REVIEW ARTICLE

The post-transformational enterprise[1]

JAMES P. BLEVINS

University of Cambridge

Peter W. Culicover & Ray Jackendoff, *Simpler syntax*. Oxford: Oxford University Press, 2005. Pp. xvii + 589.

1. INTRODUCTION

Transformational models of syntax and their direct descendants have occupied a prominent place in linguistics for nearly 50 years. By some measures, these models (which can be designated collectively as 'Chomskyan') have been strikingly successful.[2] A large published literature has grown up around Chomskyan approaches to syntax, and courses that adopt a transformational perspective are firmly entrenched within the syntax programmes of many institutions. For some, Chomskyan models have come to define the field of syntax as an academic discipline. Yet a half-century of intensive study has also produced less of the stable, incremental progress that had been anticipated by early commentators. Although the general stock of knowledge about syntactic systems has advanced in certain respects during the transformational period, current theories of syntax are still some distance from providing the basis for 'a comprehensive theory of

---

[1] I am grateful to Farrell Ackerman for discussion of and comments on an earlier version of this paper and to two anonymous *JL* referees and an associate editor for criticisms and suggestions that have led to improvements in the present version.

[2] There is no fully satisfactory term for this class of models. In their treatment of morphology, they fall squarely within the Bloomfieldian tradition (Matthews 1993: 86ff.). Their approach to syntax likewise develops key aspects of Post-Bloomfieldian accounts, notably those of Harris (1951, 1965) and Hockett (1954, 1958). Hence the terms 'Post-Bloomfieldian' or even 'structuralist' (Matthews 2001) are applicable, but, like 'generative', may be construed as being too inclusive. Yet 'transformational' is too narrowly focused on a particular formal device, while the term 'mainstream' adopted by Culicover and Jackendoff is too overtly sociological.

language which may be understood in the same sense that a chemical, biological theory is ordinarily understood by experts in those fields' (Less 1957: 377).

A number of recent studies have begun to question why the Chomskyan paradigm has – so far, at least – failed to realize its early promise. One group of accounts, including Johnson & Lappin (1999), Postal (2004), Seuren (2004) and Newmeyer (2005), provide detailed critiques of current transformational models and assumptions. *Simpler Syntax* (henceforth '*SSyn*') is a significant addition to this critical tradition, setting out informed and original criticisms of the analyses and practices that the authors, Culicover and Jackendoff (henceforth 'C&J'), associate with the syntactic 'mainstream'. The criticisms are commendably free of polemic, and C&J are often generous to a fault in their assessments of transformational accounts. The historical overview in Part I stands out as an exemplary synthesis of recent theoretical developments. By tracking interdependencies between seemingly unrelated claims and assumptions, C&J trace the winding path from the 'Standard Theory' Chomsky 1965) or 'Extended Standard Theory'(Chomsky 1970) to the Minimalist programme (Chomsky 1993, 1995). Chapters 2 and 3, in particular, can be read largely on their own as a lucid synopsis of this chain of developments.

This critical commentary is not, however, presented as an end in itself, but is offered in service of the goal of diagnosing and correcting what C&J see as the core deficiencies of transformational accounts. As leading figures in a previous generation's 'mainstream' – and, in Jackendoff's case, as a seasoned veteran of the Generative Semantics campaigns (Harris 1993, Huck & Goldsmith 1995) – C&J are well positioned to assess the effect of recent innovations and to discern alternative courses of development. But the authors' closeness to the Chomskyan paradigm also limits their perspective at times, and leads them to seek 'fixes' to problems that may call for a more radical reconceptualization.

The appearance of this type of critique 'from within', so to speak, signals the emergence of a secondary literature that attempts to place transformational models within a broader intellectual context. At the same time, the narrowly technical character of the solution space explored in *SSyn* testifies to the degree to which the Chomskyan paradigm often continues to frame the terms of syntactic debate. A review of some of the basic questions raised in *SSyn* is thus of use in clarifying the nature of the reappraisal that is currently underway and in identifying areas where there appears to be scope for a more far-reaching reassessment.

The discussion of these issues below is organized into three main sections. Section 2 takes the *SSyn* critique as the point of departure for a more general appraisal of the current state of transformational approaches. Section 3 relates C&J's Simpler Syntax Hypothesis to other approaches with broadly compatible aims and orientation. Section 4 closes by drawing some of the

bolder conclusions that can be drawn from *SSyn* and the emerging secondary literature.

2. FROM THEORIES TO PROGRAMS

SSyn offers some refreshingly clear-minded criticisms of transformational assumptions, although the clarity of the objections can lend them something of an 'Emperor's New Clothes' flavour. The first part of *SSyn* reviews familiar shortcomings of individual transformational accounts, as well as more general practices, such as the use of scare quotes as substitutes for definitions of notions like 'strong' and 'weak' features. At the same time, *SSyn* presents sustained and thoughtful critiques that attempt to diagnose and correct more subtle errors.

2.1 *Filters or patches?*

A notable example concerns constraints on extraction, which are sometimes regarded as a towering achievement of generative syntax. It is indisputable that previous traditions did not formulate intricate constraints to bar the questioning of deeply embedded oblique dependents, the relativization of individual conjuncts from larger coordinate structures, the clefting of subconstituents within clausal dependents of nouns, and other similarly aberrant patterns. However, as C&J note, it is also true that few if any antecedent traditions assumed that learners would massively overgeneralize the questions, relative clauses and cleft constructions that they DO encounter, and thereby run the risk of producing the types of patterns barred by explicit constraints on extraction.

