, subcategorizes for CP with **that** or **for**  
**believe (that), say, prefer (for), ... (aware, eager, ...)**
  - (b) exceptional case-marking (ECM) verbs:  
V, subcategorizes for IP  
**believe, ...**
- (176) for semantic type  $\langle st, t \rangle$ :
- (c) modal auxiliaries:  
Infl, subcategorizes for VP  
**must, may, ...**
  - (d) raising verbs (and adjectives):  
V(or A), subcategorizes for IP  
**have (to), need (to), ... (likely (to), ...)**
  - (e) verbs (and adjectives) with expletive subjects:  
V (or A), subcategorizes for CP with **that** or **for**  
**seem, ... (possible, ...)**

In groups (d) and (e) we can also include the passives (without by-phrase) of verbs in (b) and (a). (E.g. **There are allowed to be cars on the beach; It is known that ...**). The syntactic concepts used in this classification are open to revision, of course. Also the list is not meant to be exhaustive.

What additional semantic types do we find in lexical items that create intensional contexts? If we consider verbs like **tell** (in **John told Mary that ...**), we presumably need a type with a second type-e argument, i.e.  $\langle st, \langle e, et \rangle \rangle$ . What about the type of the intensional argument; does this always have to be a proposition ( $\langle s, t \rangle$ )? In particular, could it also be  $\langle s, et \rangle$ , i.e. a 1-place

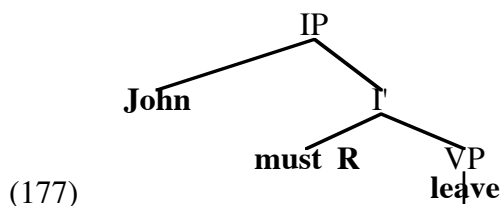
---

<sup>1</sup>In this section, we concentrate on modal operators that are syntactic heads, i.e. verbs (and adjectives). We will completely disregard adverbs (like 'possibly', 'necessarily').

<sup>2</sup>Actually, our type for the modals themselves is not  $\langle st, t \rangle$ , but rather  $\langle \langle s, st \rangle, \langle st, t \rangle \rangle$ .  $\langle st, t \rangle$  is really the type of the constituent which already includes the modal and its covert restrictor. In this section, we will consistently talk as if the covert restrictors were already included in the lexical items which require them. So strictly speaking, when we talk about the type of a modal verb, we really mean the type of the node dominating the verb and the R-variable.

property? What might be examples (or candidates for examples) of predicates whose extensions are of, say, types  $\langle\langle s, et \rangle, t \rangle$ ,  $\langle\langle s, et \rangle, et \rangle$ , or  $\langle\langle s, et \rangle, \langle e, et \rangle \rangle$ ? [hint: see H&K ch. 12]

If we were not assuming the VP-internal subject hypothesis, then our first guess about the semantic type of modal auxiliaries would probably have been  $\langle\langle s, et \rangle, et \rangle$ .<sup>3</sup> This would be the type to make them straightforwardly interpretable in syntactic structures like (177).



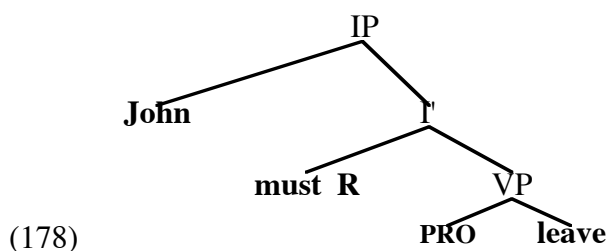
When we adopted the raising analysis of modal auxiliaries (and verbs like **have to**, **need to**), we did so without explicit motivation and discussion of alternatives. In this section, we are trying to fill this gap. We will compare the empirical predictions of the raising analysis (type  $\langle st, t \rangle$ ) with the predictions of at least one kind of non-raising analysis. We will find that the behavior of the modal auxiliaries and of many modal main verbs (including **have to** and **need to**) is indeed accounted for more successfully by the raising analysis than by that alternative. We will also ask whether any lexical items in natural languages behave as predicted by the alternative analysis, and we will conclude with some speculations about universal constraints on possible lexical items.

