FORMAL LINGUISTICS AND THE ORDINARY WORKING GRAMMARIAN*

Geoffrey K Pullum
University of California, Santa Cruz

0. Introduction

Exactly thirty years ago this month, in October 1969, a conference was held at the University of Texas at Austin on the 'Goals of Linguistic Theory' The goals of linguistic theory were not a matter of agreement then, any more than they are now, as the proceedings volume (Peters 1972) makes very clear Indeed, some participants were questioning the very idea of there being a defensible syntactic theory The pessimistic final section of Postal's paper expressed the opinion that "so little is understood of linguistic structure" that early transformationalists were making a mistake when they "naively assumed that it was actually possible at the time to construct generative grammars for human languages" (p 160) Mindful of the problem of what grammarians will do if they do not write grammars, he invokes "a no doubt never-to-be-written paper" by Lakoff, Postal and Ross called "What to do until the rules come," which was to discuss what grammarians should do with their research time if they weren't going to be writing grammars (p 168, n 50)

Fillmore's paper in the volume caught the mood, coining the term "the New Taxonomy" for "an era of a new and exuberant cataloguing of the enormous range of facts that linguists need eventually to find theories to deal with" (p 16), and introducing the Ordinary Working Grammarian a shadowy figure whose views and reactions are discussed throughout the paper Fillmore expresses the hope that the Ordinary Working Grammarian will not be judged by "his ability to demonstrate that his grammar does everything that generative grammars have been said to have to do" (p 18), because that seems impossible "Having a good time" as a practitioner of the New Taxonomy might have to be sufficient, and after all, he concludes, "It is possible to remain happy, for a while, without well-defined goals"

How long should we wait? How many years is enough to spend waiting for the rules to come? Exactly ten years ago this month, twenty years after the Texas conference, I decided that we had been waiting long enough Speaking by invitation to the Western Conference On Linguistics (WECOL) on the topic of what we should expect in linguistics in the 1990s, I reviewed some back-of-envelope computations to show that there had probably been some ten thousand or more person-years invested in generative linguistic research by then, with results hardly sufficient to justify this outpouring of effort For example, I complained, "there is absolutely no sign of generative grammar reaching the point where randomly selected practitioners will give approximately equivalent answers when asked for the syntactic surface structure of simple English sentences" Part of the trouble, I felt, lay with the heady idea that linguistics was plumbing the mysteries of the mental I observed that "It is hard to get across to people who think they have glimpsed a principle of the linguistic faculty of the human mind that they will have to substantiate that by exhibiting descriptions of linguistic phenomena which both appeal to the putative principle and equal or surpass previous descriptions in breadth or depth of insight"

I pointed out the growing evidence of retrogression — indeed, one might call it circumitineration — were evident in the unacknowledged plundering of generative semantics that was going on in the late 1980s

---

within MIT linguistics I was worried that having spent the Nixon administration failing to agree that the ideas of generative semantics were correct, syntacticians were now spending the Bush administration circling nostalgically around those ideas again, without acknowledging it (Pullum 1996 has some further discussion on this) [1] What I didn't know then was that things would get worse It is sobering to look back now at Postal's paper, with its claim that a theory with semantic representations and surface structures but no level of deep structure should "because of its *a priori* logical and conceptual properties" be regarded as "the basic one which generative linguists should operate from as an investigatory framework, and should be abandoned, if at all, only under the strongest pressures of empirical disconfirmation" (Postal 1972, p 135) The published version of Chomsky's paper at the 1969 conference argued firmly against this, claiming it to be misguided in principle

> "Improvements from the worst possible case will come by placing more possible conditions on the choice of grammars, limiting the kinds of rules that can appear in them and the ways in which these rules can operate   Thus it is misleading to say that a better theory is one with a more limited conceptual structure, and that we prefer the minimal conceptual elaboration, the least theoretical apparatus Insofar as this notion is comprehensible, it is not in general correct " (Chomsky 1972, p 68)

Today, as is well known, Chomsky takes the opposite view the syntax defines a derivation relating a phonetic form directly to a logical form without a level of deep structure He goes on to say (Chomsky 1995, p 187)

> Ideally, that would be the end of the story each linguistic expression is an optimal realization of interface conditions   Any additional structure or assumptions requires empirical justification [W]e may ask whether the evidence will bear the weight [of supporting deep structure], or whether it is possible to move toward a minimalist program

Not a word is added to remind the reader that this is a capitulation to Postal's 1969 a priori conceptual arguments in favor of the "homogeneous" theory as a default — the idea that deep structure was an extra assumption that would need special evidence to support it [2]

---

[1] It is of course this complete lack of acknowledgment, and not just the fact that Chomsky has changed his mind, that is worrying about the quotations at the beginning of this paper There is no reference to Postal (1972) — or Chomsky (1972) — in the bibliography of Chomsky (1995) Surely anyone making the kind of U-turn in syntactic theory that Chomsky has made should signal the turn, or at least admit that the turn has been made What we are seeing here is not Orwell's Problem (Chomsky 1986) but Winston's Problem references to the generative semantics era are being treated like the newspaper clipping that Winston Smith in Orwell's *1984* discovers (and immediately feels he must destroy because it clearly falsifies a crucial Party claim Orwell 1949, part One, §VII) "The immediate advantages of falsifying the past were obvious," Winston reflects, "but the ultimate motive was mysterious " He takes up his pen and records this thought *"I understand HOW I do not understand WHY "* Quite so There is a short-term saving of face from not bothering to acknowledge debts to opponents from the 1960s, but ultimately it is not clear to me what anyone has to gain by denying that reversions to 1969 theorizing are taking place

[2] The two theories are not identical, of course Postal was advocating a grammar in which logical/semantic representations were base-generated and transformations mapped them into phonetic ones Chomsky defends a view that still has syntactic structures generated independently of logical or semantic considerations A careful contrasting of the two theories might be enlightening, but Chomsky is apparently not going to be the one to provide it Seuren (forthcoming) provides some illuminating discussion

2

I could not foresee anything quite so brazen in my 1989 lecture But I did what I could to discern something of the future of the parts of the discipline I knew about And in conclusion I pushed the crystal ball aside and switched from prediction to a much easier task, exhortation I finished up — with more ultimate consequences for my own future that I could not then know — in these terms

If I had to name one thing that I felt would be most valuable for the health of linguistics in the 1990s, I would say that what was needed was a large-scale theoretical synthesis and description effort on the syntax of a single language studied in depth — probably English   Linguists are not pulling together the ideas they entertain The discipline of a team effort to lay out a serious reference grammar of English has been lacking for too long   The task will be a large one, and difficult to organize, but it would be worth it

Those who are inclined to dismiss such encyclopedizing work as relatively dull when compared to the exploration of the origins of the universe or the probing of human cognitive capacities and their genetic basis should reflect on the fact that astronomers and cosmologists have spent the better part of the last decade constructing detailed maps of the universe, and scientists who work directly on the foundations of genetics have decided that they will spend several billions of dollars over the coming decade or two constructing a complete map of the human genome An exhaustive account of what we now know about the syntax of English will be a small job by comparison with these giant cartographic endeavors We can spare a few hundred person-years, surely

I had forgotten the key principle governing such things as wise behavior at faculty meetings be very cautious about making passionate speeches advocating that novel and arduous tasks should be undertaken, because *you will be taken to have volunteered*

I was not then aware that across the Pacific in Australia, Rodney Huddleston had become convinced that a new major descriptive grammar of English was vitally needed He had written a review for *Language* of the comprehensive grammar produced by Randolph Quirk and his team (Huddleston 1987), and that effort had taught him that the Quirk grammar will not do Praiseworthy as it might be, it does not have a theoretically coherent view of what syntax is about, or even a consistent terminology It is far too wedded to traditional mistakes, commits far too many misanalyses, and despite its bulk, omits several topics and treats others carelessly

Huddleston obtained funding from the Australian Research Council and set up a mainly Australian team to begin work on a new grammar in the early 1990s The project was, as I had predicted, both large and difficult to organize In late 1995, when it became clear that the grammar was leaking and more hands were needed at the pump, Cambridge University Press contacted me to invite me to join the team The plan was for a massive 1800-page descriptive grammar to be called *The Cambridge Grammar of English* (Huddleston and Pullum, forthcoming, henceforth *CGE*) It was to be a comprehensive synchronic description of the syntax and morphology of present-day general-purpose international Standard English, intended for a general readership I demurred for a while, but by early 1996 I had realized that as the author of the above quotation I had no choice I agreed to join the team

I have never seen a summer since that day By July 1996 I was in Australia working through the southern winter on the project with Huddleston at the headquarters for the *CGE* project, his home office in Kenmore, Queensland All my summers since 1995 have been Australian winters, and some of my autumns have been tropical springs in Queensland I have now spent more than a full 100%-time person-year on the project, over twelve months full time, and it seems like a lot more It is the project that ate my life But it has been the hardest and most interesting work I have ever done in linguistics It turns out to be very exciting to be attempt a complete description of a language that one knows natively and about which vast quantities