> Generalizing to the conclusion that the language allows *Move wh* (or even more generally, *Move α*) is warranted only if learners are predisposed to generalize wildly beyond their experience. We see no reason to believe that learners are so predisposed. (333)

The basic issue that C&J are raising is whether close to a half-century of research into island constraints has yielded genuine insights into extraction constructions or whether this work has merely clarified descriptive shortcomings of unrestricted movement operations. Although C&J refer to the later operations *Move wh* and *Move α*, non-local movement devices are present from the earliest transformational accounts. In *Syntactic Structures*, the movement transformation T_w applies to a string of the form $X - NP - Y$, 'where X and Y stands for any string' (Chomsky 1957: 69), and derives a string of the form $NP - X - Y$. The body of proposals that have come to be known as 'island constraints' represent attempts to restrict variables such as X. The first general constraint of this type was the 'A-over-A' principle of

Chomsky (1964), which barred extraction of an element of a given category from within a constituent of the same category. To overcome problems with this excessively general condition, Ross (1967) proposed a set of more specific constraints that referred to particular categories, such as the Complex NP Constraint and the Sentential Subject Constraint, or to particular constructions, as in the case of the Coordinate Structure Constraint. The 'Ross constraints' led in turn to attempts to generalize the restrictions on extraction, beginning with the programmatic conditions on transformations in Chomsky (1973). The goal of unifying island constraints even played a role in the development of some short-lived frameworks, such as Chomsky (1982), which explored the 'functional determination' of null elements, or Chomsky (1986), which developed a notion of 'barriers' to extraction. But a 'grand unification' of the Ross constraints remained elusive, as each of the initiatives explored a different strategy for expressing the disjunctions that held together disparate sets of corrective conditions.

Those working within the Chomskyan tradition may discern progress in these developments, and argue – at the very least – that this work has succeeded in closing off avenues of research that might have seemed to have some initial plausibility. But those who are not already convinced of the central importance of constraining constituent structure displacements may find little evidence of incremental progress. C&J invite this audience to entertain the possibility that the entire enterprise has been based on the wrong intuition. The problem raised by extraction phenomena takes on an entirely different character if one starts from the assumption that word order dislocation is local in the general case, restricted to domains such as finite clauses, and that non-local dislocations represent extensions of local patterns. From this perspective, one might expect different non-local extensions to become established in different languages, much as different clause-internal word order patterns do. There may be syntactic – alongside functional or communicative – motivation for particular extensions, just as there is for clause-internal patterns. So the challenge of characterizing extraction phenomena would not go away, though the challenge could be met by a more language- or construction-specific account. The task of describing (or acquiring) extraction patterns would not involve the 'parameterization' of general constraints on variables or domains but would instead involve the learning of non-local word-order alternations, in essentially the same way that local alternations or other constructional patterns are learned (Culicover 1999).

A constructional treatment of extraction phenomena, one which focuses on overt, observable patterns that are present in a learner's input, is particularly compatible with the conservative learning strategies that learners appear to employ (Pullum & Scholtz 2002, Tomasello 2003). Hence the shift in perspective to an input-driven or usage-based perspective would avoid the need to prevent speakers from 'generalizing wildly beyond their experience'.

From this standpoint, the key insight expressed by generative constraints on extraction would just be that any account that assumes unconstrained generalization requires corrective and ultimately heterogeneous conditions to rein in this process.

The same shift undermines the motivation for 'post-syntactic' counterparts of extraction in languages such as Chinese and Japanese, which exhibit no 'overt' dislocation of question phrases. A distinctive feature of Chomskyan accounts from the early 1980s is the idea that '*wh*-phrases in these languages move covertly, to occupy the same position in LF as they do in English' (71f.). As C&J note, covert movement 'was motivated by the claim that the distribution of such in situ *wh*-phrases is constrained by the same principles as overt movement'. However, they also observe that the empirical support for post-syntactic movement is greatly weakened by discrepancies between overt and LF movement.

> [E]vidence appeared quite early on … that the constraints on *wh* in situ are indeed not identical to those on extracted *wh*, undermining the force of the empirical arguments for LF considerably. (72)

Efforts to generalize constraints on extraction also had a more subtle effect on the development of Chomskyan approaches, as they fostered a division of labour between unrestricted movement rules and conditions that 'filtered' rule output. As the task of constraining movement became increasingly central to the transformational agenda, there was an accompanying narrowing of focus to syntactic phenomena perceived as relevant to that task. Correspondingly, the traditional goal of broad descriptive coverage receded in importance, and the formal architecture of the language faculty reflected a progressively narrower focus. By Chomsky (1981), transformational accounts dedicated an entire grammatical 'component' to locality (Bounding Theory), devoted a second component to anaphoric dependencies (Binding Theory), and expropriated the terms 'Case' and 'government' for theory-internal notions that bore little formal or functional relation to their traditional namesakes. The enhanced prominence of pronominal items, which had scarcely figured in the models of Chomsky (1957) and Chomsky (1965), was particularly striking, especially given that these items form small inventories, with fairly circumscribed – hence, presumably, learnable – patterns of distribution. But the importance of anaphoric dependencies derived ultimately from more theory-internal factors.

A critical step was the idea that 'filler-gap' dependencies could be assimilated to cases of binding such that the 'filler' functions as an antecedent, and the 'gap' as an anaphor or pronoun (Chomsky 1976) By treating extraction as a type of anaphoric dependency between a dislocated 'operator' and a 'variable' in an extraction site, Chomskyan accounts arrived at a solution to the longstanding problem of 'derived constituent structure'. As had been noted by Less (1957) and Stockwell (1962), early Chomskyan models defined

transformations as mappings from phrase markers to phrase markers. Yet the rules formulated in these models specified only the ORDER of constituents in the output, with no information about the structure of these constituents. The notion that extraction constructions contained a dislocated filler 'binding' a gap provided the first general solution to the problem of assigning derived structure to transforms, though this solution is by no means empirically unproblematic (Blevins 1994).

