Review Article

The Cambridge Grammar of the English Language. By Rodney Huddleston and Geoffrey K. Pullum. Cambridge: Cambridge University Press, 2002. Pp. xvii, 1842.

Reviewed by Peter W. Culicover, The Ohio State University*

1. Introduction. The Cambridge Grammar of the English Language (CGEL) is a monumentally impressive piece of work.[1] Already published reviews of this work do not overstate its virtues[2]: 'a notable achievement'; 'authoritative, interesting, reasonably priced (for a book of this size), beautifully designed, well proofread, and enjoyable to handle'; 'superbly produced and designed'; 'one of the most superb works of academic scholarship ever to appear on the English linguistics scene… a monumental work that offers easily the most comprehensive and thought-provoking treatment of English grammar to date. Nothing rivals this work, with respect to breadth, depth and consistency of coverage'. I fully agree with these sentiments. Huddleston, Pullum and their collaborators definitely deserve a prize for this achievement.[3]

I try to convey here a sense of what it feels like to work with and through CGEL, and what one might plausibly conclude from this exercise about how language works. I also outline the theory of grammar that is explicit and implicit in CGEL and speculate a bit on what we might conclude from this theory about what it means to know a language. In particular, I explore the possibility that CGEL is actually the basis for a complete

description of the knowledge that a native speaker of English has of English, and the consequences of that possibility.

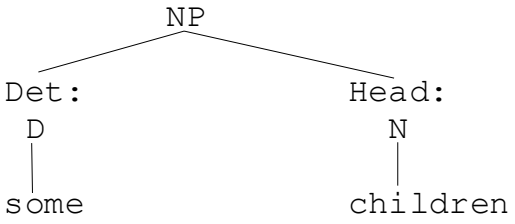
2. What is in CGEL. CGEL is organized into twenty chapters: Ch. 1 'Preliminaries', Ch. 2 'Syntactic overview', Ch. 3 'The verb', Ch. 4 'The clause: complements', Ch. 5 'Nouns and noun phrases', Ch. 6 'Adjectives and adverbs', Ch. 7 'Prepositions and preposition phrases', Ch. 8 'The clause: adjuncts', Ch. 9 'Negation', Ch. 10 'Clause type and illocutionary force', Ch. 11 'Content clauses and reported speech', Ch. 12 'Relative constructions and unbounded dependencies', Ch. 13 'Comparative constructions', Ch. 14 'Non-finite and verbless clauses', Ch. 15 'Coordination and supplementation', Ch. 16 'Information packaging', Ch. 17 'Deixis and anaphora', Ch. 18 'Inflectional morphology and related matters', Ch. 19 'Lexical word-formation', Ch. 20 'Punctuation'.[4]

CGEL is similar in its organization to the work that it aims to supplant, Quirk, et al. 1985 (Q85). Although the two are of roughly the same length (Q85 has 1789 pages and CGEL 1859), Q85 feels almost superficial compared with CGEL. The level of detail of CGEL is such that the reader may begin to feel that s/he is being told everything that one could possibly know about the topics that it covers. And while common sense tells us that this cannot be, in many cases it is difficult to think of what else one would want to say in factual terms about a particular construction, or form. (Theoretical excursions are something else entirely, of course.) There were only a few points here and there that I felt could have been mentioned but were not; further research almost invariably showed that they were in fact covered somewhere in the text.

Ch.1 introduces a number of critical points that guide the approach throughout. There is a basic introduction to the concepts of constituent structure and syntactic (lexical

and phrasal) categories. The typical **clause** is composed of a **noun phrase** followed by a **verb phrase.** Crucially, each constituent of a phrase has not only a category, but a grammatical function. These are represented simultaneously on a phrase marker, as in (1) (26:[13]).

(1)



Here, some is of the syntactic category D(eterminative), and bears the grammatical function Determiner in the phrase, while children is of the category N(oun), and bears the grammatical function Head of the phrase. Other functions include Subject, Predicate, and Object.

In the view of CGEL, syntactic categories are determined strictly by formal and distributional criteria; function is completely orthogonal.[5] The distinction is of course not novel, and is even found in some theoretical work (cf. Specifier and Complement in Chomsky 1972), but CGEL observes it rigorously, in the easy cases and in the hard cases. Sometimes the results are familiar, sometimes they are quite novel. For example, pre-head adjectives are modifiers (happy dog), pre-head nouns are modifiers (biology syllabus), and they are of different syntactic categories (537). The, a, this, that, some, etc. are determinatives, not adjectives, on the basis of their distributional characteristics; so are many, few, much and little (539). In a sleeping child, sleeping is a verb since it

cannot function as a predicative adjective, in contrast to disturbing in some disturbing news. Similarly for pre-head heard and worried (541). Clausal complements are not NP objects. The reasons: (i) V PP S is canonical, V PP NP is not, (ii) Some verbs take only S, not NP: marvel, vouch, wonder, charge, (iii) V P NP is grammatical, while V P S is not.

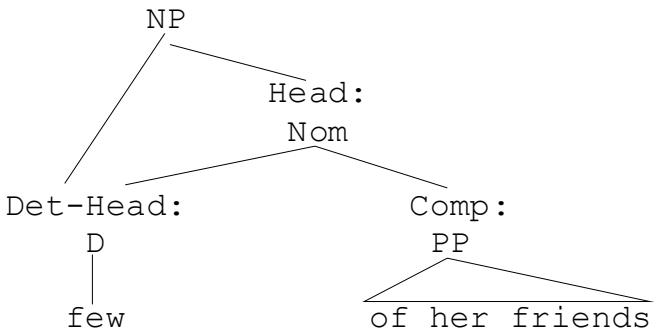
Pursuing the logic of categorization, CGEL argues that the prepositions that head phrases like before I got home are just that, prepositions, and not traditional 'subordinating conjunctions'. They argue that just as remember is a verb regardless of whether it takes an NP complement or a clause complement, so after is a preposition (600). Then, given that prepositions take such complements, and are not nouns or verbs, and are the heads of phrases that function as adjuncts, CGEL arrives at the conclusion that there are many prepositions besides the familiar before, after, in, to, at, on and so on. There are intransitive prepositions, such as downstairs, prepositions derived from adjectives, such as opposite, ahead, contrary, and prepositions derived from verbs, such as owing (to), barring, counting, including, excluding, given, concerning, provided, etc. (606-610).