3

of evidence can be obtained and a huge library of literature is available The difference between this kind of work and the theoretical work that I was formerly more used to is quite dramatic

I have been forced to learn the peculiar kind of self-effacingness that the Ordinary Working Grammarian must cultivate (that habit so familiar to Fillmore, and so alien to me) Theoreticians gain kudos by taking apparently ordinary facts and discovering in them something surprising, descriptive grammarians have to take apparently surprising facts and fit them into the ordinary pattern with least fanfare This is because of what the user of a grammar of this sort expects those who look up a topic in a reference grammar are not looking to be handed 1969-style awe and wonderment ("Wow, these data are just so cool, they blow every known theory away") or 1999-style pretentious biologism ("The facts are subtle and will mostly be ignored here but through them I think we can glimpse a property of the computational system of the mind/brain") Ordinary users want conservativity and continuity, so there should be no wholesale abandonment of traditional assumptions except where the motivation is extremely strong (an ordinary grammar user cannot be expected to re-learn the terminology of grammar every three months) They also want authoritativeness But no hype, no approximation, no suppressed exceptions, no copout

Attempting to work to such desiderata is a real exercise in discipline Theoretical syntacticians can enjoy the luxury of saying in a footnote, "There are certain subclasses of verbs for which this is not true, but I will not consider them here," while a descriptive grammarian — at least, given the ground rules we have established for *CGE* — has to list every one of those verbs, and find any generalization that governs them More generally, theoretic syntacticians can ignore or sidestep humdrum sets of facts that don't seem relevant or interesting, whereas descriptive grammarians have to maintain clarity of focus and egalitarian coverage throughout, describing the apparently humdrum just as carefully as the supposedly fascinating (not that one can tell the difference before giving serious attention to a problem, I have found)

I want to discuss one or two examples of the changes in my perceptions of my job as a grammarian that have followed from my intensive involvement in the work of producing *CGE*, and to try to assess the extent to which my background in theoretical linguistics has been a help or a hindrance, a source of inspiration or a source of shame I will argue that the no-doubt-never-to-be-written paper by Lakoff, Postal and Ross (which did indeed never get written) basically got things right [3] The great achievement of transformational-generative grammar these past forty years has been to equip the Ordinary Working Grammarian with a vastly improved set of conceptual tools, analytical tests, argumentational strategies, rules of thumb, and criteria for the evaluation of syntactic analyses Today we can be far more certain of our decisions regarding whether this or that grammatical analysis is the more reasonable one, we can at last improve on the descriptions provided in even the most assiduous traditional grammars, and be sure we are introducing an improvement With these new techniques of investigation and rigorous argumentation we have in fact been able to open up whole fields of grammar that are brand new we are obtaining results on topics that were not even topics in grammar fifty years ago

It is the business of theory construction and the establishing of broader metatheoretical principles that has stalled I believe that many of the supposed accomplishments of generative linguistics are mythical The intuitional methods that replaced corpus methods in the 1950s are outmoded, the "argument from poverty of the stimulus" that is supposed to have linked linguistic results to the explanation of language acquisition has not in fact been given substance, the conception of generativity that most linguists have been working with is the wrong one, having tempted us into confusions on quite basic claims about language — as basic

---

[3] Perhaps I should say "would have gotten things right", it is difficult to know exactly how one should talk about a never-to-be-written paper

as the claim that natural languages are infinite, and the signature achievement, the development of transformational rules, was just a mistake that has wasted an inordinate amount of time

What is so ironic about this, and what makes it worth commenting on, is that what is believed outside our discipline is basically the opposite of all this What generative grammar has in fact done best lies in areas like broadening our understanding of what is a grammatical fact, and developing new and successful kinds of grammatical argument Such work is very little known to the journalists, rhetoricians, literature professors, psychologists, neurologists, and philosophers who write about the supposed Chomskyan revolution, what they all concentrate on is the supposed cataclysmic impact of the discovery of transformations, and the alleged power and insight of the generativist-Cartesian-cognitive-biological world view for which Chomsky has supposedly made a compelling case

Ironically, it is what generative linguistics has done least well that has gained it the most fame the construction of very general theories with universal and multidisciplinary import The outside world seems to have it backwards Linguistics is being lionized for what it has not achieved instead of what it has

1. Elusive part of speech assignments

To illustrate what I think linguists are doing well, in this section I will examine some simple part-of-speech distinctions that ought to have been clearly settled long ago but for various reasons were not Traditional English grammar has had the part of speech of a large number of frequent items wrong, and for others had the right answer but for no very good reason I think that today we can put right the misanalyses and provide clear and definite grammatical arguments in support of our categorizations

If this seems like a trivial enterprise, by the way, think again If we cannot definitively determine and support syntactic categorizations for given items in a well-studied language, we can do nothing else in computational terms, we cannot even do tagging, let alone parsing, in representational terms, we cannot even draw the bottom tier of a tree, let alone the branches closer to the root, in minimalist terms, we do not even know what the numeration is of a sentence is, because we cannot say what lexical categories are represented by the words occurring in it, X-bar theory determines nothing about what phrasal projections there are if we cannot name the lexical categories Nothing could be more fundamental to syntax than getting lexical items correctly assigned to defensible lexical categories It is analogous to getting straight about identifying the elements for a chemist

1 1 Diagnosing prepositions
The study of English grammar owes a lasting debt to Joseph Emonds (1972), and to useful supplementary work by Ray Jackendoff (1973, 1977), for making it fully clear for the first time in the history of English grammatical studies that a serious error has traditionally been made concerning the categorization of items such as the following

(2)	around	back	by	down	forth
	home	in	off	on	out
	over	past	round	through	up

Because under the traditional view a word can only be a preposition if it has an NP complement, in each of these pairs the underlined word is traditionally preposition in the [i] case and adverb in the [ii] case

(3) a [ı] *I had come <u>across</u> a couple of mimes*   [ıı] *I had come <u>across</u> a couple of times*
    b [ı] *A ship sailed <u>by</u> each bay*   [ıı] *A ship sailed <u>by</u> each day*
    c [ı] *I had gone <u>up</u> a mountain before*   [ıı] *I had gone <u>up</u> a moment before*
    d [ı] *The bus went <u>past</u> every single stop*   [ıı] *The bus went <u>past</u> every single hour*
    e [ı] *He was cast <u>in</u> a bit part*   [ıı] *He was cast <u>in</u> a bit later*

Notice that the meaning of the items in question here does not vary *down* means 'down' What varies is only whether the NP that follows them is their object (as in the [ı] cases) or a time adjunct that does not form a constituent with them (as in the [ıı] cases) A comparable analysis for verbs would assign two parts of speech to (say) *eat* while called a verb in transitive clauses like *I plan to eat this macaroon*, it would be categorized in some different way when it occurred without a direct object, as in *I plan to eat this afternoon* How could any grammarian be fully satisfied with an account like this?

What Emonds did was to provide criteria to defend a new analysis, developing them into rigorous arguments that the underlined words in the [ıı] cases of (3) are prepositions One test that he introduced involved the use of *right* as a pre-head modifier ('specifier') [4] *Right* does not premodify any category but preposition in contemporary Standard English, neither adjectives nor undisputed adverbs take it

(4) a [ı] *It was right out of sight*   [ıı] *It fell right in our area*   [P]
    b [ı] **It was right invisible*   [ıı] **The place where it fell is right local*   [Adj]
    c [ı] **It was done right invisibly*   [ıı] **It fell right locally*   [Adv]

The relevance of *right* modification is seen when we test it on items like the underlined ones in the (4a–e[ıı]) *Sailed right by* is just as good as *sailed right by each bay*, *fell right down* is just as good as *fell right down the stairs*, *went right past* is just as good as *went right past every stop*, and so on The traditional account would have to be (5a), when surely (5b) makes a lot more sense

(5) a   The traditional view of *right*
        Right can be used as a pre-head modifying adjunct only with
        (a) prepositions, or
        (b) the members of a select and ad hoc class of adverbs, all of which lack the *-ly* suffix with which
            the central members of the adverb class are characteristically formed, and all of which are
            homophonous with prepositions
    b   The post-Emonds view about *right*
        Right can be used as a pre-head modifying adjunct only with prepositions

Other arguments urge us in the same direction Emonds points out that the constructions illustrated by *Into the pool with him!* or *Up the mast with that flag!* are restricted to having a PP as the initial constituent, and what do we find but *In with him!*, *Up with it!*, etc And the inversion construction illustrated by *Out of the hole popped a mouse* or *Into the room ran a messenger* demands a PP as the initial constituent, but we find *Out popped a mouse* and *In ran a messenger*