### 5.1 A type- $\langle\langle s, et \rangle, et \rangle$ analysis for modals?

Suppose we adopted a syntactic representation as in (177). Here the modal (including its restrictor) combines first with a complement whose intension is a 1-place property, and then with an external argument which denotes an individual. Therefore, the denotation of the modal+restrictor will have to be of type  $\langle et, et \rangle$  (if we are going to apply plain FA) or type  $\langle\langle s, et \rangle, et \rangle$  (if we are going to use IFA). Can we write suitable lexical entries with at least one of

---

<sup>3</sup> Even if we do stick to the VP-internal subject hypothesis, the raising analysis of modals that we have so far taken for granted is not our only option. The subject internal to the VP could be a PRO, as in (178), and it then depends on how we treat PRO.



If PRO is semantically vacuous as suggested in H&K, ch. 8, (178) is the same as (177) in its implications for the semantic type of **must-R**. If PRO is a pronoun that needs to be bound (hence treated as a variable of type  $e$ ), the modal's type would need to be  $\langle st, et \rangle$ . We will disregard these options for the moment, but see below.

these types? Here is a proposal.

(179) For any  $w \in W$ :

- (a)  $[[\text{must}]]^w = \lambda R_{\langle s, st \rangle}. \lambda P_{\langle s, et \rangle}. \lambda x_e. \forall w' [w' \in R(w) \rightarrow P(w')(x) = 1]$
- (b)  $[[\text{may}]]^w = \lambda R_{\langle s, st \rangle}. \lambda P_{\langle s, et \rangle}. \lambda x_e. \exists w' [w' \in R(w) \& P(w')(x) = 1]$

So for the modal+restrictor node, we will have (180). Compare this with (181), which is what we had for the same node on the raising analysis.

$$(180) \quad [[\text{must}]]^{w, [R \rightarrow R]} = \lambda P_{\langle s, et \rangle}. \lambda x_e. \forall w' [w' \in R(w) \rightarrow P(w')(x) = 1] \quad (= G)$$

$$(181) \quad [[\text{must}]]^{w, [R \rightarrow R]} = \lambda p_{\langle s, t \rangle}. \forall w' [w' \in R(w) \rightarrow p(w') = 1] \quad (= F)$$

The function defined in (181) – let's call it  $F$  – takes a proposition; the function defined in (180),  $G$ , instead takes first a property and then an individual. Apart from this difference, these two functions perform similar jobs, as we can see when we define one in terms of the other:

$$(182) \quad G = \lambda P_{\langle s, et \rangle}. \lambda x_e. F(\lambda w. P(w)(x))$$

(182) is a recipe for constructing  $G$ , given  $F$ . What (182) says intuitively is this: if you want to know what value  $G$  yields when applied to  $P$  and  $x$ , construct the proposition that  $x$  has property  $P$ , and then apply  $F$  to that proposition.

Given that the new type- $\langle s, et \rangle, et \rangle$  meaning of the modal is constructed in this way from the old type- $\langle st, t \rangle$  meaning, the predicted truth-conditions for the new structures in (177) will be systematically the same as for the old raising structures. In the raising analysis, **John** and **leave** formed a syntactic unit<sup>4</sup>, therefore the composition rule applying to this unit determined that we should calculate the proposition that John leaves. The modal then applied to that proposition. In the new analysis, **John** and **leave** are separate pieces, so we don't know from the composition rules alone that we should put them together into the proposition that John leaves. But the new lexical entry of the modal, which takes both of these pieces as arguments, has it written into it that we are supposed to construct this proposition. The division of labor between composition rules and lexical meanings is different, but the net result of the calculation is the same.