And while one might not initially be inclined to say that barring, as in barring accidents for example, is a preposition, it is hard to argue with the distributional facts. It does not have the full range of forms of a verb and lacks control (*Having barred accidents, we would have succeeded), it certainly heads an adjunct that alternates with PPs, it is in the same head position as a prototypical preposition. Granted, the prototypical preposition typically has a thematic function, but that could be taken to show

that is that there are at least several semantically differentiated subclasses of prepositions, some thematic and some not.

A consequence of strict application of the form-function distinction is that a single form may have more than one grammatical function at the same time. Such a situation is what CGEL refers to as 'fusion'. For instance, in few of her friends, few bears the Head function as well as the Determiner function. This is shown in ((2)) (412:[7a]).

(2)



Similar analyses are given for such expressions as someone (Det-Head, fusing some one), (the) second (Mod-Head, fusing second one), what (I said) (Head-Prenucleus, fusing thing which), and the rich (Adjective-Head, fusing rich folk).

Fusion is a clean but not very deep solution to the problem of how to analyze these constructions. It neatly sidesteps the question of whether there is deletion (the rich folk → the rich ~~folk~~) , empty proforms (the second *pro*), movement and substitution (for free relatives – I won't try to give a blurb for a derivation here). Moreover, it takes the forms

to be sui generis, which avoids the problem of explaining why they don't always mean what they would mean if the derivation did not occur.

In addition to syntax, there is a lot of informal semantics in CGEL. Sometimes the presentation is simple and elegant; for example, the rule for the interpretation of the (368) is that the speaker expects the hearer to be able to identify the referent. Similarly, the rule for indefinite a (371) is that '[t]he addressee is not expected to be able to identify anything.' (The complexity is then presumably in defining under what circumstances one can reasonably hold these expectations.) At times the presentation is complex and detailed. Although space is limited, I must cite one representative passage because without an example it is difficult to appreciate just how much detail there is.

> The perfective/imperfective contrast is particularly important in the present tense because of the constraint that a present time perfective interpretation is normally possible only when the situation is of short enough duration to be co-extensive with the utterance:
>
> [4] i His daughter mows the lawn. [salient reading: serial state]
>
> ii His daughter is mowing the lawn. [salient reading: single occurrence]
>
> Mowing the lawn does not satisfy that condition, so that a single occurrence reading is not normally available for [i], which we interpret as a serial state, with habitual lawn- mowing. The imperfective meaning in [ii], by contrast, allows for T_d to be **included** within T_{sit}, giving the interpretation where a single occurrence of

mowing is now in progress. In the present tense, therefore, the progressive is much the more frequent aspect for dynamic situations. It would, however, be a mistake to see 'habitual' vs 'non-habitual' as the difference in meaning between [i] and [ii] (or, worse, between the present non-progressive and the present progressive generally). A single occurrence interpretation of [i] is not semantically excluded, but merely pragmatically unlikely: it could occur as a timeless or historic present or as a futurate - and if embedded, for example in a conditional construction, it could easily take a single future occurrence interpretation. Nor does [ii] exclude a serial state interpretation: compare His daughter is mowing the lawn until he is well again. The 'habitual' vs 'non-habitual' contrast is thus a difference in salient interpretations arising from the interaction between the meaning of the aspects and the pragmatic constraints on present perfectivity. Note that in the preterite the non-progressive His daughter mowed the lawn allows a single occurrence reading as readily as the progressive His daughter was mowing the lawn. (164)

Much of the discussion of the semantics of time, aspect, negation, modality, and quantity is on this order of detail, as is that of verbal semantics. Remarkably, I find myself constantly in agreement with the distinctions drawn, down to the very finest points. I say 'remarkably', because the question of how CGEL and I (and, I would presume, virtually all other native speakers of English) managed to come to the same judgments about what things mean down to the finest details is a non-trivial one. I come back to it in §4.

There are many other interesting proposals in CGEL, many of which turn on the basis for categorization. For instance, by the logic of CGEL, do is an auxiliary verb. It is not a modal, because it can be used with use to, while the modals like will and can cannot be (105), and it shows agreement while the modals do not. But it shares other properties with modals, for example, it cannot appear in the infinitival form, and it takes a bare infinitival complement, unlike have and be. In this case, and in fact in many other cases, we might want to ask, what follows from the categorization? If α and β share all relevant properties (like dog and cat), then it is reasonable to conclude that they are members of the same category. Suppose that they differ on one property, like have and be. While for some speakers have may function as a main verb or an auxiliary verb when it bears main verb function (I don't have a clue, I haven't a clue, as well as I haven't got a clue), be typically does not function as a main verb (*I don't be hungry, I'm not hungry). Are we then to conclude that have and be are not both auxiliary verbs? Suppose that there are two areas of difference, or more. At what point does the categorization break down?

This is not just a question about what to call things in a descriptive grammar. It is a question about what constitutes a category, and what it means for α and β to be members of the same category. I argued in Culicover 1999 that the categories are not rigid, and that they are defined precisely by the possible clusters of properties that heads, in particular, may possess. In general, CGEL appears to adhere to this view, although not rigorously.