It can also be pointed out that there are verbs subcategorized to take PP complements verbs like *put* demand a locative PP complement (*Put these on your hands*), and verbs like *head* demand a directional PP complement (*We headed toward the harbor*), but we also find *Put these on* and *We headed in* And there are

---

[4] *Straight*, *clear*, and *smack* (or *smack dab*) can be used similarly, but I concentrate here on *right* because of its wider distribution

negative tests that can confirm that some item is **not** a preposition For example, the verb *become* never allows a PP complement we get *become a maniac* and *become insane* but not **become out of your mind* And sure enough, the behavior of intransitive prepositions is in line with this we do not find **become out/in/up/through/to* etc These alleged adverbs fail to occur with *become* in exactly the way as we predict if instead they are prepositions

In addition to these tests for prepositionhood that work for intransitive (non-complement-taking) prepositions, there are of course others depending on the syntax of those prepositions that do have complements Notably, we can appeal to the fronting of PPs containing *wh*-NPs, generally known as pied piping When a preposition has a *wh*-NP complement and it is relativized or questioned, prepositions optionally appear in clause-initial position along with the *wh*-NP Adjectives never show this behavior This contrast permits us to distinguish *due to the bad signposting*, where *due* is a preposition, from *new to the bad signposting*, where *new* is an adjective

(6) a   *We soon got lost, due to the bad signposting*
      b   *There was some bad signposting, <u>due to</u> which we soon got lost*   [fronted PP]
      c   *We soon got lost, new to the bad signposting*
      d   **There was some bad signposting, <u>new to</u> which we soon got lost*   [fronted AdjP]

In short, we are armed today with a whole battery of useful diagnostic tests that together make an overwhelming case for categorization of certain words as prepositions in English We do not have to do this on the basis of the extremely vague universalist definitions of 'preposition' that traditional grammar attempts to provide, there are clear-cut characteristic behaviors internal to English grammar that we can use

1 2 Diagnosing adjectives
Maling (1984) is concerned with applying arguments of the general sort just reviewed in an attempt to get straight the categorizations of certain words that might be taken for either adjectives or prepositions I will in fact argue that her decisions are not the right ones, but it should not be overlooked that her paper opened up a nice problem and made some very useful contributions

Maling reviews several different criteria for adjectivehood and ruling out prepositionhood, concentrating on those that are purely syntactic Among the phenomena she suggests we pay attention to are the use of *very* and *very much* as pre-head adjuncts (what she calls specifiers) adjectives take *very* but not *very much*, while with prepositions the reverse is the case (compare *very affectionate*, **very much affectionate*, *very much in love*, **very in love*) This diagnostic is often helpful, though we will see below that it can let us down

Maling also notes that the distribution of the word *enough* is peculiarly useful because it refuses to premodify adjectives (we get *good enough* rather than **enough good*) (The latter test has a complication nothing is permitted to separate a lexical head from its NP complement, so that we should not expect to see *enough* after any adjective that can take an NP complement, and there are such adjectives, as we shall see below) Finally she notes that comparative inflection, taking the negation prefix *un-*, and occurrence in prenominal attributive modification are also useful tests in some contexts

Not noted by either Emonds or Maling is a further very useful test for distinguishing adjectives (and participles in certain constructions) from prepositions, pointed out to me by Rodney Huddleston It is partially semantic, because it turns on the existence or nonexistence of **predicative** readings for fronted adjuncts Fronted PPs are capable of functioning as **nonpredicative** sentence adjuncts, whereas AdjPs, NPs, and VPs that occur as preposed adjuncts are always predicative Consider these examples

7

(7) a   *Ahead of the ship, the captain saw an island on which to land*   [PP]
     b   *Tired of the ship, the captain saw an island on which to land*   [AdjP]

Both are grammatical and meaningful, but there is a difference (7b) entails that the captain was tired of the ship, but (7a) does not necessarily entail that the captain was ahead of the ship It could have such a meaning, but what is important is that another reading is also possible, one in which *ahead of the ship* is not predicated of *the captain*, but merely indicates where the island was sighted To put it another way, (7a) does not guarantee the truth or the grammaticality of the sentence *The captain was ahead of the ship*, but (7b) does entail the existence and the truth of *The captain was tired of the ship* In (7b), only a **predicative** reading exists if (7b) is true, then the description "tired of the ship" applies to the captain This fact correlates with the categorization of *ahead* as a preposition and of *tired* as an adjective The contrast is seen again in this pair

(8) a   *Due to the terrain, Kim soon got lost*   [PP, nonpredicative reading]
     b   *New to the terrain, Kim soon got lost*   [AdjP, predicative reading]

We do not read (8a) as entailing the strange claim that Kim was due to the terrain (i e the terrain somehow caused Kim to exist) We do, however, read (8b) as entailing the claim that Kim was new to the terrain This is an indication that while *new* is an adjective (as confirmed by its semantically regular comparative and superlative inflected forms *newer* and *newest*), *due* is a preposition in the contemporary language

Grammaticality differences can result from the property under discussion For example, *owing* was at one time just the gerund participle of the verb *owe*, but has long since become a preposition The gerund participle of the verb *owe* can occur in participial constructions like *Owing several thousand dollars on his credit card, Jim was paying a lot of interest* These are always predicative (in the example just given, *owing him several thousand dollars* is predicated of *Jim*) The preposition *owing* takes a PP complement with *to*, but the verb takes a direct object as well (as in *You owe $5 to Kim*) Thus we find this syntactic contrast

(9) a   *Owing to my stupid bank, there's no money for the payroll this Friday*
     b   **Owing money to my stupid bank, there's no money for the payroll this Friday*

In (9b), the direct object *money* ensures that *owing* must be the gerund participle of the verb *owe*, but in that case *owing money to my stupid bank* is a VP and must be predicative (in traditional terms, it needs an 'understood subject'), and nothing in the following main clause provides any appropriate NP to be the target of the predication But (9a) can be understood with *owing* as a preposition PPs can be nonpredicative sentence adjuncts, so we can understand *owing to my stupid bank* as "because of my stupid bank"

To summarize, when a fronted adjunct has only predicative readings, that is sufficient to indicate that it is not headed by a preposition or an adverb, when it has only nonpredicative readings, that is sufficient to indicate that it is not headed by an adjective

1.3 Telling adjectives and prepositions apart
I summarize in (10) all the reliable tests I am aware of for identifying prepositions and adjectives, including those mentioned in the foregoing sections All of the diagnostics below state sufficient conditions for belonging, or for not belonging, to a certain category

(10) a Sufficient conditions for being a preposition
    i Dedicated preposition premodifiers *right*, *straight*, *clear*, and *smack*
      A word occurring with one of these as pre-head modifier is occurring as a preposition
    ii Pied piping
      An occurrence of a word optionally fronted along with an 'extracted' item in an unbounded dependency is a preposition occurrence
  b Sufficient conditions for being an adjective or adverb
    i Comparative inflection
      An occurrence of a word with grade inflection (comparative *-er* or superlative *-est*) is an adjective or adverb occurrence
    ii *Very* intensification
      An occurrence of a word with *very* as pre-head adjunct is an adjective occurrence
  c Sufficient conditions for **not** being a preposition
    i *Become* complementation
      An occurrence as head of the complement of *become* is not a preposition occurrence
    ii Premodifier function in AdjP or AdvP
      A pre-head modifier occurrence in AdjP or AdvP is not a preposition occurrence
  d Sufficient conditions for **not** being an adjective or adverb
    i Premodification by *enough*
      An occurrence of a word with *enough* as pre-head modifier is not an adjective occurrence
    ii *Very much* intensification
      An occurrence of a word with *very much* as an intensifying pre-head modifier is not an adjective occurrence
  e A necessary condition for being a preposition or an adverb
    If a word is a preposition or an adverb, it will have nonpredicative readings when heading a fronted adjunct
  f A necessary condition for being an adjective
    If a word is an adjective it will have predicative readings when heading a fronted adjunct

This is the toolbox I will use to repeat Maling's experiments on two problematic items, *near* and *worth*, and show that her results are in error

Problem I *Near*

The *Oxford English Dictionary* (*OED*) treats *near* as a preposition in those cases where it has a complement and (following the traditional analysis criticized above) an adverb in those cases where it is similarly used but has no complement, and a separate entry for *near* as an adjective is given to cover attributive uses like *a near relative* or *the near future* But this does not quite get it right, as the following results show

    Since *near* passes the tests in (10a) for prepositions, occurring freely with the premodifier *right* (*right near the wall*) and participating in pied piping (*the wall near which it fell*), we know it is a preposition

    Near also passes the tests in (10b) for adjectives and adverbs even when it has an NP complement, inflecting for grade (*nearer the wall*, *nearest the wall*) and taking *very* intensification (*very near the wall*), so we know it also occurs as an adjective or adverb