So far, our type- $\langle s, et \rangle, et \rangle$  analysis predicts the same truth-conditions as the old raising (type- $\langle st, t \rangle$ ) analysis. However, we will see now that this is no longer true when we move beyond examples with referential (e.g., proper name) subjects.

Suppose the subject is a quantificational DP, as in (183).

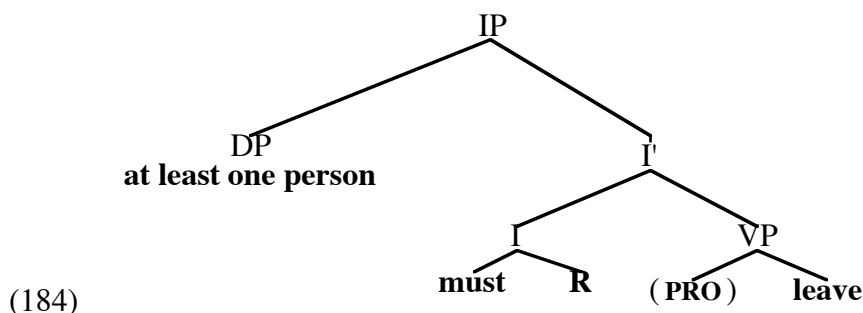
(183) **At least one person must leave.**

The syntactic structure at both S-structure and LF on the current analysis will have to be (184).<sup>5</sup>

---

<sup>4</sup>We are assuming here that we have raising followed by "reconstruction" (deletion of upper copy and coindexing). Without reconstruction, the relevant syntactic unit consists of **leave** and a variable. But after the variable is bound by Predicate Abstraction and the predicate abstract is predicated of **John**, the result still is the same.

<sup>5</sup> Ignore the parenthesized PRO. (See fn. 3)



This structure is interpretable by means of our new lexical entry in (181), and it expresses the *de re* reading. (Exercise: Verify this.) But what about the *de dicto* reading? Can we derive another LF which expresses that?

If reconstruction is the operation of deleting the upper copy (and coindexing link) of a movement chain, then of course, this operation is not applicable in a non-movement structure like (184). But the problem runs deeper than this. Even if we did allow some syntactic operation or other by which the DP **at least one person** got lowered past the modal (e.g., to substitute into the slot of the PRO, or to adjoin to the VP), that wouldn't help. It would not give us a representation of the *de dicto* reading, but only an uninterpretable structure. The sister node of **must-R** would now have a proposition as its intension (viz., the proposition  $\lambda w'. \exists x[x \text{ is a person in } w' \ \& \ x \text{ leaves in } w']$ ). But a proposition is not a suitable argument; as we see in (180), **must-R** needs a property. Nor would it help to allow a syntactic operation that would raise the modal above the subject. The obstacle to generating the *de dicto* reading is not in the syntax, but in the semantic type of the modal. As long as **must-R** has a denotation of type  $\langle\langle s, et \rangle, et \rangle$ , *de dicto* readings are predicted impossible.

It is fair to conclude, then, that our original raising analysis was superior. The existence of *de dicto* readings for subjects provides an empirical argument in its favor.<sup>6</sup>

---

<sup>6</sup> A note of caution: We have not shown that the raising analysis is superior to any non-raising analysis, only that it is superior to the particular non-raising analysis we considered – part of which was the assumption that the modal's outer argument is of type *e*. The situation changes if we consider other (fancier) choices of semantic type. Those authors who have actually defended non-raising analyses of modals in the literature have proposed instead that the modal's type is  $\langle\langle s, et \rangle, \langle\langle s, et \rangle, t \rangle, t \rangle$ . Where we had *e* in (179), this has the type of the intension of a generalized quantifier. (It is not clear exactly who should be credited with this idea. D. Dowty ('Governed transformations as lexical rules in a Montague Grammar,' *LJ* 1978) has an analogous entry for ECM ("raising-to-object") verbs, but his entry for raising-to-subject verbs has the simpler type in (179). Maybe the earliest published application to modal auxiliaries and raising-to-subject verbs is in E. Klein & I. Sag (1982) "Semantic Type and Control," in M. Barlow et al. (eds.) *Developments in Generalized Phrase Structure Grammar*, Bloomington: Indiana Linguistics Club. Anyway, the idea was widely familiar by the early 1980s. See e.g. the discussion in ch. IV, sec. 9, of Chierchia's thesis (G. Chierchia, *Topics in the Syntax and Semantics of Infinitives and Gerunds*, UMass Amherst Ph.D. thesis, GLSA 1984), and in D. Dowty (1985) 'On recent analyses of the semantics of control,' *L&P* 8.)