Working backwards through these developments, one can see the formulation of constraints on extraction, and the subsequent efforts to isolate those constraints, as decisive steps in the evolution of Chomskyan approaches. But instead of yielding new insights about language, the focus on movement led, via a succession of models with steadily decreasing empirical coverage, to a programme in which it seems appropriate to explore the question 'How "perfect" is language' (Chomsky 1995: 221). The opposing view that C&J present is that extraction constructions are learned, like other constructions (Tomasello 2003, Goldberg 2006), and that the constraints on extraction in a given language reflect a process of conservative generalization from the patterns that are established in the language. From this perspective, the study of a priori notions of 'perfection' and deviations from this ideal have as little to do with the study of language as they do with biology. C&J do not deny that their alternative may be regarded as 'uninteresting' by syntacticians who have imprinted on the Chomskyan aesthetic. But they maintain that the issue is an empirical one and that the available evidence – developmental as well as syntactic – supports their position.

2.2 'Hidden' levels and honest accounting

More generally, as the other case studies in *SSyn* show, the quest for universal principles linked by a rich deductive structure has led to steadily declining empirical coverage, as a small set of 'core' phenomena have assumed a progressively larger role, culminating in the aptly termed 'Minimalist' accounts. A key role in the discrimination of core from 'peripheral' phenomena is played by what C&J term 'hidden' levels of analysis. The postulation of a 'pre-syntactic' level of 'deep structure' (Chomsky 1965) or 'D-structure' (Chomsky 1981) may initially have reflected a genuine belief that 'the structures for which universality is claimed may be quite distinct from the surface structures of sentences as they actually appear' (Chomsky 1965: 118). Yet the hope of using transformational analysis to discover an underlying unity remained almost completely unfulfilled. The earliest transformational accounts were acknowledged to be far too construction-specific: a descriptive apparatus designed to classify strings and tree structures (technically string sets) turned out to be ill-suited to the task of capturing common properties across constructions in different languages, as argued with particular cogency in the early Relational Grammar literature (Perlmutter 1983). By the

late transformational period, the search for a deep commonality had effectively been called off, and, as *SSyn* points out in chapter 3, the universal properties of underlying structures came to be imposed by fiat, with only the most cursory empirical justification. The level of 'logical form' or 'LF' (Chomsky 1976) was never suited to the task of discovering common patterns, given that its structure was designed to reflect specific assumptions about extraction constructions. Furthermore, as C&J show, the arguments advanced for LF were undermined not only by discrepancies between 'overt' and 'covert' movement but also by the fact that the use of 'covert movement … to syntacticize quantifier scope … cannot account for the full range of phenomena' (533).

As C&J acknowledge, their own model incorporates a hidden level of sorts, in the form of a 'grammatical function (GF) tier' that associates dependents with a ranked set of grammatical functions. The elements of the GF tier are more like the argument structure (ARG-ST) terms of HPSG accounts (Pollard & Sag 1994). than like the grammatical relations of RG accounts (Perlmutter 1983). or the grammatical functions of LFG accounts (Kaplan & Bresnan 1982), in that they 'are NOT explicitly labeled subject, object, and indirect object' but 'are just ranked positions' (191). However, the distinctions expressed by the grammatical function tier are broadly similar to those expressed in the relational networks of RG, in the functional structures of LFG, or, indeed, in traditional grammars (Jespersen 1937). Moreover, because the grammatical function tier classifies the functions of surface structures, it does not shift the locus of grammatical analysis away from observable structures onto more abstract, idealized representations, such as underlying D-structures or post-syntactic LF-structures.

The sustained argument in *SSyn* against idealized 'hidden' levels that diverge significantly from observable structures thus echoes the verdict reached by Hockett (1987) in his discussion of 'the great agglutinative fraud'. Although Hockett had been instrumental in developing 'item and arrangement' analysis, he came to regard it as an empty formal exercise in which an agglutinative structure was foisted onto a language by 'devising an "agglutinative analog" of the language and formulating rules that would convert expressions in that language into the shapes in which they are actually uttered'(Hockett 1987: 83). In the continuation of this passage, Hockett notes the success of the agglutinative ideal in jumping hosts and establishing itself within the generative paradigm, where it fostered a more abstract and theory-internal style of analysis. Theoretical analyses ceased to address the facts of a given language directly, but instead constructed idealized analogues, which could then be brought into correspondence with the actual patterns in the language. Hence, the focus of linguistic analysis shifted from the study of patterns of cross-linguistic variation to the study of the properties of abstract analogues and the nature of the devices needed to mediate between these idealized forms and attested surface patterns.

[T]he promotion of 'theoretical underlying forms' to the status of hidden realities was carried over into syntax in the earliest transformational-generative theory, diverting attention from problems of real substance, concealing but not eliminating the unstated assumption of underlying agglutination … Even today, many of the issues to which linguists in the Chomskyan tradition devote their attention are, it seems to me, no more than artifacts of their frame of reference. (Hockett 1987: 84)

The contemporary relevance of Hockett's remarks is confirmed by the many cases discussed in *SSyn* in which theoretical artifacts divert attention from issues of real substance. The treatment of constituent structure offers an instructive range of cases and issues. C&J note that the late transformational preference for binary branching structures rests on the extremely slender evidence offered in Kayne (1983). They also stress that arguments based on the putative 'restrictiveness' of binary branching must be assessed in the context of a 'cost-benefits analysis' over the grammar as a whole. A 're-striction' in one part of the grammar plainly achieves little if it imposes limitations that must be overcome by greater complications elsewhere. As C&J observe, constraints on branching present a particularly direct trade-off, given that 'simplifying the branching possibilities requires the tree to have more nodes, and reducing the number of nodes requires more branch-ing possibilities' (113). This point brings out a recurring theme of *SSyn*, namely that the evaluation and comparison of analyses should be guided by a principle of 'honest accounting' that counts global as well as local consequences of analytical choices. Although C&J are surely right about the virtues of honest accounting, the fact that they feel compelled to propose such a self-evident principle at this stage goes some way toward explaining why the generative paradigm has failed to live up to its early promise.