3. What's not in CGEL? At least as important as what is in CGEL is what is not in CGEL. As comprehensive as it is, CGEL does not purport to be an in-depth exploration

of English syntax from the perspective of contemporary syntactic theory. However, in contrast to every other reference grammar of English, it is extraordinarily conscientious about many of the most influential discoveries of the past half century. It contains mention of the basic constraints on movement, i.e. Ross' Complex NP Constraint, the Left Branch Condition, and the that-trace effect (Ross 1967). While it does not relentlessly pursue the question of what else wh-movement and topicalization cannot do, there is coverage (248) of the impossibility of moving an indirect object to the left (by wh-movement) or the right (by Heavy shift), and numerous other illustrations of particular locutions that are awkward or 'hardly possible'. There is discussion of multiple extraction (1091), nested dependencies and crossing constraints (1094-5), reflexives in picture NPs (1486) (but not extraction from picture and non-picture NPs as far as I can tell), and extraction from subjects (1093). At the same time, it must be said (not critically) that there is much syntactic theory that is not touched on by CGEL. And most of the more technical questions that have been raised in the literature, with reams of associated examples and grammaticality judgments, are not addressed.[6]

It is reasonable to conclude that limitations on the theoretical scope of the work is an inevitable consequence of space limitations. After all, even at 1859 pages, the CGEL cannot tell us about everything, especially everything that cannot be said. But this is only part of the story. A position that we might attribute to CGEL is that in the ideal case, after saying what is possible, the impossible goes without saying. Such a view assumes that a language learner constructs a grammar based on its primary linguistic experience, within the general constraints of complexity imposed by the Language Faculty; then what is not encountered and falls outside of a well-supported generalization is outside of the

grammar. The reader may recognize this perspective as essentially that of Chomsky's Aspects, a point that I pursue in more detail in §4. I think that is the best way to understand why CGEL works as well as it does in describing what a native speaker of English knows about the language.

CGEL does not cite particular syntactic theories or syntactic analyses that have appeared in the theoretical literature, although it does argue against plausible alternatives to its own analyses. For instance, it argues (955) that that is not the head of a subordinate clause on the grounds that (i) that can be omitted (implicitly ruling out the possibility that there is an empty variant of that), (ii) what appears to be dependent in a subordinate clause is the finite/subjunctive marking (implicitly ruling out the possibility that there are two that's, governing different inflection), and (iii) an adjunct of the subordinate clause can precede that. 'The fact that it can precede indicates that the that is construed as part of the subordinate clause itself, not as a head element taking the subordinate clause as its complement' (956). Such superficial argumentation, while it may be well-founded, cannot do justice to the complexity of the issue, nor to the efforts that have been invested in it by others. Moreover, it does not offer the (non-specialist) reader a reasonable way of exploring the issue further.

This case is not unique: for example, there are arguments that negative auxiliaries are inflections, not contractions (91); subject-auxiliary inversion has a 'gap' (i.e. a trace) in canonical position (97); the tense that contrasts with the preterite is present tense, as opposed to non-past tense (134); the subject is a complement of the verb (215); V-P combinations such as refer to are not constituents; expressions such as [a number [of people]] are bracketed as shown, and not as [[a number of][people]] (351-2)*;* determiners

are dependents, not heads (357-8), hence NPs are NPs, not DPs; in all three proposals, all is the determiner (434), there is no system of voice in the NP (476); complement clauses are not NPs and not objects (1018-22); and many others. In virtually every instance CGEL is constrained by limitations of space to a brief mention of some of the more salient considerations. Given this, it would have been better, it seems to me, to have pointed the interested reader to some of the relevant literature.

This point can be extended to many of the factual observations in CGEL; the expert reader will no doubt often experience with a shock of recognition some data in CGEL that was first pointed out by, say, Chomsky[7] in 1977, or Klima in 1964, or whatever.[8] While the particular citations are no doubt of limited interest to the reader who is consulting CGEL as a reference grammar (for that is what it is), and not as a treatise on syntactic theory, it would be unfortunate if readers formed the view that CGEL is the first place where such data is cited, the disclaimers in the preface and the section on Further Reading notwithstanding.

4. The theory of CGEL. As the foregoing suggests, CGEL clearly stakes out a myriad of positions regarding the acceptability of sentences, the interpretation of particular expressions, the categorization of elements, and the phrase structure analysis of complex expressions. I find the judgments and interpretations to be right on the mark in the overwhelming majority (well over 99%) of the cases. I also find myself in sympathy with many of the particular categorizations and analyses (but not as overwhelming a majority). But underlying all of this, and to my mind even more important than the

individual details, is the fundamental theoretical posture of CGEL, some of which is explicit and some of which is not.

4.1. Concreteness and minimalism. As I have suggested elsewhere (Culicover 1999; Culicover and Nowak 2003), the most radical form of minimalism in linguistic theory is one that starts from the premise that other things being equal, if you can't 'see' it, i.e. if it isn't concrete, it isn't there. What is concrete in the case of language is meaning and sound. Syntactic structure, while abstract, is clearly required, as is phonological structure and conceptual structure. Structure that does not correspond to sound is more abstract, and more problematic.

This view, which I have called 'Concrete Minimalism', requires that strong empirical justification has to be provided for abstract entities and structure. It is founded on the hypothesis that a language learner is capable of forming sound-meaning correspondences (in the sense of Jackendoff 1990 and elsewhere) at any level of generality on the basis of concrete experience. Some correspondences pertain to individual words, some to classes of words, some to particular complex expressions (idioms and 'fossils' in the sense of CGEL), some to syntactic configurations with relatively complex specifications (i.e. constructions), and some to syntactic configurations with relatively simple and general specifications (i.e. constructions such as wh-questions and topicalization).