I leave it as an exercise to show that this fancier type accommodates *de dicto* readings without raising or reconstruction. A *prima facie* problem that arises in exchange is that only *de dicto* reading are generated. A follow-up exercise is to solve this problem. See the references cited above.

## 5.2 Reminder: standard tests for control and raising

In introductory syntax texts, the distinction between Raising predicates and Control predicates is often motivated by a battery of "tests":

	<i>raising</i>	<i>control</i>
<i>selectional restrictions</i>	imposed by lower verb only	imposed by both verbs
<i>passivization</i>	preserves truth-conditions	changes truth-conditions
<i>expletive subjects</i>	okay	impossible
<i>de dicto subjects</i>	possible	impossible

See e.g. Radford (1988).<sup>7</sup> Here are some examples to remind you of the data that the table above is intended to summarize.

Selectional restrictions:

- (192) (a) **John likes physics.**  
 (b) **#This number likes physics.**  
 (c) **This number is divisible by 5.**  
 (d) **#John is divisible by 5.**
- (193) (a) **John seems to like physics.**  
 (b) **#This number seems to like physics.**  
 (c) **This number seems to be divisible by 5.**  
 (d) **#John seems to be divisible by 5.**
- (194) (a) **John tries to like physics.**  
 (b) **#This number tries to like physics.**  
 (c) **#This number tries to be divisible by 5.**  
 (d) **#John tries to be divisible by 5.**

Passivization:

- (195) **Mary is likely to hire John.**  
**John is likely to be hired by Mary.**  
 (equivalent)

---

<sup>7</sup>To be more accurate, only the first three of these tests are found in Radford and other books on that level. The 'de dicto subjects' test is normally omitted, since it would be harder to explain to readers with no knowledge of semantics. But it is found in the original literature on which these textbooks draw (see for example May (1977)).

- (196) **Mary is eager to hire John.**  
**John is eager to be hired by Mary.**  
(not equivalent)

Expletives and idiom chunks:

- (197) (a) **There seem to be two people in the booth.**  
(b) **The cat seems to be out of the bag.**
- (198) (a) **\*There tried to be two people in the booth.**  
(b) **The cat tried to be out of the bag.** (only non-idiomatic)

De dicto readings:

- (199) **Several accomplices appeared to be involved.**  
(But in reality, there weren't any accomplices.)
- (200) **Several accomplices wanted to be involved.**  
(#But in reality, there weren't any accomplices.)

Exercise: How does the raising analysis that we have favored account for the behavior of raising predicates under the first three of these tests? (We have already answered the question for test number four.) Where necessary, state appropriate auxiliary assumptions about what it means to be an "expletive" or an "idiom chunk".

What about a non-raising analysis along the lines of (179) (adapted here to *seem*, *appear*, *be likely* etc., but crucially keeping the semantic type the same, i.e.,  $\langle\langle s, et \rangle, et \rangle$ )? We have already seen that this fails to capture the existence of *de dicto* readings. What about the behavior of raising verbs under the other three tests?

