

John Benjamins Publishing Company



This is a contribution from *Chomskyan (R)evolutions*. Edited by Douglas A. Kibbee.
© 2010. John Benjamins Publishing Company

This electronic file may not be altered in any way.

Permission is granted by the publishers to post this file on the author's webpage and/or institutional server, to distribute it through mailing lists and to use it in any other way that serves the promotion of the publication.

Tables of Contents, abstracts and guidelines are available at
<<http://www.benjamins.com/>>www.benjamins.com

Noam and Zellig

Bruce Nevin

In this paper, we explore when and how the work of Noam Chomsky diverged from that of his mentor, Zellig Harris, and identify the origins and character of their differences. Considering evidence that they never fully understood each other, the rhetorical vehicle for this exploration is speculation as to how much this divergence is due to differences of temperament, to “post-war” generational differences, or, with more importance for the field, to different conceptions of the proper conduct of science. A number of questions are addressed, such as: Is truth attained by winning arguments, or is the appropriate role of argument to do everything possible to prove oneself wrong? Is a theory of language a prerequisite to or an outcome of linguistic analysis? And what are the fundamental data of linguistics?

Here at the outset, please understand that by the use of first names in the title and throughout this paper I mean no affront or disrespect for either Zellig Harris or Noam Chomsky, but rather the intimation that we are talking about them as persons. Discussion of their respective achievements cannot entirely avoid the quasi-mythical personae that have become institutionalized around them in public discourse, so it has seemed well to place the divergences in their work more accurately in context of their relationship. The central theme of this paper is that they never fully understood each other, due to differences that we may variously ascribe to personality, or temperament, or cognitive style. This is an essentially irenic supposition, but to demonstrate and give substance to it requires several excursions into regions where it is possible that responses to what I say may be less pacific. Let me make explicit here, for it is implicit throughout, that the same considerations of comprehension and incomprehension apply as well to those responses. I know this does not apply to you, dear reader, whose perspicacity and fair-mindedness are unquestioned; it is those other readers that concern me.

I must consequently ask your patience with the cyclical structure of this paper. It is a truism of psychology that we understand in terms of what we have previously understood, and sometimes perhaps we understand too quickly. In order that they might insinuate themselves through this perfectly normal cognitive hedgerow, themes and topics recur in a recycling kind of way, each time from a different angle, with a different emphasis, or in a different combination.

That said, let us begin at the beginning.

1. B'Reshit¹

In 1945 or 1946, William Chomsky (1896–1977) asked Zellig Harris (1909–1992), a neighbor and family friend teaching and doing linguistic research as a Professor at the University of Pennsylvania, to take his brilliant but erratic son Avram Noam (b. 1928) under his wing.² Both families had immigrated in the first decade or so of the century (Zellig's family when he was four years old) from the terrible political and social conditions in Ukraine to this tight-knit neighborhood in Philadelphia. Zellig had obtained his degrees at Penn as a specialist in Semitic linguistics in 1934 and 1936, then taught and continued research there, becoming Assistant Professor in 1943. Funding for war-time and post-war linguistic research enabled him to develop a wide range of courses encompassing his many research interests.

It was literally the case that in the beginning years Zellig was the department, and its formation involved no more than bringing together, as a program in 'linguistic analysis,' a number of courses that Zellig had developed during the war years under the aegis of either the anthropology or the Oriental Studies Departments of the University. (Hiž n.d.).

He established the new Department of Linguistic Analysis in 1946 and was advanced to Full Professor in 1947. The concern that Mr. Chomsky brought to his friend was that young Noam had just entered Penn as an undergraduate, but was impatient with his classes, and was considering dropping out. Maybe he would move to a kibbutz in Israel. Noam was then about 17 years of age, Zellig in his late thirties. The Harrises essentially took the teenager into their family, sharing meals with him and so on. Zellig supported and encouraged him to study mathematics, logic, and philosophy, sponsoring him to

-
1. The first two words of Genesis in Hebrew, usually translated "In the beginning".
 2. Noam has said that his "formal introduction to the field of linguistics was in 1947, when Zellig" gave him the proofs of Harris (1951), but in the preface that was signed in January 1947 he is credited for help with reading those proofs. Either his memory is wrong, or the preface was revised after the signature date. In any case, he almost certainly means the manuscript rather than the proofs, that is, galley proofs from a publisher, which would not have been available until perhaps 1950. Harris 2002.2 says the book was "completed and circulated in 1946, though it appeared only in 1951", so copies of the manuscript were available. Zellig seems always to have had numerous manuscripts in progress concurrently, and considerable time could elapse before publication, particularly with the austerities of the war years. See for example (Wells 1947: 81n1): "The central importance of the problem of immediate constituents was driven home to me in many valuable conversations with Zellig S. Harris, who also let me read a number of his manuscripts, of which not all have yet been published." This was his continued practice with students. In the late 1960s, I and other students had copies of the manuscript of Harris (1968), which we discussed in seminars with him, and of course copies of various papers and monographs.