    We find that *near* can head the complement of *become* (*Soon it became near quitting time*), and thus we further confirm by (10ci) that there are non-preposition occurrences, in addition to the non-prepositional

occurrences found premodifying adjectives and adverbs as noted by the *OED* (*near complete failure, near perfectly camouflaged*)

The marginal occurrence of *near* with degree premodifiers *enough* (*?The gunshots were enough near the house that we were worried*) and *very much* (*?It was very much near the house*) weakly confirms by (10d) that there are non-adjective occurrences, though clearly (and not unexpectedly) there is a strong preference for the adjective *near* when the sense is modified in terms of degree

Finally, since *near* can have a nonpredicative reading when heading a fronted adjunct (*Near the wall but out of his reach he saw a crowbar*) we know that it has prepositional occurrences, while because it can have a predicative reading when it heads a fronted adjunct (*Near death, the lone survivor staggered to a nearby farmhouse*) we know it sometimes has non-preposition and non-adverb occurrences

(11) a [i] *It fell right near the wall*   [ii] *the wall near which it fell*
      b [i] *nearer the wall, nearest the wall*   [ii] *very near the wall*
      c [i] *Soon it became near quitting time*   [ii] *near perfect(ly)*
      d [i] *?enough near us to make me nervous*   [ii] *?very much near us*
      e [i] *Near the wall but out of his reach*   [ii] *Near death, he staggered to a nearby*
             he saw a crowbar   *farmhouse*

The conclusion is clear Maling's claim, that *near* is solely an adjective, is incorrect The word is dually categorized as an adjective and as a preposition (the meanings being apparently identical) As an adjective, *near* optionally takes an NP complement and inflects for comparative and superlative grade, as a preposition, it obligatorily takes an NP object and does not so inflect Its behavior under the usual tests for class membership is mixed in just about exactly the way we would expect for an item with a dual classification

Manning and Schutze (1999 11–13), in an introductory chapter of a textbook on statistical natural language processing, use *near* as an example of a word illustrating "non-categorical phenomena in language", namely "blending of parts of speech" They note examples like *We will review that decision in the near future* as evidence of attributive adjective use, and examples like *He lives right near the station* as evidence of prepositionhood They then cite examples like *He has never been nearer the center of the financial establishment* (with both an NP complement and adjectival inflection) as evidence of the overlapping of prepositionlike and adjectivelike properties But they miss what is really crucial here that when the adjective behavior is at its most unambiguous, in particular when the word is inflected in the comparative or superlative, the most unambiguous prepositional behavior disappears

(12) a   *It is right near the wall*   [*right* modification Preposition]
      b   *It is nearer the wall than it was*   [comparative inflection Adjective]
      c   **It is right nearer the wall than it was*   [no categorization possible]

This follows immediately from the dual categorization account together with the preposition-only limitation on *right* and the fact that grade inflection occurs only on adjectives

Problem II *Worth*

Now I consider an item, *worth*, that is more problematic but still capable of being categorized on the basis of fully convincing evidence

Maling's conclusion about *worth* is introduced with the remark "As counterintuitive as it may appear, *worth* is best analyzed as a preposition" and a footnote saying "The fact that our first intuitions about *worth* and *near* turn out to be wrong shows how misguided the attempt to provide notional definitions of categories is " The intuitions to which she refers are that since *near* is semantically locational it should be a preposition and that since *worth* has no sense that is in any way locational it should be an adjective Notional definitions may or may not be misguided, at least at a parochial level (rather than as part of an attempt to link part of speech assignments in different languages to each other as a contribution to universal grammar), but Maling is wrong about *worth* it is a further example of an adjective taking an NP complement, this time one that does not also have an analysis as a preposition It is a rather unusual adjective — and its entry in the very traditional *OED*, presumably based on notional definitions, gets this right

One way that *worth* is unusual among adjectives is that its complement is absolutely obligatory (There are a few others *fond, desirous*, etc) The obligatory complement is either an NP denoting some kind of index of value, as illustrated in [24i], or else a gap-containing clause understood as a value-determining property of what the subject denotes

(13) a   *That book turned out to be worth seventy dollars*
      b   *I think you'll find this worth your time*
      c   *This idea is worth giving some thought to __*
      d   *The house is certainly worth your going to see __*

Given the almost complete prohibition against attributive use of adjectives with subcategorized complements, this entails that *worth* is restricted to predicative function

A second oddity is that *worth* is an exception to the strong tendency for monosyllabic adjectives to take grade inflection **worther the money than the other one* is completely impossible Comparison is periphrastic with *worth It was more worth the money than the other one you bought*

Worth is also fairly incompatible with *very* (?*very worth the money*) and yet accepts *very much* However, this diagnostic, while it often points in the right direction, cannot be relied upon crucially, because it turns out that there are adjectives that allow *very much*, a clear example is *alike* there is nothing wrong with *The two are very much alike* Moreover, non-gradable items always take *very much* with the sense "decidedly" rather than "to a high degree", as in *The ship is very much unique in its class*, which means not "the placement of the ship on the scale of uniqueness in its class is very high", but rather "the appropriacy of describing the ship as unique in its class is very high" Thus the *very much* test cannot be relied upon

However, the generally reliable criterion of *enough* placement confirms that *worth* is an adjective *enough* will not premodify adjectives, and sure enough, as anyone may verify, corpus examples of the sequence '*enough worth* NP' are not found at all

The test provided by predicative readings in fronted adjuncts is particularly important in confirming this As fronted adjuncts, *worth* phrases are always interpreted predicatively

(14) a   *Worth five minutes, the article will tell you a lot about snorkeling*
      b   *Within five minutes, the article will tell you a lot about snorkeling*

It is entailed by (14a) that the article is worth five minutes (of your time), this is a predicative reading, with *the article* as target of the predication But it is not entailed by (14b) that the article is within five minutes, the *within* phrase in (14b) is interpreted nonpredicatively A consequence of this is that if we change the

11

examples to introduce a dummy *it* subject in the matrix clause, we get an ungrammatical result in one case but not in the other

(15) a *Worth five minutes, it will become obvious to you that snorkeling is fun*
      b *Within five minutes, it will become obvious to you that snorkeling is fun*

A dummy subject cannot be the target of a predication, so (15a) is ungrammatical

Summarizing our results, the full picture looks like this

(16) a [i] *It was right worth the money*   [ii] *I paid $75, worth which I thought it*
     b [i] *worther the money*   [ii] *very worth the money*
     c [i] *It became worth taking him seriously*   [ii] (no attributive use)
     d [i] ?enough worth your time   [ii] ?very much worth the time spent on it
     e [i] (no nonpredicative use)   [ii] *Worth five minutes, the article will*
                                                            tell you a lot about snorkeling

Most of these results are neutral, for example, all we learn from (16a) is that the test fails to show *worth* is a preposition, and all we learn from (16b) is that the test fails to show it is an adjective However, (16c[i]) definitely tells us that *worth* is not a preposition, and (16e) definitely confirms this

Worth is quite unusual in bringing the property of being an adjective together with five others that are highly unusual for adjectives

   (a) like *near*, it permits its complement to be an NP,
   (b) like *loath*, it selects a complement obligatorily,
   (c) like *awake*, it is completely excluded from prenominal attributive modifier function,
   (d) like *main*, it has the syntactic behaviour characteristic of strictly ungradable adjectives, and
   (e) like *extra*, it is inert with regard to adjectival derivation processes

All of these help to disguise its adjectival status, but at the same time, none of them are unrepresented elsewhere among uncontested adjectives *Worth* is strikingly unusual, perhaps unique, in having all five of these unusual properties at once Nonetheless, it is possible to confirm the *OED* claim about its part of speech classification with overwhelmingly greater confidence using novel syntactic and semantic tests

In such ways we can defend categorization decisions with complete conviction and multiple lines of argument, even with difficult and marginal cases *Near* is both an adjective and a preposition (not just an adjective as Maling claims) and *worth* is solely an adjective (not a preposition as Maling claims) My point about these small results in syntax is that today, as a result of the past several decades of intensive syntactic research, we can provide more solid support for claims of this sort than was ever possible before, and we can correct with confidence both some traditional claims and some generativist claims

2. New grammatical territory

I have suggested so far that the development of far more rigorous and detailed ways of arguing for elementary claims like categorization claims is a major achievement of modern linguistics I now want to mention briefly — too briefly — a second major advance, which consists in the opening up of new domains of facts Generative grammatical work has brought into focus several entirely new domains of facts — not just new facts in an area of grammar that had been inadequately mapped by traditional grammarians, but whole new areas that were unnoticed by traditional grammarians