### 5.3 Raising, control, and constraints on the lexicon

We have seen that a non-raising analysis using type  $\langle\langle s, et \rangle, et \rangle$  is not empirically adequate for the modals and other raising verbs that we have looked at. It fails to account for *de dicto* readings and expletive subjects. So as far as accounting for the data of English (at least as far as we have looked at them) we can just forget about it. But there is a different and deeper question to consider. As linguists we are not merely concerned with finding the right analysis to predict the behavior of a given set of lexical items. We also aim to make predictions about the range of possible lexical items allowed by Universal Grammar. So, for example, we must look again at the already discarded type- $\langle\langle s, et \rangle, et \rangle$  entry for **must** and ask ourselves: Granted that this isn't the meaning of **must** – could it still be the meaning of some verb or adjective, in English or in some other natural language? Are there any words out there that do behave in precisely the way that **must** would have had to for us to maintain this entry for it? And if not, why not?

So suppose there were a verb **xxx** with just the lexical entry we entertained in section 5.1 for **must**, repeated here from (179):

$$(216) \quad [[\mathbf{xxx}]]^{w, [R \rightarrow R]} = \lambda P_{\langle s, et \rangle}. \lambda x_e. \forall w' [w' \in R(w) \rightarrow P(w')(x) = 1]$$

How would this verb behave in regard to the tests that are commonly employed to distinguish so-called "raising" verbs from so-called "control" verbs? We already saw in section 5.1 that it would never allow *de dicto* readings for its subject, and in this respect it would look like a "control" verb. What about the other tests?

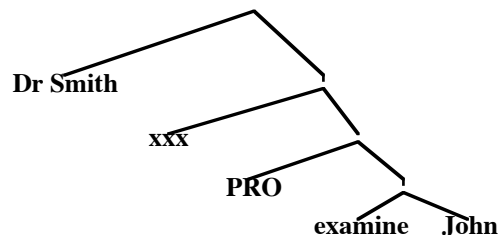
Another characteristic of "control" verbs is that active-passive pairs like the one in (217) are truth-conditionally non-equivalent (in contrast with analogous pairs with "raising" predicates like (218), which are always equivalent).

- (217) (a) Doctor Smith hopes to examine John.  
(b) John hopes to be examined by doctor Smith.

- (218) (a) Doctor Smith is likely to examine John.  
(b) John is likely to be examined by doctor Smith.

How will our hypothetical verb **xxx** behave in this regard? The two LFs for the active-passive pair will look as follows:<sup>8</sup>

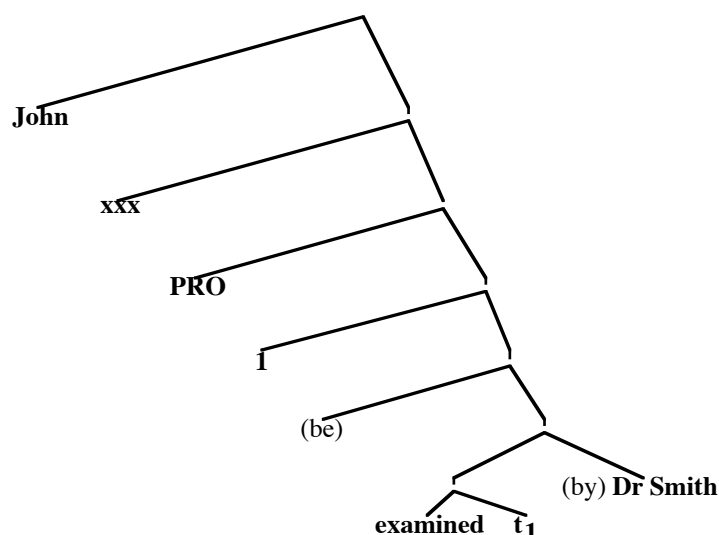
(219)



<sup>8</sup>We are assuming a movement analysis of the passive here, on which the passive morphology is semantically vacuous. An alternative would be a lexical analysis, which treats the passive morpheme as a relation-changing operation ( $\lambda f_{\langle e, et \rangle}. \lambda x. \lambda y. f(y)(x)$ ) and base-generates the two argument DPs in their surface hierarchical order. This would not make any difference to the discussion in this section.