excellent teachers, including Nathan Fine (1916–1994) for mathematics, and Nelson Goodman (1906–1998) and Richard Martin (1916–1985) in philosophy and logic.³

What was it like, having this unusual young man in their home? One family member⁴ says “He would follow you from room to room arguing, arguing, arguing you down to dust!” and also told me, with sadness, that they tried for many years “to establish a human relationship with him”. These clues, while conclusive of nothing, suggest a disparity of temperament. The aim of this paper is to explore how differences in temperament suggested by these and other indicators may underlie fundamental differences in their ways of working and help to explain the sundering of the results of their work and the peculiar extension of ‘graduate student amnesia’⁵ through an entire academic lifetime.

2. Two uses of argument

It can hardly be doubted that Noam is a master of debate. His extraordinary prowess as a polemicist is well documented. More, it is very important to him to *win* arguments.

3. The Wikipedia article on Richard Milton Martin has this relevant passage about an important influence on Noam who has been hitherto invisible in the standard accounts:

Martin was especially fond of applying his first-order theory to the analysis of ordinary language, a method he termed logico-linguistics. He often referenced the work of the linguists Zellig Harris (admiringly) and Henry Hiz (more critically); Martin, Harris, and Hiz all taught at Penn in the 1950s. Yet Martin was dismissive of the related theoretical work by Noam Chomsky and his M.I.T. colleagues and students. Ironically, Martin appears to have been Chomsky’s main teacher of logic; while a student at Penn, Chomsky took every course Martin taught.

4. Who asked not to be identified.

5. The phrase is Lila Gleitman’s:

I rapidly fell under the influence of Harris, whose thinking has guided the rest of my intellectual life. In light of that fact, I have been surprised, looking back over my own writings, to find that citations and references to Harris are conspicuously absent from most of them. ... I have tried to think about why. The answer is that so much did Harris’s approach to language get into my skin, become the sure and self-evident basis of my own thinking, as eventually to feel like my own quite clever inventions; that is, to lead to the well-known academic malady called Graduate Student Amnesia. (Gleitman 2002: 209)

Needless to say, this cutely named ‘malady’ is neither universal nor inevitable, and while it is one of the joys of teaching (and a mark of its success) when a student makes one’s ideas their own, taking them in their own directions, nonetheless a failure to credit sources is wrong, as Gleitman obviously agrees.

As Bob Ingria said to me once, “If you want to talk to Chomsky, wear boxing gloves.” He has been called an intellectual bully,⁶ and has been accused of all sorts of intellectual malfeasance, from incorporating the supposedly vanquished positions of his opponents into his own ‘revised standard view’ without acknowledgement, to peppering his argument with citations and references that cannot be checked until too late (and which then turn out, it is said, to be not quite as represented), to outright lying.⁷ And if he cannot win, the argument or the terms proposed are dismissed as unimportant, or trivial, or uninteresting.⁸ Countless anecdotes are told, and many have been published. Though such testimony abounds, I will not put any weight on what could after all be no more than the discontent of poor losers thwarted by a brilliant mind. That said, I doubt that any who have engaged with or witnessed Noam in action would gainsay the proposition that when he does enter an argument, winning it is very important to him.⁹

6. Said e.g. in Barsky, forthcoming.

7. All these points are documented in e.g. R. Harris 1993, 1998.

8. See e.g. Lin (1999) for examples.

9. Any verbatim transcript of dialog affords examples to the attentive eye. Because the issues and the progress of the encounter are especially accessible and transparent, and perhaps more importantly because the exchange is neutral with respect to linguistics, consider the interaction reported in MacFarquhar (2003: 66). The situation is that Noam has invited any student present to articulate an opposing view. One takes him up on it. He then counters with “Suppose the goal is to liberate Iraq. How come it’s not proposed in the United Nations?” The student starts to respond, “There are a lot of answers to that, like I think —” and Noam interrupts with “Really? I don’t know of any” and launches into a proposal that the US should support Iran in an invasion of Iraq. The student tries to respond with “But —” and is cut off with “Excuse me...” (not everyone is allowed to interrupt) and what turns out to be a *reductio* ploy rolls on to the rhetorical question “What’s the downside?”