12

2 1 Unbounded dependencies

There are few better examples that could be cited than that of **unbounded dependency constructions** or **UDCs** A UDC is a syntactic construction in which a designated subconstituent is, to put it intuitively, missing a phrase is missing in a place where that sort of phrase would be expected under the ordinary principles of internal syntax Let us call such missing phrases **gaps** (in order to remain neutral on the controversial question of whether there are traces — phonetically null but syntactically real constituents — at the positions in question) A typical example will be something like (17a)

(17) a   *I would imagine most people would enjoy a job like yours*
  b   **I would imagine most people would enjoy*
  c   *A job like yours, I would imagine most people would enjoy __*
  d   **A job like yours, I would imagine most people would enjoy being an airline pilot*

With a strictly transitive verb like *enjoy* we would expect a direct object, as in (17a) In (17b) it is missing, and the sentence is ungrammatical But (17c), with a new fronted NP, the direct object of *enjoy* is missing and the sentence is nonetheless grammatical, the gap is marked by '__' And as (17d) shows, the sentence actually becomes ungrammatical if a direct object for *enjoy* is added

Examples like (17) give a slightly misleading impression they suggest that what is involved in an unbounded dependency is a phrase that is misplaced from where it would normally have been positioned For some cases this is not a correct description of the situation to give just one example, in (18a) the constituent in clause-initial position cannot be substituted for the gap, as the ungrammaticality of (18b) shows

(18) a   *We cannot afford Susan, brilliant analyst though she is*
  b   **We cannot afford Susan, though she is brilliant analyst*

(I return to this topic below) What this suggests is that we need to characterize this phenomenon not in terms of phrases that are "shifted" from their canonical or normal positions, but rather, in terms of the existence of syntactic domains that are required to contain gaps

2 2 Island constraints

The plot thickens when we note that gaps have to be in certain syntactic positions It has been customary to give transformational theories in a way that suggests that as a first approximation gaps can be anywhere, but this default is overridden by the existence of certain constraints on their positioning I now think this is backwards Unconventionally (following an important insight of a neglected paper, Cattell 1976), I think the default picture can be given in positive terms, as follows

- Gaps in English must always correspond to constituents of the clause that are semantic arguments (subjects, objects, or other complements) or post-head modifiers, they can never be determiners or pre-head adjuncts
- Gaps are primarily permissible only if connected back to the root of the domain by a chain of internal complements and/or phrasal head relations (Keeping in mind that direct objects are complements, and that I take the VP of a clause to be the head of that clause, the gap in (17c) is complement of a complement of the head of a complement of a complement of the head of the domain)
- To a very limited extent can a gap be buried inside post-head adjuncts (*Which day is he arriving on __ ?*, but probably not ?**Which book did you get angry because she had lost?*)
- For a gap to be buried inside left-branch material is not permissible at all (**Who did George's not liking surprise you?*)

13

From this characterization there follow a variety of specific prohibitions that have long been observed under the heading of island constraints gaps cannot be inside clausal subjects, or in temporal adjuncts, or in relative clauses, or in genitives, or in subcoordinates of a coordinate structure, for example This discovery of this complex web of restrictions and exceptions to restrictions is a post-1960 syntactic discovery of major importance

3.3 Multiple gaps

The account sketched thus far applies only to cases with single gaps There are two circumstances in which multiple gaps are found in a single UDC One is across-the-board cases like *Tell me one thing that you like __ and she doesn't also like __* The other is parasitic gap cases like *Which was the memo that you tore __ up without even looking at __?* Each has quite specific conditions The across-the-board constructions are those in which each subcoordinate of a coordinate structure has a gap in it and the gaps are all controlled by the same superordinate structure — for example, they are all associated with and licensed by a single relative clause construction The parasitic gap cases have extra gaps in positions where anaphoric pronouns would generally also be grammatical, and those extra gaps have to be inside constituents that are sisters to domains including ordinary gaps

To go into these details would take vastly more time than I have here All I want to make of this topic is this point, which I take to be uncontroversial the whole subject of gaps and where they can occur is completely missing from all grammars in the first half of this century There effectively was no such topic There was no terminology for it because the phenomena had scarcely been noted The whole cluster of phenomena surrounding gaps, unbounded dependencies, island constraints, across-the-board facts, and parasitic gaps represents an entire new region of grammatical territory that was not even discovered, let alone mapped, until the efforts of the transformational-generative period of linguistic research began, but now has been the subject of dozens if not scores of significant monographs and hundreds if not thousands of articles

And the facts are real Finding them — discovering there was so much new grammar out there — is a genuinely important accomplishment of late 20th-century theoretical linguistics

3. Generative mythology

However, when we turn to the other side of generative grammar, the side that gets the kudos from philosophy, psychology, and other areas of cognitive science, things are different Here we encounter the mythos of the paradigm, the legends and pieces of hallowed dogma that it has handed down These are not nearly as well supported as people imagine Unlike good myths, they can actively impede progress by sowing confusion The very material that has made Chomskyan theoretical linguistics famous — the topics that have excited philosophers, enraged psychologists, perplexed computational linguists, inspired cognitive scientists — do not stand up to objective scrutiny There are many aspects that could be discussed (for some attempt to survey them, see Pullum, in press), here I will just pick four

3.1 Introspection and asterisks

Consider first the methods of investigation that remain standard within generative grammar How do we obtain the facts about English that it is our job to describe? It is a rather extraordinary fact that generative grammatical research is still being done today in the same way it was being done forty years ago, with a recipe that begins, "Take one large armchair "

To illustrate what is possible in the language, the investigator sits in the armchair and uses imagination to develop an example of it and intuition to confirm that it is well-formed and makes its point To illustrate

14

what is not possible, the investigator constructs a string that would have the characteristics of the bad construction in question and uses intuition to confirm that indeed the constructed string is impossible As Tom Wasow put it to me (in conversation) the *N* (number of experimental subjects) is 1, the subject knows the purpose of the experiment, and the subject is committed to a particular hypothesis about the phenomena[1] This is not a methodology that can be taken seriously in cases of controversy

Moreover, even when considered at face value as a use of introspective data gathering, it is being abused as Schutze (1996 50) perceptively notes, linguists appear to assume subjects' intuitive judgments are veridical they ask themselves, "What must be in subjects' minds in order for this sentence to have the status they claim it has?" when it would be more appropriate to ask, "What must be in subjects' minds in order for them to react this way to a sentence?"

If there was an excuse for this forty years ago, it was that our tools for maintaining corpora of any reasonable size were so poor that armchair reflection was actually far better at ensuring broad coverage and representation for rarer constructions Certainly that was still true in the early 1960s, when according to an ex-NASA programmer of my acquaintance NASA had to make do with a total of one kilobyte of RAM Today it is common for a cheap desktop PC on a student's desk to have *sixty-four thousand times* as much memory as that And disk storage is now measured in gigabytes (billions of letters), even on cheap machines We can now store corpora larger than the corpus-based linguists of the first half of this century ever dreamed of, and search them at speeds that would have been thought science fiction just a decade ago And yet syntacticians are for the most part not using corpora at all

I am not advocating a practice of basing grammars rigidly or mechanically on corpora, so that if available corpora do not contain an instance of the *tough* construction embedded in a *wh*-relative embedded in a subject we have to gerrymander our grammar to disallow *Anyone who thinks John is easy to please had better think again* Use of a corpus does not have to make one irrational Once you have a good description of adjective phrases and a good description of relative clauses you are entitled to assume that any of your adjective phrases could fit into adjective-phrase slots in any of your relative clauses (that *easy to please* will be grammatical in the above example because *easy* or *tall* would be grammatical there), unless you find that this yields definitely unacceptable results

But two great benefits accrue to the grammarian who uses corpora in addition to using common sense and native speaker intuition First a presentational point it is far more convincing to illustrate grammatical structures with examples chosen from a collection of sentences that have already been attested in natural contexts, claiming them as a part of the language has much greater persuasiveness if it can be shown that they are repeatedly used by speakers and writers of the language Second, it is far more convincing when a certain construction type is claimed not to be permitted in the language if that claim can be made into an empirical prediction that strings of certain types will never be found in corpora (except perhaps sporadically as errors)

Certainly, this will miss the distinction between the extremely rare construction type and the occasionally encountered error, but my point is that syntacticians do not even use corpus checks on their work when studying quite frequent construction types And one can hardly take as a key empirical datum a reported negative intuitive reaction by the very person who (a) invented the example and (b) wants to convince of the hypothesis that the example's ungrammaticality will support It would be greatly preferable if the theoretician took an intuition of ungrammaticality to be nothing more than the basis for a conjecture

15

about what the language does not permit,[5] and looked for evidence to support or disconfirm it on the basis of a corpus of attested utterances Certainly, the feeling expressed by "Sentences containing a clause beginning with *that for* do not sound good to me" constitutes grounds for at least some suspicion that clauses cannot begin with *that for*, but a demonstration that in a hundred million words of diverse prose the sequence *that for* cannot be found in clause-initial position at all constitutes a powerful vindication of that suspicion — which still might, of course, be wrong, but is not nearly as likely to be wrong as unaided intuition