Notice, by the way, that if we did not have a PRO-subject in the complement of **xxx** (but no subject at all, not even a semantically vacuous one), then the movement analysis would be unavailable to us here and we would be forced to choose the lexical analysis of passives. This is the reason why many authors conclude that a non-raising analysis of "raising" verbs is incompatible with a movement analysis of passive. The conclusion may hold on their assumptions, but not in the context of the H&K treatment of PRO (where PRO itself is vacuous, but capable of leaving a non-vacuous trace when moving).

(220)



Calculating the interpretation up to the step that depends on the lexical entry for **xxx**, we know that (219) is true in  $w$  iff  $[[\mathbf{xxx}]]^{w, [R \rightarrow R]}(\lambda w'. \lambda x. x \text{ examines John in } w')(Smith) = 1$ , and that (220) is true in  $w$  iff  $[[\mathbf{xxx}]]^{w, [R \rightarrow R]}(\lambda w'. \lambda x. Smith \text{ examines } x \text{ in } w')(John) = 1$ . Given the entry for **xxx** in (1), and given that, for all  $w'$ ,  $[\lambda x. x \text{ examines John in } w'](Smith) = [\lambda x. Smith \text{ examines } x \text{ in } w'](John)$ , it follows that  $[[ (219) ]]^{w, [R \rightarrow R]} = [[ (220) ]]^{w, [R \rightarrow R]}$ . So the truth-conditions of the two structures (in a given utterance context) are identical. In the passive test, then, our hypothetical verb **xxx** is predicted to act like a "raising" verb.

A caveat is in order here: The prediction that (219) and (220) are equivalent means only that their truth-conditions will coincide if they are evaluated w.r.t. to the same contextually supplied accessibility relation. If (219) were used in a context which supplies one accessibility relation  $R_1$ , and (220) in a context which supplies a different one,  $R_2$ , then of course they could well differ in truth-value. For example, it may happen that the active sentence (219) is uttered with the tacit understanding that we are quantifying over those possible worlds in which Dr Smith fulfills his duties, and that the passive sentence (220) is uttered with the understanding that we are talking about those worlds in which John fulfills his duties. Then these two utterances will express different propositions, and maybe one of them is true and the other false.

This implies that we will have to be very careful when we evaluate speaker's judgments as to whether a given verb does or doesn't pass the equivalence-under-passivization test. If there is (suspected) context-dependency of the accessibility relation, we'll have to make sure that people are really resolving it in the same way for the two sentences. If we just present the sentences out of context, this should not be taken for granted. People will then have to exercise their imagination to fill in a natural context, and in doing so they may well be influenced in a roundabout way by the very fact that one sentence is in the active voice and the other one in the passive. Passivization, even though it doesn't affect truth-conditions (in simple sentences), does have some effect on non-truth-conditional aspects of meaning or discourse-coherence. These effects are not well understood, but there is, speaking very roughly and vaguely, a *ceteris paribus* tendency for the surface subject to denote the topic of the discourse. This might indirectly affect the relative salience of different competing candidates for the accessibility relation. In other



words, if Dr Smith is referred to with the subject of the sentence, and therefore is perceived as the discourse topic, then the accessibility relation which is defined in terms of Dr Smith's duties is *ceteris paribus* more salient than the one defined in terms of John's duties. So a verb with the semantics defined in (216) may appear to fail the equivalence-under-passivization test, due to an indirect chain of connections between the discourse function of passivization and the strategies for the resolution of context-dependence. (And, for the same reason, even a verb with the simpler type <st,t> entry that is our official proposal for **must-R** may appear to fail it!)