The student looked baffled. “Are you honestly advocating that we help Iran invade Iraq?” he asked. “No. You are.” [...] Chomsky continued to berate the student for a long time, ignoring his attempts to break in. People cried out “Let him talk!” but to no avail. Another student stood up and called out a request that he be allowed to help, but Chomsky ignored him. People made loud, disgruntled noises in protest at this treatment, but Chomsky ignored those too. Finally, the first student sat down.

Other examples are documented in Huck & Goldsmith (1995), e.g. on p. 70 “Chomsky never directly addressed this facet of McCawley’s argument, which clearly constituted a serious problem for an Interpretive approach”, in R.A. Harris (1993, 1998) and elsewhere.

Further, Noam believes that argument is the way that you arrive at truth. An anecdote brought this home to me.¹⁰ He and a friend were discussing the conflict of Israel and the Palestinians. The friend said that it was not just a matter of a rational negotiation of interests, and that because such intense and deeply founded emotions were involved it could not be resolved by logical argument. Somewhat to the friend's surprise, this occasioned a pause, and as Noam thought about it he grew sad, saying that it may be then that it can never be resolved. This suggests that he sees argument as the *only* way (or at the least, the best and the most preferred way) to arrive at truth.¹¹

This, I submit, is a temperament well suited to philosophical disputation. I take it as no accident, then, that Noam has on occasion characterized linguistics as a branch of philosophy, too immature yet to be considered a science.¹² Let us take a brief excursion to consider the other side of this distinction which he has drawn between linguistics as philosophy and linguistics as science.

Science seeks truth without claiming to arrive at it. Scientists do this by interrogating nature and then corroborating what they find, an inherently communal process. The tension between individual exploration and collaborative confirmation is well expressed in the following excerpt from a letter that Zellig wrote to Albert Goetze (December 27, 1940):¹³

Thanks for your paper and review. I have just read them hastily, and will soon go over it point by point. Let me assure you that not only do I not consider it 'unpleasant' but am glad of the controversy. No person, certainly not I, can be

10. As told me by a friend who asked to remain anonymous, and who didn't want to indulge in psychological speculation because of not wanting to lose Noam's friendship — a telling comment in itself.

11. To those readers who share these assumptions — that the road to truth is by winning arguments (or if you're not as good at it as Noam is, then by agreeing with those arguments that have won) — I am not trying to dissuade you, nor is any denigration intended, of you or of Noam. After all, I only ask you to suspend disbelief, for the present anyway, that there may be alternative *modi operandi*; and I ask this not least because otherwise you may find this paper rather incomprehensible. Mindful attention to the kindling of counterarguments, etc. may disclose for you a kind of implicit meta-point of the paper concerning how that incomprehension works.

12. Although more recently he has considered linguistics to be a subpart of psychology, apparently without concern for arguments e.g. in *LSLT* (Chomsky 1951, 1955a, 1956a, 1975) for the autonomy of linguistics.

13. I am indebted to Robert Barsky for bringing to my attention this and other letters preserved at Yale University, by way of an unpublished ms. of his that was provided for my review by Seymour Melman.

sure of his judgments as ‘always right’; the best way to get closer to the ‘truth’ — after I have figured out whatever I could — is to get the divergent opinions which arise from a different scientific analysis. The only fun in science is finding out what was actually there.

The place of argument in this is important, but its role is distinctly secondary, properly relegated to the design of experiments and the interpretation of results. Indeed, the aim of the scientist — when doing science — is to try as hard as he can to lose the argument. This is because a working scientist knows the perils of wishful thinking. Consequently, before a finding or hypothesis is published, the scientist will, ideally, try to destroy it in every way that can be imagined, and when it is published, the presentation must point out its vulnerabilities and weaknesses, inviting others to devise challenges that the author failed to imagine.¹⁴ This is not for the sake of constructing an

14. “Science is not a monument of received Truth but something that people do to look for truth.” (Overbye 2009). Science and the arts both require patronage. Patrons expect a return on investment, to be sure, and too loose oversight of their beneficiaries can be exploited and abused. It is possible that excessive ‘enjoyment’ of military funding for linguistics in the 1960s contributed to poisoning the well. But critics of funded projects have too often demonstrated their incomprehension of the nature of science, as witness some of Senator Proxmire’s Golden Fleece Awards for ‘wasteful’ government spending. Science has means of internal discipline that are far less available to the arts, or perhaps only far less well defined, but peer review cannot be more free of ulterior motives than the scientists themselves are who constitute that peerage. *Quis custodiet ipsos custodes?*