Significant as it was, the adoption of binary branching was only one factor in the development of increasingly abstract constituent analyses in 'Principles and Parameters' (P&P) models (Chomsky 1991). Binary branch-ing structures pushed the syntactic analyses in P&P models away from the relatively flat, 'bushy' trees which were characteristic of earlier transform-ational models and are retained in traditional grammars and constraint-based frameworks. By reducing branching options, P&P models incorporated a larger number of nodes into increasingly 'spindly', bamboo-like trees. These structural developments were reinforced by shifting assumptions about the content of individual nodes.

Reflecting their origins in the distributional model of Harris (1951), the models of X-bar theory derived from Chomsky (1970) regulate only in-formation about word class within a phrase. So the initial Chomskyan models of X-bar theory enforced a match between the category (N, V, A, or P) of an endocentric phrase and its head, but did not regulate any other grammatical features, such as tense or agreement. This restriction on the type

of information governed by the X-bar conventions fed the use of constituent-structure displacements to distribute other types of feature information, notably via the advancement of heads to higher head positions or movements linking heads to 'specifier' slots. However, the unbundling of the former 'INFL' node proposed in Pollock (1989) offered a different way around the restrictions imposed on X-bar models. By treating 'functional categories' as abstract HEADS, P&P models invoked the X-bar conventions to regulate the distribution of tense, aspect or agreement features. Reifying morphosyntactic features as heads is consistent with the distributional approach to syntax that underlies X-bar theory, as these heads rehabilitate the 'abstract morphemes' that played a significant role in the distributional model of Harris (1951). At a more technical level, recasting inflectional features as functional heads allows the X-bar conventions to regulate a larger class of features. But this coercion is only necessary because the class of X-bar features was arbitrarily restricted in the first place. As the X-bar models developed in GPSG (Gazdar et al. 1985) and HPSG (Pollard & Sag 1994) show, traditional notions of endocentricity can be captured by generalizing the X-bar conventions to regulate a broader class of 'head' features. Classifying person, number and gender as nominal head features permits generalized X-bar conventions to accommodate constructions in which agreement features of a noun are preserved on the noun phrase that it heads. Classifying tense, aspect and mood as verbal head features likewise allows the inflectional features of a verb to be shared with the verb phrase or clause that it heads.

A principle of honest accounting would dictate that any benefits obtained by restricting the X-bar conventions to word-class features should be balanced against the cost of reclassifying entire inventories of morphosyntactic properties as word-class features. Constraining the X-bar conventions so that they only regulate word-class features has no meaningful effect if there are no restrictions on which properties can be classified as word-class features. Similar cost-benefit calculations apply to other division of labour choices. There is plainly a trade-off between the density of feature information present at individual nodes and the number of nodes required to represent a fixed quantity of information. Any analysis must face the substantive challenge of identifying the features that are distinctive in a system and assigning appropriate features to particular items or constructions. But the way that these features are distributed over a representation will tend to reflect more general considerations. Constituent structures can be simplified by increasing the density of feature information associated with individual nodes. Or the complexity of individual feature structures can be reduced by distributing simpler feature 'bundles' over a larger number of nodes. Given that comparable grammatical information can be distributed over spindly or bushy trees, one might have expected some discussion in the P&P literature regarding the formal or empirical benefits of the spindly variety. Yet, as

becomes clear from C&J's attempts to impose a coherent narrative on the developments within late transformational models, there is almost no discussion of this kind.

C&J bring out even more striking cases of accounting lapses in their summary of 'VP-shell' analyses (in which verbs are evacuated from their initial positions and advanced into higher 'shells' in a layered VP). From their origins in Larson (1988) as arboreal reimplementations of Montague grammar accounts (Bach 1979), VP-shells became progressively more influential in transformational models, extending to patterns traditionally regarded as derivational. C&J review published criticisms of the initial proposals, including the key objection that the use of shells serves mainly to preserve the view that anaphoric dependencies are subject to hierarchical, not linear, constraints. They then consider the remarkable extension of VP-shells in Hale & Keyser (1993), in which a sentence with a denominal verb such as *shelves* is treated as containing an underlying noun *shelf* that becomes a verb in the course of a syntactic derivation. C&J run through some of the most glaring deficiencies of this account, a number of which appear to be freely conceded by its proponents. Specifically, the account does not attempt to explain which nouns will participate in the alternation, what meaning changes may be effected by the derivation or even whether there will be any phonological effects (as in the voicing of *shelf* to *shelve*). The even-handed tone that C&J maintain in their review only makes their criticisms more devastating: there is no escaping the fact that the Hale & Keyser account fails to address basic questions that it would be expected to answer. Just as extraordinary is the fact that the account does not set out to solve any outstanding problems whose resolution might that would balance the costs associated with the substantial problems that it itself creates.