I believe that the central question is to what extent it is possible for a learner to acquire all of these correspondences without positing the existence of abstract entities (and derivations). If abstract entities are required, then so be it. If they are not, then if

they are to be justified, the justification must come from elsewhere, and must be evaluated on grounds other than factual correctness.

Culicover and Jackendoff 2005 argue that in fact the primary motivation for just about every abstract entity in modern Mainstream Generative Grammar (MGG) is the drive to maximize uniformity. It is important to understand this drive for uniformity in order to fully appreciate the significance of CGEL. There are several kinds of uniformity.[9] There is Derivational Uniformity, in which the fact that different sentences have some of the same properties is attributed to the hypothesis that they share some derivational operations. The best example is the analysis of long distance dependencies in Chomsky 1977. In this analysis, the fact that wh-questions, topicalization, clefts, relative clauses and comparatives all share certain properties, in particular sensitivity to island constraints, is attributed to the assumption that they are all derived through the application of wh-movement.

This example illustrates the fact that maximization of uniformity in one area does not come for free. If we want to say that there is wh-movement in a relative clause such as the people that you were telling me about ___, then we have to posit some abstract relative operator that undergoes the movement, i.e. the people OP_i that you were telling me about t_i.[10]

A second type of uniformity is Interface Uniformity. This is where we posit the same structure when we have the same interpretation. A classical instance is the assumption that there is an empty but fully structured VP in cases of ellipsis; for example, George believed every word that the professor said and John did ~~believe every word that the professor said~~ too. Interface Uniformity can be carried to an extreme position, as in

Merchant 2001, who argues that the structure of a sentence like Ted said that he was looking for someone who can translate Syntactic Structures into some obscure Caucasian language, but I can't for the life of me remember which has the analysis Ted said that he was looking for someone who can translate Syntactic Structures into some obscure Caucasian language, but I can't for the life of me remember which ~~obscure Caucasian language Ted said that he was looking for someone who can translate Syntactic Structures into which obscure Caucasian language~~. (Single strikeout indicates deletion; double strikeout indicates the phonetically empty copy (or trace) of movement.) Such an analysis is of course not without precedent (see Ross 1969) which attests to the venerable character of this type of uniformity.[11] It goes back at least as far as Katz and Postal 1964.

A third type of uniformity is Structural Uniformity. This is where we posit the same structure for the same grammatical function. E.g., if something is understood as a subject, then there must be a syntactic subject that represents it, as in the case of PRO in control constructions; the analysis of Sharon expected to fix the computer has the analysis Sharon expected [$_{IP}$ PRO to fix the computer] because expect takes a sentential complement in some cases, and the interpretation of the infinitive is the same as that of a sentential complement (Interface Uniformity) and all sentences must have subjects. More generally, Structural Uniformity leads to the hypothesis that all phrases of all categories of all languages fall under the X′-Schema (Jackendoff 1977).

Invocation of uniformity considerations is typically justified by the belief that in doing so we are simply following the guidelines of good science, applying Occam's Razor in the interest of eliminating unnecessary entities in the service of explanation.[12] But this is only half the story. If the phenomenon is of sufficient complexity, imposing

uniformity in one area will require the multiplication of entities in another. We can see this in the examples given above: empty operators for relative clauses, empty VP structure for ellipsis, and PRO for control. Occam's Razor works when there is a net savings. And we cannot (or should not) use the rhetorical trick of justifying our postulation of one type of abstract entity by explaining that it is not a net cost because it is a special case of an abstract entity that we have already postulated on the basis of an earlier application of the Razor.

No, the issue comes down to whether there are abstract entities or not. And the answer given by CGEL appears to be that in order to describe the entirety of a native speaker's knowledge of English there is no need to resort to abstract entities. Period. Granted, one might argue that CGEL's treatment of movement posits traces, since there is explicit reference to the gap corresponding to the canonical position of the moved constituent. But given that the canonical position of the trace is independently given by the phrase structure, and consequently is concrete in some sense, it would be straightforward to justify traces in terms of chains, i.e. relations between structural positions. Beyond this, there is no empty structure.

4.2. What is this thing called C_{HL} (Or, Yes, but is it interesting)? At this point it is not uncommon for theoreticians to say something like, 'Well, it may be true, as you say, that there's nothing more to be said about the computational system for human language (sometimes abbreviated as C_{HL}) beyond the description given by CGEL. But if that is the case, then it's simply not interesting, and if it turns out that language is not interesting, then I (for one) would prefer to be studying something else.'[13]

Accepting for the moment that what CGEL says about some phenomenon is all that there is to say, it is somewhat puzzling that as scientists we would have a serious notion of what would be more interesting than the truth. For instance, it would definitely be more interesting to discover that the moon is made almost entirely of green cheese than that it is made of rock and dust, especially given that it looks like it is made of rocks and dust, and the samples that have been brought back are – rocks and dust! It would be more interesting to learn that pigs cannot fly because their wings are made of an invisible substance that is too insubstantial to support their weight, rather than that they simply lack the anatomical and physiological wherewithal in the first place. It would be far more interesting to discover that chimps appear to lack human language because their religious beliefs prohibit the expression of personal thoughts (as opposed to feelings) to other creatures, rather than whatever the true answer is, which is probably some mundane story about neural organization, computational capacity, conceptual structure, and the like. But granting that the less interesting explanations are the right ones, scientists do not give up the good fight and turn to other pursuits. Why should linguists?[14]

I contend that what CGEL tells us is interesting, even if it more or less says what everyone thought was going on before MGG came along. Although CGEL appears to be a descriptive grammar of the traditional sort, there are vast differences in detail, and in this case the devil is in the details.[15] The problem for the theorist is to explain how in the world someone could have all of this stuff in his or her head. This is, to my mind at least, a vastly interesting question, and it becomes more interesting, the more detail there is that has to be gotten into the head. For it is clear that most of this knowledge cannot, I repeat, cannot be a priori. That is, it cannot arise from a setting of parameters, where the

learner, given some exemplars from the language, settles quickly on a generalization that accounts for all, or virtually all, subsequent exemplars from the language.