The influence of patents and profits in the life sciences provides striking illustration. Roger J. Williams refused to patent any of his discoveries, notably pantothenic acid (one of the B vitamins), in order to ensure that no ulterior motive should undermine his integrity as a scientist, nor any imputation of such sully his reputation. Since patent law was changed in the 1980s, corporations, universities, and individual scientists now routinely patent new discoveries, even new living organisms and bits of DNA. A consequence is a shutting down of that sharing of information which is essential to science as a communal endeavor. (See e.g. Butkus 2009 for discussion.) The epidemic of Lyme and other tick-borne diseases provides a tragic demonstration of consequences that can follow. Pharmaceutical companies seek a circumscribed disease definition so they can more easily design drugs and get them through FDA trials. Insurance companies demand a narrow definition to exclude profit-sapping long-term treatment of patients with indeterminate responses. Diagnostic tests are designed to excessively narrow definitions. Researchers and physicians are paid as consultants by manufacturers of test kits as well as by pharmaceutical and insurance companies. And there are the usual all too human ego issues — my theory, my disease definition, the organism I discovered, etc. The devastation to sufferers of tick-borne diseases, a population exceeding that of AIDS victims, has until very recently been denied and relegated to baseless diagnoses (Epstein-Barr syndrome, fibromyalgia, neurosis) that have no effective treatment, and remains controversial (Weintraub 2008). It can be seen, then, that the ideals of science are serious stuff indeed, with serious consequences, and not ‘mere idealism’. We will see farther on that even

unassailable position, but in order to find out “what is actually there.” And while logic is important in these challenges, its import is as to whether or not one’s conclusions are warranted by empirical observations and sound methodology.

The operative words here are “when doing science” and “ideally”. There is no doubt that scientists also do philosophy, and that they energetically argue for their favored views as means of promoting their careers, attracting students, ensuring that students are recognized by others to be *their* students, and so on. And there are many corrosive influences making the ideals of science difficult to maintain in practice. The strengthened demand by government funding agencies and patrons of science in industry that deliverables of research be specified in advance and subsequently delivered on spec and on schedule is understandable, but at odds with the essentially exploratory nature of research, the role of serendipity, and so forth. This reflects a too-common misapprehension of the hand-off between science and engineering. A related corrosive influence is the social demand that the scientist be an Authority. Certainty is a cardinal requirement of engineering; uncertainty is indispensable for doing science. It is too little understood that science proves nothing, and that proof is possible only for logic and mathematics. The complex interrelationships of mathematics, science, and engineering have obvious pertinence to the politics of science and to our present discussion.

It has been demonstrated many times that peer review, one of the bulwarks against investigator bias, can easily become instead a bulwark of shared bias. Referees become gatekeepers. This is the meat and contested bone of any discussion of revolutions in science. The ultimate safeguard of the methods and results of science therefore rests in the integrity of the individual researcher. Confirmation bias is a well-documented hazard.¹⁵ The common or garden variety — seeing what you expect to see — can become greatly strengthened in science and in philosophy, because when data are fitted into an intellectually satisfying explanatory system, data that don’t fit may be selectively set aside. Thomas Huxley wryly identified “the great tragedy of science — the slaying of a beautiful hypothesis with an ugly fact.” Given this universal human foible, the price of the pursuit of an ever-expanding, ever-receding truth is vigilance of a peculiar kind, and to this peculiar attitude of counter-advocacy some are more alert, or more suited by temperament, than others.

in linguistics there are parallel effects on funding, reputation, precedence, and intellectual property, albeit with less dire consequences, to be sure.

No one ever said that doing science was easy.

15. Hymes & Fought (1975: 172) give a striking example of Hockett’s own students misreading a paper of his in consequence of their being “in the grip of preconceptions in a climate of opinion.”