SSyn attributes the acceptance of the Hale & Keyser account to its use of an established technology (VP-shells) and to the fact that it 'purports to explain semantic parallelism without having actually to invoke a structured theory of semantics'(56). The first of these factors is of additional interest for the way that it highlights the pivotal role that 'exemplary analyses' have played in Chomskyan accounts. As C&J note in section 2.2, the treatment of active-passive alternations provided a template for early transformational accounts and subsequent exemplary analyses served a similar function in later versions. Transformational analyses of passive constructions exemplified the notion of 'significant linguistic generalization' in the relation between active and passive clauses, and clarified the role of underlying structures and derivational strategies. But even in early (and comparatively explicit) models, the use of a 'canonical' account to exemplify principles of analysis was not unproblematic, and contributed to what are sometimes known as the 'linguistic wars' (Harris 1993). At least part of the rancour generated by what ought to have been an intellectual dispute derived from the fact that proponents of Generative Semantics regarded their proposals as

legitimate extensions of exemplary transformational analyses. Starting from the assumption – widely accepted at the time – that underlying structures were the primary locus of semantic interpretation, the Generative Semantic programme explored the hypothesis that underlying structures were themselves semantic representations. By expanding the scope of transformations (which mapped semantic representations onto surface structures), this hypothesis seemed to extend earlier exemplary analyses, which had shifted the division of grammatical labour away from Immediate Constituent structures to transformational analyses. The 'lexicalist' response in Chomsky (1970) did not so much refute the Generative Semantics positions (Huck & Goldsmith 1995) as establish a different exemplary model, based on analyses of derived and gerundive nominals.

Moreover, as transformational models became more programmatic, the task of inducing principles from 'canonical' individual analyses became substantially harder. By Chomsky (1993), the reader is faced with the daunting challenge of trying to 'reverse engineer' constitutive principles from exemplary analyses that consist largely of implementation details, key terms in scare quotes and little actual linguistic data. Although C&J offer a characteristically charitable assessment of the 'complex and impressive' (56) achievements of Chomskyan approaches, their account of the evolution of 'mainstream' models makes sobering reading. From an intellectual movement that once seemed destined to revolutionize the study of language in much the way that plate tectonics revolutionized geology, Chomskyan theories have become increasingly detached from the concerns of traditional descriptive and theoretical approaches as well as from those of allied disciplines. A central question raised by this volume is whether the flaw lies in the execution of the Chomskyan programme, or whether the basic conception is misguided.

3. THE SIMPLER SYNTAX HYPOTHESIS

As *SSyn* shifts from the diagnosis of problems to an exploration of the 'Simpler Syntax Hypothesis', C&J make a spirited case for the view that mainstream accounts represent a flawed implementation of what is actually a viable research programme. The general architecture of their alternative has much in common with the design of constraint-based formalisms such as LFG, GPSG (Gazdar et al. 1985) or HPSG (Pollard & Sag 1994). Like these frameworks, the Simpler Syntax model is constraint-based rather than derivational, posits no 'hidden' syntactic levels (apart from the grammatical function 'tier' discussed in section 2), and distributes combinatorial complexity throughout the components of the grammar. In common with construction-based approaches such as Goldberg (2006), the Simpler Syntax model also recognizes a continuum of regularity, from completely idiosyncratic expressions to fully regular 'rules'.

The core intuition of this model is expressed by what C&J call the 'Simpler Syntax Hypothesis' (SSH). In its most concise form, this hypothesis states that

> [t]he most explanatory syntactic theory is one that imputes the minimum structure necessary to mediate between phonology and meaning. (5)

Determining exactly what should count as 'structure' and how much of this structure is 'necessary' are both highly contentious matters, as is the notion that syntactic structure is motivated exclusively by its role as a conduit for meaning.

However, the general import of the SSH can be seen in somewhat more neutral terms as a hypothesis about the place and scope of the syntactic component. With respect to grammar design, the SSH favours an architecture in which the syntactic component has the same status as other linguistic components, so that a given expression receives an analysis in parallel in the phonological, morphological, syntactic and semantic components. This conception differs from what C&J term the 'syntactocentric' design of transformational approaches, in which the syntactic component has a privileged combinatoric role and other components merely interpret the output of the syntax. With respect to the classification of phenomena, the SSH favours reducing the scope of the syntactic component by shifting a larger part of the descriptive and explanatory burden onto non-syntactic components and factors. To illustrate the effects of this revision, C&J offer an analysis of Bare Argument Ellipsis (BAE) in which they argue that acceptability reflects semantic and/or pragmatic considerations.

More generally, C&J argue that the SSH is motivated by its empirical effects rather than by the type of a priori conjectures about 'optimal' design features that guides some Minimalist accounts. The initial discussion of BAE sets the general tone of the argumentation, as *SSyn* runs through analyses of a comparatively wide range of phenomena, including discontinuous dependencies, control and raising constructions, various types of ellipsis and anaphora and a number of idiosyncratic construction types. The breadth of construction types covered is reminiscent of the transformational literature of the late 'Standard Theory' (as reflected in textbooks of the period, such as Soames & Perlmutter 1979) and contrasts with the narrower conception of 'core' constructions adopted in subsequent Chomskyan accounts. Many of the individual proposals merit a more detailed discussion than can be given in a review article of this nature, though it is possible to comment on some of the properties of the analyses as a group.

Overall, the strengths of the Simpler Syntax model tend to concentrate in the areas where it develops insights from outside the transformational paradigm, while its weaknesses cluster in those areas where it is most conservative. When C&J argue persuasively against derivational mechanisms in general and derivational treatments of discontinuous dependencies in

particular, they reinforce a position that has been held for over 20 years within the constraint-based tradition. Their arguments for a parallel architecture that distributes computational complexity through the grammar support an equally well-established alternative. In developing a construction-based treatment of various syntactic 'nuts', *SSyn* similarly places itself within the constraint-based tradition and breaks with the transformational view of constructions as epiphenomena. C&J likewise set out clear and compelling grounds for preferring an analysis of elliptical constructions based on what they term 'indirect licensing' over an account that treats elliptical expressions as remnants of full sentences. However, the claim that indirect licensing 'is discourse-based, i.e. it is not strictly a part of sentence grammar' (298) appears to endorse transformational ideas about the primacy of sentence grammar, which raises questions about the theory of discourse in which indirect licensing is embedded and perhaps invites a reconsideration of the relation between discourse grammar and sentence grammar.