Let me consider one real case in which corpus checking was sorely needed but not employed the attempt by Higginbotham (1984) to argue that English is a non-context-free language by virtue of a class of sentences that he calls *such that* relatives It is a necessary premise of Higginbotham's argument that every *such that* relative contains a pronoun anaphorically linked to the head noun *a result such that no one could believe it* is grammatical but crucially *a result such that no one disagreed* is not But the search for evidence here is absurdly simple one searches text for the word sequence *such that* and sees what comes up And if you try it on any reasonable collection of written English (I used the ACL's *Wall Street Journal* corpus) you will find sentences like these

(19) a *Speculation in platinum futures has been a driving force **such that** an equivalent of 81 million ounces were traded on the New York Mercantile Exchange last year*

  b *Global warming has reached a level **such that** we can ascribe with a high degree of confidence a cause and effect relationship between the greenhouse effect and observed warming*

  c *These regulatory offenses create more complexity **such that** we get away from the old-fashioned types of crimes that everybody can understand*

It isn't just easy to find examples like this in abundance, it's trivial [6] And they aren't just troublesome for Higginbotham's argument, they are fatal Perhaps there was an excuse for working without looking for corpus evidence in 1984 (I don't really think so, because you can readily find examples of the crucial sort in the *Oxford English Dictionary*, see Pullum 1985 294), but there is certainly no excuse today

Things are even clearer when constructions are claimed to be impossible My proposal would be that the asterisk should be given a straightforward empirical interpretation putting an asterisk on an example constitutes a claim that the construction type illustrated will never be found in an error-free corpus of material from the language in question Above I noted that if *enough* refuses ever to premodify adjectives and *worth* is an adjective (as I claim), we have a very simple prediction the word sequence *enough worth* should not occur at all in even the largest corpora (except perhaps by accident across a constituent boundary, as in *Are students who are not clever enough worth worrying about?*) It is easy enough to test this claim nothing more than a fixed-string search is needed to pick out the candidate examples

It is extraordinarily difficult to convince linguists of this They are equipped with many knee-jerk reactions They will say that people say so many wild things that a large corpus is likely to contain examples of just about anything, including just about everything you could think of that is not grammatical But this isn't true, as experience with corpus methods will show anyone who chooses to make the effort to find out

---

[5] This corresponds to the second interpretation for the asterisk given in Householder (1973) "I have never seen or heard a sentence of the type X and hereby wager you can't find an example "

[6] These sentences were found in the corpus of text from *Wall Street Journal* articles on the CD-ROM made available by the Association for Computational Linguistics I have edited them to reduce their length, but only in ways not relevant to the point

They will say that using a corpus restricts one's purview, which is also not true — using a corpus is liberating, and expands one's understanding rather than contracting it, because of all the examples that turn up that exhibit relevant usages that would not have immediately come to mind What they will not say is that they cannot be bothered, but I suspect that is what is going on

I am not suggesting that intuitions of grammaticality be banished from our armory of tools for investigating language Heaven forfend I am saying that a better methodology for a grammarian today involves a back-and-forth interplay between hypothesis, intuitive reflection, corpus searching, refinement of the intuition, prediction concerning what will be found in the corpus, further searching, and so on — combined with occasional recourse to informants or even acceptability surveys

3.2 The alleged poverty of the stimulus
The term "argument from poverty of the stimulus" appears to have dropped into linguistic discourse when Chomsky (1980 34) referred to "a classical argument in the theory of knowledge, what we might call 'the argument from poverty of the stimulus'," citing no references other than Socrates' elicitation of knowledge from the slave boy and Descartes's argument in the *Dioptrics* that "there is no need to suppose that anything material passes from objects to our eyes to make us see colors and light " From these inexplicit remarks a tradition has somehow grown up of asserting that Chomskyan linguists have shown that human infants learn things about their first languages that they were provided with no evidence for I will not discuss this topic in detail here, though I have treated it in a preliminary way elsewhere (Pullum 1996) Suffice it to say that although reference books in philosophy and cognitive science now contain articles that attempt to outline the argument from poverty of the stimulus and sketch the support linguists are supposed to have offered for it (Garfield 1994 and Marcus 1999 are two examples), I see few signs of anyone attempting to provide such support in a serious way in the domain of syntax

The argument could in principle be tested I take it that a specific instance of applying the argument to particular phenomena would say something like the following for a specific speaker S, a fact F, a class D of sentences of a language L, and a class E of fact-based 'empiricist'-style learning procedures

(20) a F is a fact about language L, and S is a speaker of L who can be shown to know F
     b   It can be demonstrated that no empiricist-style learning procedure of the type E can learn F from a corpus of utterances from L unless that corpus includes a sample from the specific domain D of crucially relevant utterance types
     c   It is known that in the process of acquiring L, there was never at any time at which S was exposed to data from the domain D
     d   Therefore, S did not acquire L using a learning procedure of type E

There is much to be specified precisely here the class E of empiricist-style learning methods must be explicitly characterized, that F is really true must be confirmed by descriptive work on L, careful informant work or psycholinguistic experimentation must be done to show that S really does know F, the domain D must be explicitly defined, the unlearnability of L from D-free corpora by methods of type E must be proved as a theorem of learnability theory, and somehow it must be shown convincingly that S never encountered data from the domain D during the acquisition period — an empirical matter involving longitudinal investigation of input to (and ideally uptake in) the learner A tall order, but one can see in principle what it might be like to accomplish all this However, when we search the literature of linguistics for an example of this program being carried out for some syntactic fact, we find virtually nothing

In Pullum (1996) I consider the only close approach I then knew of Let F be the fact that subject-auxiliary inversion in English is structure-sensitive (it fronts the main clause auxiliary rather than the first

17

auxiliary in the string), and let D be the class of English sentences containing two auxiliaries in a configuration that permits us to tell the difference between fronting the main clause auxiliary and the first auxiliary (i e , sentences like *Could those who are leaving early sit near the door?*) I suggest that there will be great difficulty in exhibiting a learner S who can be guaranteed never to have been exposed to utterances of the type illustrated by D, because (as Sampson 1989 suggested might be the case) such utterances are quite easy to find in any reasonable-sized corpus

At present my claim stands as a challenge that no one has taken up At least two authors of recent books on how language might be learned (Cowie 1999 and Sampson 1998) take the view that the onus is now on generative linguists to respond to the challenge I have laid down My feeling is that linguists are lucky that philosophers and psychologists have been so credulous on this point, because there are few signs of real substantiation of the claim on which the argument from poverty of the stimulus trades Linguistics is getting some credit here that it simply does not deserve

3 3 The infinity myth
Mention of size limits brings me to a generativist myth that has not been questioned in forty years the claim that natural languages are infinite This topic will be treated more fully elsewhere (Pullum and Scholz, in preparation), but I will give a brief digest

The view that natural languages are infinite is unsupported by any sound argument, empirical or formal A typical defense of it is put by Stabler (1999) this way

> Although there are obvious practical limitations on the lengths of sentences that any human will ever
> pronounce, these bounds do not seem to be linguistic in nature, but rather derive from limitations in
> our life span, requirements for sleep, and so on As far as the grammars of natural languages go, there
> seems to be no longest sentence, and consequently no maximally complex linguistic structure, and we
> can conclude that all natural languages are infinite

But we cannot validly conclude that The key to seeing why lies in a fairly elementary point about model theory Call the language in which a formal grammar is written a **description language** A description language must have a semantics if grammars are to make identifiable claims about what is in the language under description Grammars of the type I will call **production systems**, of which the rewriting systems introduced by Post in the 1940s are an example, have a semantics in terms of set definition Rules like $PP \rightarrow P\,NP$ have a similar status to move-permitting statements in the definition of games like chess 'a pawn may be moved one square perpendicularly ahead' To the extent that a model-theoretic semantics for them can be provided, it defines the entire language at once No individual rule of a production system can be interpreted as making any statement about an individual sentence The rule $PP \rightarrow P\,NP$, for example, does not claim that prepositions are required to precede their NP complements (there could be a rule $PP \rightarrow NP\,P$ in the grammar as well), just as 'a pawn may be moved one square perpendicularly ahead' does not say that pawns have to move perpendicularly (there is also a rule 'a pawn may take an opponent's piece that is adjacent diagonally ahead,' which permits a pawn to move one square diagonally) The only claim that is made by the rules of a production system is that the entire set of rules defines the entire collection of sentences for which the system provides derivations

There is an alternative Over the last twenty years a different kind of description language has been developed, one that provides for declarative constraints that have a model-theoretic semantics, individual sentences (or sentence structures) being the models Call these **constraint systems** An example is provided by the grammars defined by Johnson and Postal (1980), to take one of the earliest examples of such a framework each grammatical constraint is a material conditional in a first-order predicate calculus in which