However, even if this effect is quite strong and systematic, we can still tell the difference between a verb like **xxx**, which guarantees equivalence under passivization in a fixed context, and a verb whose meaning is defined in such a way that it "hard-wires" a distinctive contribution of the subject to the truth-conditions. The difference will be that we can, in principle, give independent contextual clues about the intended accessibility relation and thereby counteract and override the effect of the active/passive choice.<sup>9</sup>

Let us return now to the main thread of our discussion. What do we learn from our investigation of the predicted behavior of our hypothetical verb **xxx** w.r.t. to standard diagnostics for "control" and "raising"? We could continue to proceed through the whole list of tests, but already we have seen enough to conclude that **xxx** displays a mixed pattern of behavior w.r.t. these standard diagnostics: On at least one test (availability of *de dicto* readings), it acts "control"-like, and on at least one other one (equivalence under passivization), it acts "raising"-like. This is interesting, because it makes **xxx** a counterexample to an assumption which is widely shared (though often left implicit) in the syntactic literature, namely the assumption that the characteristics targeted by the standard tests "cluster together"; in other words, any given lexical item will either act "raising"-like on all test, or else act "control"-like on all tests.<sup>10</sup> Let us call this the "Clustering Assumption". More precisely, **xxx** is relevant to this assumption in the following way: First of all, the very fact that we were able to write down an explicit and coherent lexical entry like the one for **xxx** in (216) shows that the Clustering Assumption is not conceptually necessary, but rather constitutes a substantive empirical generalization. Given this, we now are motivated to look for real counterexamples in the vocabularies of the world's languages, and the hypothetical

---

<sup>9</sup>See Bhatt 1998 (Rajesh Bhatt, "Obligation and Possession," in H. Harley (ed.) MITWPL 32: *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*, pp. 21 - 40. [esp. section 4]) for related discussion. Giving proper attention to the complexities of strategies for the resolution of context-dependence may also help us make some sense of the inconclusive evidence presented by Jackendoff (1972; pp. 104f.). See also Brennan 1993 (Virginia Brennan, *Root and Epistemic Modal Auxiliary Verbs*. Ph.D. Dissertation, University of Massachusetts at Amherst. [Chapters 1 and 2 on the argument structure of various kinds of modal auxiliaries]).

<sup>10</sup>This is not meant to exclude the possibility that a given word may be ambiguous between raising and control meanings. (Many authors in fact have posited pervasive systematic ambiguity of this sort; see e.g. Brennan 1993 and references cited there for modal auxiliaries, and Jacobson 1992 and references cited there for main verbs like 'promise,' 'permit', etc.) The assumption stated says that such an ambiguous word will act consistently raising-like on one of its readings, and consistently control-like on the other reading. This is not the same as displaying control behavior as well as raising behavior on one given reading. (Though admittedly the difference will not be easy to detect in practice.)

example of **xxx** gives us some concrete idea of what we should be looking for. And if this search turns out to be unsuccessful (i.e., we don't find real counterexamples), then the example of **xxx** will give us some guidance in our endeavor to identify the principle(s) of Universal Grammar from which the Clustering Assumption follows.

Let us skip here over the survey of known data and the many non-trivial problems in the interpretation of that data, and let us proceed on the assumption that the Clustering Assumption is empirically correct. This means that **xxx**, as interpreted in (216) above, is not a possible word in natural language. Why not?

Two kinds of answers come to mind. One conceptually attractive hypothesis is that there are simply no lexical items of type  $\langle\langle s, et \rangle, et \rangle$  – perhaps more generally no lexical items with any arguments of types other than  $e$  or  $\langle s, t \rangle$ . This has been proposed (most recently, to my knowledge by Larson, den Dikken & Ludlow<sup>11</sup>), but rejected by many because it rules out a widely favored analysis of control verbs (most prominently defended in Chierchia 1984). Another answer is that, even though  $\langle\langle s, et \rangle, et \rangle$  is a possible lexical type, there is a relevant universal constraint against meanings of this type with a certain formal property ...

---

<sup>11</sup> see <http://semlab5.sbs.sunysb.edu/~rlarson/larson-papers.html>, listed under "in progress, 1997"