The metaphor of a 'tinkerer' (5) that C&J suggest for the language faculty (based on an earlier characterization of biology) applies equally to aspects of the Simpler Syntax programme. A number of the formal proposals in *SSyn* are more tactical than strategic, in that they endorse Chomskyan objectives while taking issue with the manner in which problems are approached or solutions are implemented. The treatment of constituent structure provides a clear case in point. In common with many transformational models, C&J 'take syntactic structure to be a linearized hierarchical tree structure whose nodes consist of syntactic features' (108). They go on to add that they 'do not take the terminal nodes in the tree to be full lexical items' but instead interpret the arc between a preterminal such as N and a terminal such as *dog* as meaning that N 'is LINKED to corresponding phonological and semantic structures' (109, emphasis added). Although, as C&J note, a 'linking' interpretation fits with the parallel architecture adopted in *SSyn*, it is far less clear what is at stake here.

Any serious attempt to characterize the mapping between syntactic structures and the lexicon raises a host of fundamental questions – not the least of which concerns the adequacy of the post-Bloomfieldian model of morphology that *SSyn* assumes. But it is hard to see how any substantive issues are affected by implementation-level decisions about whether lexical items are 'dominated' by preterminals in an integrated data structure or treated as the values of 'pointers' from preterminals. These types of 'assembly-level' details are the sort of theoretical 'artifacts' that Hockett (1987) was referring to, especially when – as in the present case – they are associated with no discernible empirical consequences.

Perhaps the most obvious weakness of *SSyn* is one that C&J freely acknowledge, namely that its analyses deal almost exclusively with English, reflecting the authors' primary expertise. Motivating a theory on the basis of

data from a single language reflects the generative idea that sufficiently close study of one language can yield insights into the general language faculty. The hazards of this strategy may be more evident in domains such as morphology, where the patterns in English provide a poor guide to the attested variation in the world's languages. But in the domain of syntax, one can also discern the distorting influence of English on 'universal' constraints that require the presence of subjects, or on theories of voice alternations that treat passives as intrinsically promotional, to mention only two obvious cases. Hence, an approach to syntax that aims to correct the harmful idealizations and outright errors of transformational accounts might start with a substantially wider empirical base.

As far as syntactic argumentation is concerned, *SSyn* tends to comply with the requirements of honest accounting. C&J offer cost-benefits analyses for what they regard as the most plausible alternatives and argue directly for particular choices. The style of argumentation is reminiscent of the transformational literature from the mid-1960s through the late 1970s, in which a range of converging evidence is marshalled in support of a theoretical claim. Although this type of argument is more explicit and transparent than those that proceed from conjectures about optimal language design, it also has a somewhat anachronistic character in the current context. Given the greater availability of corpora, tools for corpus analysis, and even corpus-based grammars, such as Biber et al. (1999), syntacticians now have the opportunity – as well as the responsibility – to base their analyses on a more secure empirical foundation than their own introspection. The use of this kind of data not only offers a current snapshot of a language but also provides a means of validating the judgements on which syntactic theories are still based, thereby offering a measure of protection against the experimenter bias, intentional or not, that may devalue those judgements.

Overall, *SSyn* can be seen as the fruit of a theoretical convergence whose main preconditions were met some 25 years ago, but which did not – for a variety of reasons – happen then. By the early 1950s, the Bloomfieldians had arrived at an essentially generative conception of grammar, according to which

> [a] grammatical description … sets forth principles by which one can generate any number of utterances in the language; in this sense, it is operationally comparable to that portion of a human being which enables him to produce utterances in a language; i. e., to speak. (Hockett 1954: 390)

Early Chomskyan models represented the first attempts to formalize this conception, though their design was influenced by the somewhat distinctive division of grammatical labour in the transformational model of Harris (1957). In particular, Chomskyan approaches took over the notion

that the representational limitations of simple constituent analyses were best overcome by elaborating the control structure of the grammar rather than by relaxing representational constraints on individual representations. (This reflects the same formal aesthetic as does the preference for increasing the number of nodes in a spindly tree, rather than increasing the information content of each node in a bushier tree.) As a consequence, transformations came to perform an extremely wide range of functions in Chomskyan models, from inserting lexical items, to capturing valence alternations, to expressing local and non-local word order alternations.

By the early 1980s, it had become clear that alternative accounts of these phenomena were available and that alternative divisions of labour were possible within a formal grammar. Gazdar (1981) had shown how transformational analyses of extraction constructions could be mimicked by phrase structure grammars with complex nonterminals and elementary 'feature passing' mechanisms. At about the same time, Bresnan (1982a, b) demonstrated that a richer model of the lexicon permitted a lexical treatment of valence-changing operations and analyses of control and raising constructions. These developments were particularly compatible with the 'representational' interpretation of movement that had been proposed by Koster (1978), and suggested a natural rapprochement between transformational accounts and the emerging constraint-based approaches. There was even a sense in which constraint-based approaches were a natural development of the lexicalist proposals in Chomsky (1970).

Yet, as it turned out, Chomskyan models lurched in the direction that made them maximally incompatible with the emerging alternatives, much as they had a decade earlier in response to the Generative Semantics challenge. The role of derivations was considerably enhanced through the use of movement to effect agreement and through the introduction of various new types of head movement. The P&P and Minimalist approaches also rehabilitated a body of syntactic analyses which 'many researchers have observed … are hauntingly reminiscent of Generative Semantics of the late 1960s and early 1970s' (95). One can get a flavour of this style of analysis from the Hale-Keyser account in section 2.2 above, in which a noun becomes a verb in the course of a syntactic derivation, though the 'classic cases' summarized in section 3.3.2 of *SSyn* provide a more systematic presentation. Taken together, these changes drove Chomskyan models along a path of development that led away from convergence with constraint-based approaches and towards the abstract syntactocentric analyses criticized at length in *SSyn*. The alternative outlined in *SSyn* represents a 'third way' that consolidates what C&J see as the strengths of Chomskyan and constraint-based approaches. If C&J are right in regarding Chomskyan approaches as flawed implementations of a sound underlying conception, the Simpler Syntax model may well suggest a way forward for a subject which appears

increasingly unable to forge a consensus regarding what constitutes the relevant data for a theory of syntax, let alone agree on criteria for evaluating competing analyses.