18

the predicates denote properties of or relations between arcs (labeled ordered pairs of nodes) Each constraint is either true or false of any arbitrary individual sentence structure A sentence structure is admitted by the grammar if and only if all rules of the grammar are true of it

Making grammars strictly declarative in this way has several interesting and desirable consequences One has to do with degrees of grammaticality Consider how one might distinguish between the mild ungrammaticality of (21a), the greater ungrammaticality of (21b), and the extreme ungrammaticality of (21c) not in terms of a production system but in terms of a constraint system

(21) a *They have been informed the time of his arrival
    b *Have been informed the time of his arrival they
    c *Arrival the been have of they time his informed

Among the declarative constraints on syntactic form in English that are relevant here are those given informally in (22)

(22) a Subjects precede predicates in the clause
    b Perfect *have* takes a subjectless past participial complement
    c In a passive clause the copula takes a subjectless past participial VP complement
    d Articles precede nominal expressions in the NP
    e Lexical heads precede their complements
    f Subject pronouns are in the nominative case
    g If *inform* has an NP complement and a second complement, the latter is a PP
    h A PP second complement of *inform* is headed by *of*

Of these statements, only the last is false of (21a) Only the first and the last are false of (21b) But (with the words assigned to the obvious lexical categories) **all** of them are false of (21c) A quantitative index of approach to grammaticality is available that at least has some *a priori* plausibility degree of grammaticality is linearly correlated with number of statements in the grammar satisfied And nothing has to be added or stipulated to obtain this

So constraint systems have a certain desirability purely from the standpoint of factual coverage But there is an important further consequence of constraint systems that has gone entirely unnoticed many different collections of sentence structures will satisfy all the rules of a grammar, there is not a unique such collection In fact there is not even a clearly defined largest one, as in effect shown later by Langendoen and Postal (1984), since sentences of infinite size need not be stipulatively excluded, and thus transfinitely vast collections of sentence structures may satisfy the grammar, including collections so big that set-theoretic notions like cardinality do not apply to them

It is true that for any interestingly complex grammar (any grammar with the analog of direct or indirect recursion, i e in which a structure of type α may occur as a proper subpart of a structure of type α) there will exist infinite collections of sentence structures satisfying it But that does not mean that finite collections do not satisfy it There will be infinitely many finite sets of sentences that constitute models of the grammar

We therefore do not need to assume that English is identical with the smallest denumerably infinite set of sentences that satisfies the grammar, which is what a production system says under the standard interpretation We do not need to fix upon any finite cardinality for a unique set that is to be by stipulation the formal analog of English Instead we can say that the size of the language is **not fixed** by the grammar

Whether it is finite or infinite is neither a matter that grammatical study is concerned with nor a question that a grammar answers For linguistics, language size should be a non-issue (since it is uncontroversial that, as Stabler puts it, size bounds "do not seem to be linguistic in nature, but rather derive from limitations in our life span, requirements for sleep, and so on") Under a model-theoretic view of the semantics for description languages, it is a non-issue, since the existence of infinite models for a grammar does not imply the nonexistence of finite ones, Stabler is wrong to conclude that English is infinite If our grammatical description is given in the form of a constraint system, it can be any size, finite or infinite, and provided our constraints capture the right structural properties the same description will work no matter what the cardinality of the collection of all sentences (if we assume there is any such collection) It is only production-system grammars such as transformational grammars that mislead us into thinking that languages must be infinite (see Pullum and Scholz, in preparation, for a more careful exploration of these ideas)

3 4 The failure of movement rules

Finally let me turn to a central and definitive innovation of transformational grammar that is crucially tied to the production conception of grammars the feature that survives in all varieties that bear the name of transformational grammar, even in those where deletion rules (the main danger as regards Turing-equivalence proofs) are banished the device of movement rules Most linguists seem to recollect being convinced — often by reading Chomsky (1957) — that movement rules were an excellent idea The tyranny of procedural metaphors seems to have all thinking about syntax in its icy grip Yet the classic arguments for the necessity of movement transformations are unsound The famous Affix Hopping analysis is not compatible with the formal definition of transformations originally given by Chomsky (Sampson 1979 360–365) and simply does not work descriptively (Pullum 1979 244–247, Gazdar, Pullum and Sag 1982 613-616) Movement derivations of passives had been identified by the early 1970s as entirely unnecessary (Brame 1973, Freidin 1975, Bresnan 1978) The most solid arguments in favor of movement rules that was available twenty years ago were the kind of which Perlmutter and Soames (1979 229ff) provide a crystal-clear instance, based on data of the kind I presented above in (17), arguing that phrase structure rules simply cannot capture the generalizations involved But twenty years ago this month, while Perlmutter and Sambas book was being distributed, Gerald Gazdar realized that the argument for movement transformations to account for topicalization sentences was entirely unsound There was nothing about topicalization facts that would defeat context-free phrase-structure description

What had been missed was that context-free grammars allow arbitrary latitude as regards the content of the nonterminal vocabulary — the set of syntactic categories To put it very simply and intuitively, and compatibly with the terms suggested above, context-free phrase structure rules permit us to distinguish a category 'Clause' (for brevity, S) from a category 'Clause with an NP gap inside it' (abbreviated S/NP) We can regard (17a), *A job like yours, I would imagine most people would enjoy* __, as consisting of an NP (*a job like yours*) followed by an S/NP (*I would imagine most people would enjoy* __), thus

(23) $[_S [_{NP} a job like yours] [_{S/NP} I would imagine most people would enjoy __]$

The internal structure of an S/NP will be just like that of a clause, except that there must be some daughter that instead of having the label α that would normally be expected in a clause of this sort has the corresponding label for 'α-with-NP-gap' instead More generally, wherever a constituent of the category α normally allows a daughter sequence $\varphi^\frown\beta^\frown\psi$, a constituent of the category α/γ will allow the daughter sequence $\varphi^\frown\beta/\gamma^\frown\psi$ (The result is still a context-free phrase structure grammar, because although the size of the set of categories has been expanded, it expands only from k categories to a maximum of $k + k^2$ categories — if for any α and β, not necessarily distinct, there is a category 'α-with-β-gap' — and in practice the needed expansion is much smaller)

This makes it clear that putting an asterisk on an example like (17b), *I would imagine most people would enjoy __*, is misleading [7] this string is not ungrammatical in the sense of violating grammatical constraints It is simply not of the category S Rather, it is of the category S/NP, and thus can be used as a bare relative clause (*something I would imagine most people would enjoy*) And the ungrammaticality of *A job like yours, I would imagine most people would enjoy being an airline pilot* is due to the fact that a clause in English is not permitted to consist of an NP followed by a clause

Thus the basic arguments for movement presented in introductory books on transformational grammar are unsound it simply has not been shown to be necessary to augment phrase structure grammar by movement transformations to achieve a description of the familiar facts of English syntax that have been held to motivate movement But things are in fact much worse than that What the movement idea suggests heuristically is in fact misleading There are numerous constructions that should give pause to anyone who thinks critically about the concept We need go no further than independent interrogative clauses to see that in some cases the mover cannot be put back into the gap position with a grammatical result

(24) a   *Who do you think you are __?*
     b   **Do you think you are who?*

Thus if it was the simple intuition about (17a) that it is like the antecedently grammatical (17b) with a phrase pulled out of its canonical position that motivated us, things do not go so well here

Of course, this is not to say that no movement account of the facts in (24) can be constructed The standard account has two movements, one feeding the other the *wh*-phrase in (24a) is moved to the beginning of the sentence and this triggers movement of the auxiliary into second position But now the trouble lies with this second movement, subject-auxiliary inversion there are sentences where the wrong auxiliary turns up after movement

(25) a   **I aren't coming with you*
     b   *Aren't I coming with you?*
     c   **I aren't good enough to compete with her*
     d   *Who aren't I good enough to compete with?*

The alleged mover — the auxiliary *aren't* — occurs in pre-subject position in sentences where it would not be possible in post-subject position Both the *wh*-movement and the auxiliary movement are afflicted with the problem that they occur in post-movement position in cases where they would not be permitted in the supposed pre-movement position The mover is impossible in the position of the gap There are numerous other examples of the same sort

(26) a   *He wanted me to take over, which I couldn't __*   [cf *   *I couldn't which*]
     b   *They thought it was blue, which it wasn't __*   [cf *   *it wasn't which*]
     c   *Susan, brilliant analyst though she is __, gave up*   [cf *   *she is brilliant analyst*]
     d   *That he was there that night I am certain of __*   [cf *I am certain of that he was   ]
     e   *Who the hell do you think you are __?*   [cf *You are who the hell?]
     f   *Where else could they go?*   [cf *They could go where else*]
     g   *Whatever else you do __ in Sydney, visit the aquarium*   [cf *You do whatever else   ]