Parameter setting is in principle easy. Is the language VO or OV? Assuming that the language is consistent, just look at one instance, and there's your answer.[16] But which adjectives go with which verbs in the class that contains become, get, come, fall, come, go, turn and grow? Let's look at a chart that summarizes the possibilities.

	become	get	fall	come	go	turn	grow
asleep			✓				
awake	✓			✓			
aware	✓						✓
dead		(✓)	✓		✓		✓
alive	✓			✓			
nasty	✓	✓				✓	✓
nice	✓	✓				✓	✓
tall		✓					✓
short		✓	✓				✓
happy	✓	✓					✓
sad	✓	✓	✓			✓	✓
sharp	✓	✓				✓	✓
flat	✓	✓	✓		✓	✓	✓
smooth	✓	✓				✓	✓
sick	✓	✓	✓				✓

true	✓			✓			
mad	✓	✓			✓	✓	✓
nuts	✓	✓			✓		(✓)
silent	✓	✓	✓		✓	✓	✓
loud	✓	✓					✓
loose	✓	✓		✓			
open		✓	✓	✓			
closed		✓	✓				
red	✓	✓				✓	✓
sour	✓	✓			✓	✓	✓

Table 1, V+Resultative possibilities

This chart shows that it is possible to say things like become sad, get mad, fall sleep, go flat, turn nasty, grow silent, but not *become asleep, *get true, *fall alive, *come nasty, *turn sick, *grow open. CGEL gives some representative cases. I agree with all of its judgments, and in discussing these cases with other native speakers of English, have found that there is very close agreement about what is possible and what is not.

This is an extremely interesting fact. How did Huddleston, Pullum, me and all of the other speakers I consulted (and those that H&P consulted, and those that I would consult if I had the time, and so on) all arrive at the same judgments? Clearly we didn't learn English from the same people. Two answers come to mind immediately. The first is that the pattern in Table 1 arises from parameter setting: we hear a few instances of V+Resultative, and the whole thing falls into place. The second is that we learn this

pattern on the basis of actual experience with each of the possible pairs, or at least a large enough subset of the possible pairs that we can confidently generalize to new cases that are semantically close. By this I mean that if I experience turn red, I am going to be fairly confident that turn blue is going to be possible, and I might even say it if I have not heard it. But I will not say *fall blue on the basis of turn red or even turn blue. If I hear fall sad or fall silent I might be tempted to try fall happy or fall loud, but on the other hand I might notice that other cases involve for the most part adjectives that convey something negative, and then I won't be so tempted.

Sure, these look like simple selectional restrictions. One might argue that they are very refined selectional restrictions, but that's all they are. Some are semantically based, and some are just arbitrary. We have to learn selectional restrictions as part of learning the lexicon. That is, we have to learn, for every phrase [α β], where α is the head, which properties of β go with which properties of α. So on this view Table 1 is not a problem.

If we step back a little, we can see that there are several ways in which there can be restrictions of type illustrated here. In one case, the relevant properties are the syntactic categories of α and β, and their linear position relative to one another. In another case, the properties may have to do with inflectional classes, e.g. α may select a non-finite complement, and so on. And of course there are many other possibilities falling under the classical typology of c-selection and s-selection.

One might then argue that characterizing matters in this way loses sight of the fact that certain restrictions are not specific to individual lexical items, but hold for categories or even classes of categories. For example, in English the verb precedes the direct object, regardless of what the verb or the head of the direct object are. To allow for lexical

variance in such a domain would be to open up the possibility that languages exist in which the basic within phrase ordering relations have to be learned α-β pair by α-β pair. And surely no such languages exist.

The response at this point is that, first of all, special cases of this bizarre state of affairs do occur, and second of all, that if they are rare and virtually impossible, they should be ruled out not by stipulation but by principled explanatory mechanisms. And the explanation, in this case as in many others, is that the rare and virtually impossible situation is so highly marked that no plausible learning procedure would settle on it. That is, exceptions are costly, and generalizations are highly preferred. I return to this shortly.

In fact CGEL gives us plenty of evidence that the bizarre is possible, although rare. Consider the case of enough (445).

(3) a. soldiers enough

b. enough soldiers

c. tall enough

d. *enough tall

e. *soldiers quite enough

f. quite enough soldiers

g. *tall quite enough

h. quite tall enough

Enough can follow and precede N (3a,b), and only follow Adj (3c,d). When it follows, it cannot be modified (3e,g), but when it precedes, it can be (3f). And, strangely, when we

have quite Adj enough, quite modifies enough. This situation is by no means typical, but it is possible, therefore it must be learnable, and it should not be ruled out by a theory.

Note, in this connection, that Borer 1984 proposed that all parameters are stated over the lexicon. At first glance the current proposal seems to be simply a restatement of her original idea, and in a way it is. The difference come down to what counts as a parameter, and whether abstract 'functional' elements need to be posited. For Borer, the parameters in fact are restricted to just the functional elements; for us, and the theory that I am reading into CGEL, the parameters are in principle unbounded in their specificity, and they apply only to concrete lexical items, not to abstract functional elements.