4. THE POST-GENERATIVE ENTERPRISE?

Despite the optimistic perspective adopted in *SSyn*, it is not always easy to share C&J's conviction that syntax rests on a 'complex and impressive edifice' or that the problems they have identified are susceptible to technical solutions.

At the most fundamental level, it is not clear that there is any meaningful empirical motivation for the representational assumptions of any current formal model of syntax. While it is reasonably well established that speakers are aware of certain patterns in their language, there is at present no reliable methodology for probing the format in which that knowledge is represented or manipulated. Generative approaches accordingly construct symbolic representations that are presumed to model the 'content' of a speaker's knowledge, in the belief that the structure of this idealization may lend some insight into the structure of the knowledge that a speaker exploits in actual language use. But these representations are massively underdetermined by what is actually known. For example, in their discussion of the very minor system of geographical names in English, C&J identify four ordering patterns, which are extended to new terms, as in the minimal pair '*Gloggle Mountain* but *Mount Gloggle*' (29). Even in this trivial case, we do not at present have any way of determining whether speakers encapsulate geographical patterns symbolically in minor schemas, templates or rules, or whether they analogize from a stock of names in their mental lexicon. The problem is only worse with more general or abstract patterns, since in these cases, the extent and even the content of the pattern may be indeterminate. Thus speaker knowledge of constructions can be modelled by inheritance hierarchies, as in Sag (1997), or in terms of exemplary constructions, as in traditional grammars. In this case, it is not even obvious what information a speaker associates with a given construction type, let alone what format it is stored in.

This question of representational legitimacy arises in an acute form in C&J's critique of syntactocentrism. As noted above, *SSyn* provides a sustained argument against a conception of grammar in which syntax is the sole 'generative' component, and other components serve an essentially interpretive function. Their alternative assumes that semantic and phonological representations have an independent combinatorial structure, which is brought into correspondence with syntactic structures in various ways, as in models of LFG. Yet the basic argument is set within a very narrow frame of reference. In particular, it assumes the fundamental legitimacy of an analysis that associates elements of idealized phonological and syntactic

representations with elements of an equally idealized representation of meaning, formulated in a variety of 'Mentalese'.

It may turn out, as C&J appear to believe, that these sorts of analytical techniques may yet provide the basis for stable results, results that can be validated by more direct methodologies for probing the format of linguistic knowledge. However, one should also be alert to the possibility that the flaws in current approaches lie not with faulty execution but with illegitimate idealizations.

Moreover, even if one assumes the soundness of a representational approach, current models are shot through with underdetermined representational choices. For example, C&J usually assume that constituent structures are continuous and take pains to avoid 'resorting … to discontinuous constituency' (143) in cases where this might seem to be an option. In this respect, they are following accepted generative practice, and departing from the analytical practices of the Bloomfieldians, as discussed in McCawley (1982), among others. Yet the justification for this bias is even less substantial than the motivation for binary branching. It is difficult to find any empirical basis for the exclusion of discontinuous constituents, other than perhaps the observation that many of the constructions that the Bloomfieldians had described as discontinuous were subsequently dealt with in terms of transformations.

The conception of the lexicon outlined in *SSyn* likewise incorporates widely held but largely unargued assumptions about permissible redundancy:

> we regard the lexicon as a long term memory repository of items that are learned and stored (possibly including predictable high-frequency material); this makes for a smooth transition between theories of competence and performance. In our view, what the lexicon should NOT include is material that can be freely constructed online through fully productive rules. (188, emphasis in original)

The notion that speakers do not store regular, productive forms in their mental lexicon is firmly entrenched in generative theories of syntax, and a guiding idea in 'dual mechanism' (Clahsen 1999) and 'words and rules' (Pinker 1999) models. But surely this is an empirical question, and the well-documented phenomenon of 'morphological family effects' (de Jong et al. 2000, Moscoso del Prado Martín et al. 2004) suggests that speakers may indeed store regular forms as wholes.

Nevertheless, these types of general concerns should not be allowed to detract from C&J's considerable achievements in this volume. *SSyn* is essential reading for anyone interested in understanding where the field of syntax has been in the second half of the 20th century, and the direction that at least parts of the field will be taking in the 21st. The even-handed comparison of current Chomskyan models with constraint-based alternatives

also builds a bridge between these traditions that may yet help to facilitate a general approchement. Furthermore, by highlighting the assumptions that these approaches share, *SSyn* may also play a role in provoking readers to conceptualize the study of syntax in the 21st century in more radical terms than C&J are themselves prepared to entertain. In particular, although C&J recognize a continuum of regularity, the grammatical descriptions they present are essentially discrete, as in Chomskyan models. If speaker sensitivity to probabilistic patterns is as pervasive as some current work would appear to suggest (Bod et al. 2003, Ernestus & Baayen 2003, Hay & Baayen 2005) the idealizations that underlie discrete descriptions may require wholesale reevaluation. In this case, the status of familiar techniques for modelling discrete variation may also be called into question and much of what C&J take to be secure results of the post-Bloomfieldian tradition may need to be reassessed in a more probabilistic or information-theoretic context.