3. Founding a science

Zellig lived in a milieu of working scientists and mathematicians; his brother Tzvi and sister-in-law Susannah are immunologists; his wife Bruria Kaufmann (b. 1918) is a mathematician and theoretical physicist, and was Einstein's (1879–1955) assistant at Princeton, responsible for helping him to clarify and simplify his restatement of relativity and his work toward a unified field theory; his close friend and colleague was the Polish logician, Henry Hiż; and his interest in re-founding linguistics in linear algebra was nourished by conversations and correspondence with mathematicians including Marcel-Paul 'Marco' Schützenberger (1920–1996), the latter's student André Lentin, Max Zorn, and of course Bruria. As he pointed out in the 1986 Bampton lectures which became *Language and Information* (Harris 1988), there are two kinds of applied mathematics: calculational, of which there is very little in language, and the finding of mathematical objects in the world, of which there are many in language. His essential effort, developed consistently through almost sixty years, was "to see how a little mathematics might become linguistics," applying set theory and linear algebra to the elements and sets of linguistic analysis.¹⁶ According to Leigh Lisker,¹⁷ Zellig was teaching him and his fellow students transformations as early as the late 1930s. Subsequently, Zellig "had conversations about transformations with many people: with Piaget, and the psychologist David Rapaport, with Carnap and his follower Y. Bar-Hillel, with Max Zorn (of the lemma) to whom [he] showed the whole system at the Indiana Linguistic Institute, and with others" (Harris 2002.4).¹⁸

16. Lentin (2002: 1) says that Zellig approved the felicity of this phrase. *Mathematical Structures of Language* (Harris 1968) is the most complete expression of his effort to lay the basis for more capable mathematicians to identify homomorphisms, rings, and other structures, prove theorems, etc. One of his interests was in how language might develop additional capabilities; and as language is the (or an) interpretation of the mathematical structures that he did find, so might additional developments of the mathematics have their interpretation in extensions or specializations of language. See also (Harris 1962b) furthering an interest that he shared with Sapir.

17. In an e-mail message to the author, 1 March 2000.

18. Here, we must pause to clear up some potential terminological confusion. In abstract algebra, a homomorphism is termed more or less equivalently a linear transformation, linear map, linear operator, or linear function. (The related term 'kernel' is also from abstract algebra.) Zellig used this algebraic term 'transformation' quite literally and directly to refer to a mapping from subset to subset in the set of sentences. Noam's subsequent use of the same word to refer to deformations of abstract tree structures is only related by common subject matter, and derives not from algebra, but from Carnap's notion of 'rules of transformation' as a correlative of 'rules of formation'. The role of phrase-structure grammar in the sundering of Noam's work from Zellig's is an important thread that extends beyond the scope of this

As we have seen one critical ingredient of the scientific temperament is scepticism. Another is often called curiosity, meaning by that not the idle collection of trivia, but rather a peculiarly focused probing behind the surface appearance of things. A letter that Zellig wrote to Bernard Bloch (August 20, 1949) gives some personal insight into this aspect of his temperament:

I get my main interest or pleasure out of consuming — in my case it's specialized not into food but into subject-matter and information. It is quite important for me to find out what gives — whether with linguistics, or with human language, or with politics, or with physics.¹⁹

But while he was temperamentally a scientist, he thought of himself as a methodologist more than as a linguist.²⁰ Why a methodologist? To understand this, we must place ourselves in the milieu of linguistics in the years before, during, and after WWII. Linguists were acutely aware of creating a new science.²¹ Problems of recursion and regress then under discussion in logic and mathematics suggested that this

paper; likewise the visual impact of branching tree diagrams in the branding and packaging of Generative Grammar, and in the ready identification of in-group presentations vs. out-group presentations.

19. It was this characteristic need to “find out what gives,” which I share, that drew me particularly to Harris, although (for this and other reasons) I was already committed to linguistics when I came to Penn in 1966. The talk of an interest or pleasure being “specialized” may be in reference to the typology of Fromm.

20. The title of the first of his several major books famously begins with the word “methods.” He also supposed that linguists would not be interested in his work, though people interested in language would be, and that his work was not part of linguistics as it is institutionally defined. Such remarks possibly reflect in some measure the ways in which ‘linguist’ had been redefined by that time in the late 1960s, and the polemics then raging. However, while he knew the lasting value of what he was doing, Zellig neither demanded nor expected to retain a leadership role in the field on the strength of what he had done in the past. Anarchism means no institutionalized leadership. Many espouse anarchist and libertarian principles, Zellig embodied them in practice.

21. “What was at stake, in short, was the possibility of a science at all: the possibility of taking as problematic, exploring, and proving linguistic phenomena, rather than having inquiry cut off by conventional habits of explanation” (Hymes & Fought 1975: 162).