4.3. Aspects was right. The mention of markedness in the preceding section brings us to a consideration of the syntactic theory of Aspects (Chomsky 1965), which is extremely pertinent for several reasons. One is that in reading CGEL one might be struck by how quaint its syntactic theory is. In many respects it is an informal and simplified version of the Aspects theory. There are syntactic categories, phrase structure rules, subcategorization and selection restrictions, and a few obvious transformations (like wh-movement, topicalization, and subject-aux inversion). The two works part ways in the considerably more expressive power that Aspects lent to transformations. Aspects keeps in the forefront the powerful methodology of Interface Uniformity that saw its clearest statement in Katz and Postal 1964; this methodology leads the theorist to posit transformational relationships of various degrees of complexity to account for cooccurrence restrictions shared by different constructions (like active and passive,

raising and control). CGEL treats them as distinct constructions related by shared interpretations.

Aspects introduces the idea that the set of possible languages is far greater than the set of languages that can be feasibly attained by a learner who is confronted with primary linguistic data. The intuition of markedness theory applied to language acquisition is that highly complex grammars will lose out to grammars that take the primary linguistic data to reflect a higher degree of generality. So, for example, if we find a language in which wh-phrases undergo movement to clause-initial position, we will not posit that the language allows for the movement of wh-phrases from within moveable phrases unless there is specific evidence that shows this (an instance of the A-over-A condition, Chomsky 1964). The learner will not generalize in such a way that idiosyncrasies will be predicted in the absence of positive evidence.[17] Only very strongly supported idiosyncrasies will be able to overcome the learner's tendency to generalize.

What this is getting to is the idea that when the learner is exposed to evidence of restrictions that have limited range, the learner forms hypotheses that are correspondingly limited. Generalizations may occur, but they are constrained by the specificity of the evidence. When the evidence supports a broader generalization, then specific conditions that a learner may have previously imposed are gradually lost.[18]

5. The Cambridge Grammar Of The English Language is alive and well and living in your head. Now consider the following scenario. Suppose that H&P had set out not to create CGEL, but to create GODZILLA-CGEL. GODZILLA-CGEL contains everything that H&P knew about English as of January 1, 2003. Suppose that we had given them the

resources to do this, in principle at least, instead of requiring them to work for a living at the same time. Suppose that we gave them as many helpers and as much money as they needed. Suppose that their collaborators, by a stroke of good luck, had in their heads precisely the same knowledge that H&P had in theirs, so that there would not be any of the indeterminate 'on the one hand, many speakers would not say this, but on the other hand, many would.' GODZILLA-CGEL would not be a description of the knowledge of an ideal native speaker – it would be a description of the knowledge of some native speakers; let's call it $\text{English}_{\text{H\&P-1/1/2003}}$.

One question to ask about GODZILLA-CGEL is whether it could actually be created in finite time. I am going to assume that it can be, through a natural extension of what went into CGEL. Most of what is missing in CGEL is a list of all of the words in a typical educated speaker's passive vocabulary at some particular time and their properties. To come up with such a list is a daunting, but by no means should be an impossible task, especially if we throw unlimited resources at it and set some cutoff date.

What about the constructions of English? Is there a bounded number of these? Again, I am going to say 'yes,' reasoning that constructions are like words, except that (i) they consist of more than one word and (ii) they may be stated in terms of categories. While there may be a lot of these, their number must be finite, since they are not defined recursively.

As far as the general constructions, that is, phrase structure and movement rules, are concerned, again they appear to be finite in number. As far as can be determined, they are all in CGEL, but if somehow CGEL missed some, we'll put them in GODZILLA-CGEL.

Let us not forget that for every word and construction, GODZILLA-CGEL will have to say what it means. Let us assume that there exists a semantic theory that we can put into GODZILLA-CGEL in terms of which we can state the meanings of individual words and constructions to the extent that they are idiosyncratic, and the result of composing them into larger expressions. Such a theory would be accompanied by explicit theories of information packaging and conversational implicature, which would formalize and extend the many discussions in CGEL devoted to drawing the boundaries between narrowly grammatical phenomena, and phenomena that are related to grammar but go beyond it.

Since we have unlimited resources, we might as well count the frequency of occurrences of each problematic configuration and encode them as probabilities in GODZILLA-CGEL. We would thereby be able to flesh out the intuitively correct but generally unsupported remarks in CGEL to the effect that a certain way of saying something is possible, but relatively infrequent, and so on.

Supposing that all this is possible, let us consider a computer program that implements all of this knowledge. The computer program will assign to any sentence a structural analysis, following the rules and descriptions of GODZILLA-CGEL. It will assign an interpretation to any string of words, following the semantic theory of GODZILLA-CGEL. For any string of words that is not a sentence, the program will indicate that it is ungrammatical. In fact, CGEL often indicates degrees of ungrammaticality based on the properties of particular lexical items, and GODZILLA-CGEL should be able to draw upon this knowledge. Give the computer a meaning, and using GODZILLA-CGEL it should be able to find one or more ways of expressing it.

Here's the question: Does the computer know $\text{English}_{\text{H\&P-1/1/2003}}$?[19,20] Or, to put it another way, is what we have put into GODZILLA-CGEL a sufficient characterization of what a native speaker of English has in his/her head? This is an awkward question, since both the answer 'yes' and 'no' are troublesome, but for different reasons. If we answer 'yes', we are in effect claiming that the modified Aspects-style grammatical theory of CGEL is all that is in the head of a native speaker, as far as a grammatical theory is concerned. The grammatical theory is supposed to correspond in some way to the architecture of the Language Faculty (see Chomsky 1965 for the defining statement of this relationship). In developing a grammatical theory, we are seeking to explain why languages have the properties that they have. This explanation is mediated by the Language Faculty, which leads the language learner to construct grammars that conform to certain constraints and take a particular form. Our 'yes' answer would commit us to rejecting as irrelevant to the goal of linguistic explanation much if not all of subsequent MGG, including the Conditions framework, GB Theory, Barriers, Principles and Parameters Theory, and the Minimalist Program, not to mention the architectural offerings of HPSG, and LFG.