REFERENCES

Bach, Emmon. 1979. Control in Montague Grammar. *Linguistic Inquiry* 10, 515–531.
Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan 1999. *Longman grammar of spoken and written English*. London: Longman.
Blevins, James P. 1994. Derived constituent order in unbounded dependency constructions. *Journal of Linguistics* 30, 349–409.
Bod, Rens, Jennifer Hay & Stefanie Jannedy (eds.). 2003. *Probabilistic linguistics*. Cambridge: Cambridge University Press.
Bresnan, Joan. 1982a. Control and complementation. *Linguistic Inquiry* 13.3, 343–434.
Bresnan, Joan. 1982b. The passive in lexical theory. In Bresnan (ed.), 3–86.
Bresnan, Joan (ed.). 1982c. *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
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. 1970. Remarks on nominalization. In Jacobs, R. A. & Rosenbaum, P. S. (eds.), *Readings in English transformational grammar*, 184–222. Waltham, MA: Ginn and Company. [Reprinted in *Studies on Semantics in Generative Grammar*, 1–61. The Hague: Mouton.]
Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart and Winston.
Chomsky, Noam. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2, 303–351.
Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
Chomsky, Noam. 1982. Some concepts and consequences of the theory of Government and Binding. Cambridge, MA: MIT Press.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freiden (ed.), *Principles and parameters in Generative Grammar*, 417–454. Cambridge, MA: MIT Press.
Chomsky, Noam. 1993. A Minimalist program for linguistic theory. In Hale & Keyser (eds.), 1–52.
Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.
Clahsen, Harald. 1999. Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences* 22, 991–1060.
Culicover, Peter W. 1999. *Syntactic nuts*. Oxford: Oxford University Press.

de Jong, Nivja H., Robert Schreuder & R. Harald Baayen. 2000. The morphological family size effect and morphology. *Language and Cognitive Processes* 15, 329–365.

Ernestus, Mirjam & R. Harald Baayen. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79, 5–38.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12, 155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Goldberg, Adele. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Hale & Keyser (eds.), 53–109.

Hale, Kenneth & Samuel Jay Keyser (eds.). 1993a. *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press.

Harris, Randy Allen. 1993. *The linguistics wars*. Oxford: Oxford University Press.

Harris, Zellig S. 1951. *Methods in structural linguistics*. Chicago: University of Chicago Press.

Harris, Zellig S. 1957. Co-occurrence and transformation in linguistic structure. *Language* 33, 283–340.

Harris, Zellig S. 1965. Transformational theory. *Language* 41, 363–401.

Hay, Jennifer & R. Harald Baayen. 2005. Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences* 9, 342–348.

Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10, 210–231. [Reprinted in Martin Joos (ed.), 1957, *Readings in linguistics I*, 386–399. Chicago: University of Chicago Press.]

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.

Hockett, Charles F. 1987. *Refurbishing our foundations*. Philadelphia: John Benjamins.

Huck, Geoffrey J. & John A. Goldsmith. 1995. *Ideology and linguistic theory: Noam Chomsky and the deep structure debates*. London: Routledge.

Jespersen, Otto. 1937. *Analytic syntax*. London: Allen & Unwin.

Johnson, David E. & Shalom Lappin. 1999. *Local constraints vs. economy*. Stanford: CSLI.

Kaplan, Ronald M. & Joan Bresnan. 1982. Lexical-functional grammar: A formal system for grammatical representation. In Bresnan (ed.), 173–281.

Kayne, Richard S. 1983. *Connectedness and binary branching*. Dordrecht: Foris.

Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Foris.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19, 335–391.

Lees, Robert B. 1957. Review of *Syntactic structures*. *Language* 33, 375–408.

Matthews, Peter H. 1993. *Grammatical theory in the United States from Bloomfield to Chomsky*. Cambridge: Cambridge University Press.

Matthews, Peter H. 2001. *A short history of structural linguistics*. Cambridge: Cambridge University Press.

McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13, 99–107.

Moscoso del Prado Martín, Fermín, Raymond Bertram, Tuomo Haikio, Robert Schreuder & R. Harald Baayen. 2004. Morphological family size in a morphologically rich language: The case of Finnish compared to Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory and Cognition* 30, 1271–1278.

Newmeyer, Frederick J. 2005. *Possible and probable languages: A generative perspective on linguistic typology*. Oxford: Oxford University Press.

Perlmutter, David M. (ed.). 1983. *Studies in Relational Grammar 1*. Chicago: University of Chicago Press.

Pinker, Steven. 1999. *Words and rules*. London: Weidenfeld & Nicholson.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Stanford: CSLI & Chicago: University of Chicago Press.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20, 365–424.

Postal, Paul M. 2004. *Sceptical linguistic essays*. Oxford: Oxford University Press.

Pullum, Geoffrey K. & Barbara Scholtz. 2002. Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19, 8–50.

Ross, John R. 1967. *Constraints on variables in syntax*. Ph.D. dissertation, MIT. [Published 1986 as *Infinite syntax!* Norwood, NJ: Ablex.]

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 30, 431–483.

Seuren, Peter A. M. 2004. *Chomsky's Minimalism*. Oxford: Oxford University Press.

Soames, Scott & David M. Perlmutter. 1979. *Syntactic argumentation and the structure of English*. Berkeley & Los Angeles: University of California Press.

Stockwell, Robert P. 1962. Discussion of 'A transformational approach to syntax'. In Archibald A. Hill (ed.), *The Third Texas Conference on Problems of Linguistic Analysis in English*, 158–169. Austin, TX: University of Texas.

Tomasello, Michael. 2003. *Constructing a language*. Cambridge, MA: Harvard University Press.

Author's address: *Research Centre for English and Applied Linguistics,*
University of Cambridge, 9 West Road, Cambridge CB3 9DP, U.K.
jpb39@cam.ac.uk