---

[7] This point was made by Brame (1981 283–284)

h   <u>*Whoever*</u> *he let __ in appears to have left no fingerprints*   [cf **He let whoever in*]

In each case the alleged mover is marked by double underlining, and the gap is shown by a single underlining Putting the alleged mover back into the position it is supposed to have moved from yields an ungrammatical structure

A different type of argument is provided by the following examples, where the problem is that we have a gap but there is no visible mover

(27) a   *Some day he'll come along, the man* [*I love __*]
    b   *You're the one that* [*I want __*]
    c   *You should have seen the way* [*he looked at me __*]
    d   *You'd be so nice* [*to come home to __*]

In none of these is there any apparent mover at all To relate the occurrence of the indicated gaps in the bracketed constituents to some kind of movement, what has to be assumed is that something moves to leave the gap and then disappears through some kind of spontaneous combustion

Worse still for the intuition that is supposed to motivate movement are the cases in which several movers set off on their journey from different coordinate subparts of a coordination, leaving several gaps, but by the time they arrive at their destination they have fused into one

(28) a   *It was on a stupid TV show,* <u>*which*</u> *I hate __ and my partner loves __*
    b   *They cut up __ and threw __ to the sled dogs* <u>*the remaining chunks of bear meat*</u>

In these the double-underlined mover has to be associated with two different gaps (Actually it can be arbitrarily many *a show which I hate __, my brother dislikes __, my sister loves __, and my parents are neutral about __*) Attempts by Williams (1978) to show that such cases could be treated by a formal innovation making coordinates occupy the same linear position in sentence structure have been shown fairly conclusively not to work (Gazdar 1981, Gazdar, Pullum, Sag, and Wasow 1982)

My point in reviewing these facts, most of which are fairly well known, is not to suggest that movement theories can be refuted This of course is impossible, what I am contrasting them with is a theory that has less machinery, not more Phrase structure rules on their own, without movement transformations, can be used to describe the phenomena just cited There is no way to show that a theory with movement added would necessarily do worse, it could of course simply mimic the simpler theory using phrase structure alone, and do nothing with its movement capability, so the worst possible result for movement theories is a draw I mean only to query the intuition behind movement theories, and to offer some factual background to the following observations about the present relevance of movement rule theory to the work of the Ordinary Working Grammarian

   (i)   The original arguments for movement transformations were not sound

   (ii)   The intuition that movements explain cases where some phrase is out of its canonical position is undercut by numerous cases in which either the canonical position is not a possible one for the alleged mover or the movers and gaps are not in one-to-one correspondence

(iii) The main feature of linguists' discourse that is continually reinforced by the assumption of movement rules is the descriptively unhelpful dynamic metaphor of derivations, which has linguists talking in empirically ungroundable terms about histories for sentences instead of structures of sentences

4. Conclusion

The transformational-generative linguistics that has dominated the second half of this century has been rightly celebrated and praised, but for quite the wrong reasons The great achievements of modern grammatical work lie in what has been done that has changed the life of the Ordinary Working Grammarian of whom Fillmore spoke expanding the fact base, and adding content to the toolbox of arguments, diagnostics, and criteria of which Lakoff, Postal and Ross once planned to write Real science has been done, and real progress has been made And while it may be regarded as a piece of good luck for us linguists that our discipline has been feted by outsiders like never before in its history, I think it is actually a pity that what has most captivated outsiders has been our myths I have discussed four examples of these the counterproductive idea that intuitions are data, the falsehood that a powerful "argument from poverty of the stimulus" has been developed, the logical error that has had us parroting the view that natural languages are infinite these last forty years, and the notion that the unhelpful and unworkable device of movement transformations was a technical advance We don't need the unearned kudos we have illicitly derived from these hoary myths, we have real achievements to celebrate, achievements that we would be able to present to Sapir or Bloomfield or Jespersen if they yet lived, and discuss with some pride, and some confidence that those great men of the first half of this century would agree we had not wasted our half

References

Brame, Michael K 1973 *Conjectures and Refutations in Syntax and Semantics* New York North-Holland
Bresnan, Joan 1978 A realistic transformational grammar In Halle, Bresnan and Miller, 1–59
Chomsky, Noam 1957 *Syntactic Structures* The Hague Mouton
Chomsky, Noam 1972 Some empirical issues in the theory of transformational grammar In Peters (1972), 63–130
Chomsky, Noam 1986 *Knowledge of Language Its Nature, Origin, and Use* New York Praeger
Cowie, Fiona 1999 *What's Within? Nativism Reconsidered* Oxford Oxford University Press
Emonds, Joseph E 1972 Evidence that indirect object movement is a structure-preserving rule *Foundations of Language* 8, 546–561 Reprinted in Maurice Gross, Morris Halle, and Marcel-Paul Schutzenberger (eds), *The Formal Analysis of Natural Languages Proceedings of the First International Conference*, 73–87 (The Hague Mouton)
Fillmore, Charles J 1972 On generativity In Peters (ed) 1972, 1–19
Freidin, Robert 1975 The analysis of passives *Language* 51, 384–405
Garfield, Jay 1994 Innateness In Guttenplan 1994, 366–374
Gazdar, Gerald 1981 Unbounded dependencies and coordinate structure *Linguistic Inquiry* 12, 155–184
Gazdar, Gerald, Geoffrey K Pullum, and Ivan A Sag 1982 Auxiliaries and related phenomena in a restrictive theory of grammar *Language* 58, 591–638
Gazdar, Gerald, Geoffrey K Pullum, Ivan A Sag, and Thomas Wasow 1982 Coordination and transformational grammar *Linguistic Inquiry* 13, 663–677
Guttenplan, Samuel, ed 1994 *A Companion to the Philosophy of Mind* Oxford Basil Blackwell
Halle, Morris, Joan Bresnan, and George Miller, eds 1978 *Linguistic Theory and Psychological Reality* Cambridge, MA MIT Press

Higginbotham, James 1984 English is not a context-free language *Linguistic Inquiry* 15, 119–126

Householder, Fred W, Jr 1973 On arguments from asterisks *Foundations of Language* 10, 365–376

Huddleston, Rodney 1988 Review article on Quirk et al 1985 *Language* 64, 345–354

Huddleston, Rodney, and Geoffrey K Pullum Forthcoming *The Cambridge Grammar of English* Cambridge, U K Cambridge University Press

Jackendoff, Ray S 1977 $\bar{X}$ *Syntax* Cambridge, MA MIT Press

Manning, Christopher, and Hinrich Schutze 1999 *An Introduction to Statistical Natural Language Processing* Cambridge, MA The MIT Press

Marcus, Gary 1999 Poverty of the stimulus arguments In Wilson and Keil 1999, 660–661

Orwell, George 1949 *1984* New York Harcourt Brace

Perlmutter, David M, and Scott Soames 1979 *Syntactic Argumentation and the Structure of English* Berkeley, CA University of California Press

Peters, P Stanley, ed 1972 *Goals of linguistic theory* Prentice-Hall, Englewood Cliffs, New Jersey

Postal, Paul M 1972 The best theory In Peters (1972), 131–170

Pullum, Geoffrey K 1979 *Rule Interaction and the Organization of a Grammar* New York Garland

Pullum, Geoffrey K 1985 *Such that* clauses and the context-freeness of English *Linguistic Inquiry* 16, 291–298

Pullum, Geoffrey K 1989 Prospects for generative grammar in the 1990s In Frederick H Brengelman, Vida Samiian, and Wendy Wilkins, eds, *Proceedings of the Western Conference on Linguistics, Volume 2 1989*, 257–276 Department of Linguistics, California State University, Fresno, California

Pullum, Geoffrey K 1996 Nostalgic views from Building 20 *Journal of Linguistics* 32, 137–147

Pullum, Geoffrey K In press Linguistic models To appear in the proceedings of the conference on Mind, Brain, and Language at the University of Illinois at Urbana–Champaign, May 1998

Pullum, Geoffrey K, and Barbara C Scholz In preparation Frameworks and languages

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik 1985 *A Comprehensive Grammar of the English Language* London Longman

Sampson, Geoffrey 1979 What was transformational grammar? *Lingua* 48, 355–378

Sampson, Geoffrey 1989 Language acquisition growth or learning? *Philosophical Papers* 18, 203–240

Sampson, Geoffrey 1998 *Educating Eve* Washington DC Cassell Academic

Seuren, Pieter Forthcoming *Chomsky's Minimalism*

Stabler, Edward 1999 Formal grammars In Wilson and Keil 1999, 320–324

Williams, Edwin 1978 Across-the-board rule application *Linguistic Inquiry* 9, 31–43

Wilson, Robert A, and Frank C Keil, eds 1999 *The MIT Encyclopedia of the Cognitive Sciences* Cambridge, MA MIT Press