Obviously there is much of value in these theories, and many genuine empirical results, and those must and will appear in GODZILLA-CGEL whether or not they are in CGEL. For example, the core ideas of the Binding theory, which CGEL treats relatively superficially, must somehow be included, as well as the major constraints on extraction, which CGEL does cover. But much of the theoretical literature is irrelevant to CGEL, will not be reflected in GODZILLA-CGEL, and hence has nothing to do with the Language Faculty, if our answer is 'yes'.

Let's suppose that the answer is 'no, the computer does not know English$_{\text{H\&P-1/1/2003}}$. But we just agreed for the sake of the discussion that the computer program incorporated all the knowledge of English that we were able to extract from the heads of H&P (and their likeminded collaborators). So either the answer is 'yes', or we must reject the premise that it is possible in principle to say what a native speaker knows, using a grammatical theory, a lexicon, probabilities, a semantic theory, a theory of information packaging, and a theory of discourse structure. Unless we want to admit into linguistics some notion akin to quantum uncertainty, it may seem at first that there is no way out of this quandary except to say 'yes', GODZILLA-CGEL is indeed all that there is in the speaker's head.

6. The way out. The way out of this dilemma is to accept the conclusion that syntactic theory as we know it for the most part has nothing directly to do with what is in the head, beyond the part that we have to concede to GODZILLA-CGEL. That is, it does not correspond in any principled way to the architecture of the Language Faculty. Rather, it is a description of the way in which the Language Faculty behaves, that is, the expressions that it produces, the acceptability judgments that it comes up with, and so on. It is also a description of the possible relationships among these expressions.

It is not inconceivable, of course, that a syntactic theory may also be an integral part of an explanation of how knowledge of language is acquired or represented in the mind. While it may play no role in the description of what is in a speaker's head, it may play a role in explaining how it got there. Such a theory would be more than a description of the behavior of the Language Faculty. It would be part of the description of the Language

Faculty itself. Of course, linguists have been clear for a long time that a grammar is not a model of the language learner, or the speaker-hearer (Chomsky 1965:9). But this brings us back to the question of whether GODZILLA-CGEL knows $\text{English}_{\text{H\&P-1/1/2003}}$.

To the extent that the relationships brought out by a syntactic theory do reflect properties of the Language Faculty, they are surely not accidental. To deny that a syntactic theory is an account of the Language Faculty per se is far from a pessimistic position. We should not deny the importance or interest of the theoretical enterprise. The usefulness of a theoretical statement about the phenomena per se does not reside solely in the possibility that it might be about the Language Faculty as well. It resides in the fact that it may be a non-accidental property of a natural language for which an explanation must be found. This is what makes syntactic theory 'interesting', I believe. But, this said, it is an open question whether the explanations of the observed phenomena are to be found in the architecture of the Language Faculty, in the organization of Conceptual Structure, in an account of the constraints on real time processing of natural language (e.g. memory, parallel vs. serial processes, etc.), in the cognitive and computational constraints on language learners, in the dynamics of information structure in discourse, or in the social and cognitive dynamics of language interaction and change. Quite beyond its massive survey of the grammatical phenomena of English, CGEL has contributed significantly to our understanding of the proper role of syntactic theory by making it abundantly clear what it really means to know a language.

References

Berwick, Robert, and Partha Niyogi. 1996. Learning from triggers. Linguistic Inquiry **27**.605-622.

Borer, Hagit. 1984. Parametric syntax. Dordrecht: Foris Publications.

Chomsky, Noam. 1964. Current issues in linguistic theory. The Hague: Mouton.

Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Chomsky, Noam. 1972. Some empirical issues in the theory of transformational grammar. Goals of linguistic theory, ed. by Stanley Peters, 63-130. Englewood Cliffs, NJ: Prentice-Hall.

Chomsky, Noam. 1977. On wh movement. Formal syntax, ed. by Peter W. Culicover, Thomas Wasow, and Adrian Akmajian, 71-132. New York: Academic Press.

Chomsky, Noam. 1982. The generative enterprise. Dordrecht: Foris Publications.

Chomsky, Noam. 1986. Knowledge of language: its nature, origin, and use. New York: Praeger.

Chomsky, Noam. 1995. The minimalist program. Cambridge, MA: MIT Press.

Culicover, Peter W. 1999. Syntactic nuts. Oxford: Oxford University Press.

Culicover, Peter W., and Ray Jackendoff. 2001. Control is not movement. Linguistic Inquiry 30.483-511.

Culicover, Peter W., and Ray Jackendoff. 2005. Syntax made simpler. Oxford: Oxford University Press. In press.

Culicover, Peter W., and Andrzej Nowak. 2003. Dynamical grammar. Oxford: Oxford University Press.

Fodor, Janet D. 1994. How to obey the Subset Principle: binding and locality. Syntactic theory and first language acquisition, Vol. 2: Binding, dependencies and learnability, ed. by Barbara Lust, Gabriella Hermon, and Jaklin Kornfilt. Hillsdale, NJ: Lawrence Erlbaum Associates.

Fodor, Janet D. 1998. Unambiguous triggers. Linguistic Inquiry 29.1-36.

Fodor, Janet Dean. 2001. Parameters and the periphery: reflections on Syntactic nuts. Journal of Linguistics 37.367-392.

Gibson, Edward, and Kenneth Wexler. 1994. Triggers. Linguistic Inquiry 25.407-453.

Hornstein, Norbert. 2003. On control. Minimalist syntax, ed. by Randall Hendricks. Oxford: Blackwell.

Huddleston, Rodney D. 1988. Review of Quirk et al.: A comprehensive grammar of the English language. Language 64.345-354.

Jackendoff, Ray. 1977. X-bar syntax: a study of phrase structure. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1990. Semantic structures. Cambridge, MA: MIT Press.

Jackendoff, Ray, and Peter W Culicover. 2003. The semantic basis of control. Language 79.517-556.

Katz, Jerold J., and Paul M. Postal. 1964. Toward an integrated theory of linguistic descriptions. Cambridge, MA: MIT Press.

Landau, Idan. 2003. Movement out of control. Linguistic Inquiry 34.471-498.

Merchant, Jason. 2001. The syntax of silence. Oxford: Oxford University Press.

Pullum, Geoffrey. 1984. Linguists and computers. Natural Language and Linguistic Theory 1.311-315.

Quirk, Randolph, Greenbaum, Sidney, Geoffrey Leech, and Jan Swartik. 1985. A comprehensive grammar of the English language. London: Longmans.

Ross, John R. 1967. Constraints on variables in syntax. Cambridge, MA: MIT dissertation.

Ross, John R. 1969. Guess who. Proceedings of the fifth annual meeting of the Chicago Linguistics Society, ed. by Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 252-286. Chicago, IL: Chicago Linguistics Society.

Tomasello, Michael J. 2000. Do young children have adult syntax. Cognition 74.209-253.

Notes

---

[*] I am grateful to Ray Jackendoff and Geoff Pullum for their helpful advice and for saving me from some serious errors. I take full responsibility for any errors that remain and for the views expressed herein.

[1] Rodney Huddleston was the original conceiver and planner of this work, and Geoff Pullum joined him in the project in 1995. Although they are the main authors of this work, important contributions were also made by Laurie Bauer, Betty Birner, Ted Briscoe, Peter Collins, David Denison, David Lee, Anita Mittwoch, Geoffrey Nunberg, Frank Palmer, John Payne, Peter Peterson, Lesley Stirling, and Gregory Ward.

[2] Quotes taken from http://books.cambridge.org/0521431468.htm.

[3] And they have in fact won a prize: the LSA's Leonard Bloomfield book award for 2001-2003! (See http://www.lsadc.org/index2.php?aaa=lsanews.htm for details.)

[4] There are remarkably few errors in CGEL; see http://people.ucsc.edu/~pullum/errata.html for current errata.

[5] See Culicover 1999 for a theoretical justification of this position.

[6] I have in mind examples on the order of *The DA proved the defendants to be guilty during each other's trials (Chomsky 1995:272) or *It is certain John to seem intelligent (Chomsky 1995:365).

[7] Curiously, CGEL does not even cite Chomsky in the references given in Further Reading (1765-78).

[8] The possibility of such a reaction was not unanticipated by the author/editors. On page xvi they refer explicitly to their policy of not citing sources, even in footnotes. And on the Cambridge University Press website (http://uk.cambridge.org/linguistics/cgel/faqs.htm) one

finds the following exchange: 'I don't see any references to the literature in the grammar's pages. How come?' Answer: 'This is a reference grammar, not a monograph about linguistics. No references to the literature are given in the body of the work. A modest attempt at attribution of key ideas is made in the Further Reading section at the end, together with a list of references, but the idea of including a complete bibliography of the gigantic field of English grammar could not even be considered.' My own view is that this policy is a mistake.

[9] The ideas sketched out in the next few paragraphs are developed in much more detail in Chapter 2 of Culicover and Jackendoff 2005.

[10] Another option is deletion under identity, which is a notational variant.

[11] There are also empirical motivations for such an analysis, but as argued in Culicover and Jackendoff 2005, these can be dealt with in a natural way without complicating the syntactic analysis.

[12] For a particularly clear statement of this type of thinking, see Hornstein 2003. Hornstein argues that control should be reduced to movement, thereby eliminating the control relation from the grammar. For a critique of Hornstein's proposal, see Culicover and Jackendoff 2001, Jackendoff and Culicover 2003 and Landau 2003.

[13] See for example Chomsky 1982. Similar sentiments are often expressed. For example, a colleague writes, quoting another colleague: 'if the Jackendoffian model of language is correct, and he seems to have figured it all out, then what we end up with it not very interesting anymore. He solved it. All questions answered, nothing else to do.' A strange response to the possibility of having uncovered the truth about a major scientific mystery, the nature of human language. In other sciences people win the Nobel Prize for a lot less.

[14] I refuse to entertain the proposition that linguists are not scientists.

[15] The extent of the differences can be seen in a close comparison with Q85, which I will not undertake here. For some premonitions about the ways in which CGEL was planned to diverge from Q85, see Huddleston 1988.

[16] One of the problems that make parameter setting not quite so easy is that languages are not always that consistent; see for example Fodor 1994; Fodor 1998; Fodor 2001. There are also some difficult technical issues because the superficial facts do not always define a path to a unique and correct grammar when more than one parameter must be set (see Berwick and Niyogi 1996; Gibson and Wexler 1994), making the picture more implausible.

[17] This point is made in more detail in Culicover 1999. Ross 1967 focused the attention of the field on the broad explanatory power of constraints on transformations and consequently, away from considerations of relative complexity.

[18] See, for example, Tomasello 2000.

[19] This is not a Turing test. It's fine if you can tell that it's a computer. The question is, can you distinguish its knowledge of English from H&P's?

[20] In an eerily prescient passage, Pullum 1984 offers the following description of what linguistics is: 'Suppose you wanted to program a computer to understand plain English, like the HAL 9000 computer in Stanley Kubrick's film *2001: A Space Odyssey*. Linguistics is the subject that figures out what you'd need to know about a language in order to do that, for English or any other language, in a general and theoretically principled sort of way.'