[T]he cast given methodological ideas in the United States was probably due to their institutional value, combined with the common tendency of a younger generation to delight in shocking an established order. Extreme differentiation of ‘scientific linguistics’ from ‘philology’ probably had adaptive value in efforts to secure a novel place in the academic sun. This essential *social* role of the Bloomfieldian idiom goes far to explain its general acceptance as central reference point, even by linguists whose own work continued or developed ideas in conflict with it. (Hymes & Fought 1975: 117)

new science might be subject to unique restrictions and requirements because of its peculiar subject matter, language. It was obvious to them, as quickly becomes apparent to any student in the sciences, that the language of scientific discourse — the very means of formulating and communicating hypotheses, methods, results, theories, and so on — is constrained or disciplined in ways that general usage is not. This is because in the sublanguage of the science, objects and operations are explicitly defined, either in common parlance or in terms of logically or methodologically prior sciences,²² with the consequence that sense and nonsense are sharply distinguished, as they cannot be in general usage. How might this affect discourse in a science whose subject matter is language itself? Discussions of metalanguage vs. object language, in the air since Bertrand Russell's (1872–1970) tremendously influential essay “On Denoting” (Russell, 1905) and Alfred Tarski's (1901–83) monograph on the truth-functional semantics of formalized languages (Tarski 1933), informed thinking about the relationship between linguistic descriptions and the languages that they described:

[T]he explicit structure of statements in logic and mathematics had made it clear that the statements about this structure could not be expressed within this structure: the metalanguage of mathematics was outside mathematics. (See for example [Church 1956]. While the term ‘metalanguage’ as used in the linguistic work is an extension of the use in [Carnap 1934], it also satisfies the more stringent (finitary) condition for the term ‘meta’ in [Kleene 1952].) The structure of the metalanguage had been left undescribed, the view being that it, or its metalanguage in turn in infinite regress, has to be undescribed and indeed not fully specifiable, simply given in natural language. This conforms to the common view in philosophy that natural language is amorphous, or in any case not fully specifiable. (Harris 2002.7–8)

There were other reasons that this new science had to find its own footing. Consider its relationship to neighboring fields. Physics bears obvious relevance to acoustic phonetics, but to nothing else in language; similarly biology is pertinent only through physiology and the study of articulation.²³ Psychology as then constituted had really nothing to contribute to linguistics regarding either that which was unique to individual languages or that which was common to all.

22. In a sublanguage, such use of a prior and external metalanguage is possible, as it is not for language as a whole. Use of analogy and metaphor extending the distribution of a word from one subject-matter domain to another is very much related, as illustrated, famously, by the adaptations of nautical terminology in common parlance. The main distinction to be made is that this is a folk process, whereas definitions in science are deliberate and disciplined.

23. Setting aside undisprovable speculations about genetics as anachronisms here.

By the time Bloomfield became a behaviorist in psychology, he was committed to the belief that linguistics was an autonomous science... Thus, one's preference for one psychological theory or another did not matter to the linguistic analysis done. (Murray 1994: 121)

As Bloomfield (1933: 32) said,

The findings of the linguist, who studies the speech-signal, will be all the more valuable for the psychologist if they are not distorted by any prepossessions about psychology.... We shall all the more surely avoid this fault ... if we survey a few of the more obvious phases of the psychology of language.

Consequently, claims that Bloomfield and other American linguists were behaviorists are essentially vacuous,²⁴ and any supposed influence of anthropology or any of the social sciences on the methods or theories of linguistics is so weak as to be scarcely ever heard of.²⁵ Rather, specialists in other fields looked to the scientific study of language for answers to some of their difficulties, and still do. Thus, Noam's various claims of providing insight from language into the nature of mind says more about the paucity of such insight in psychology than it does about the converse relevance of psychology, as presently constituted, to linguistics.

At the time we are discussing, the mid 1940s, American linguists had long been confronted with the complex and daunting task of describing exotic languages in understandable and useful ways, and this task grew enormously in scope in the "hothouse atmosphere of the wartime work" (Joos 1957: 108) as both demand and

24. Even if they were in some sense adherents of behaviorism (which is debated), so what? The methodology and practice of linguistics took no direction from operative conditioning or the like, resting instead on their own foundations and grappling with the unique requirements attendant upon using language to describe language. The determining factor was patterning in language. Attempts were made to justify this preoccupation with patterning by reference to 'habits' and the like, but it came to be seen (by Zellig perhaps most clearly) that no justification is required. Even Sapir, who was very attentive to the interrelation of language with culture and with personality, plainly recognized the autonomy of this patterning in language, and affirmed (so far as he concurred with Whorf) that what influence there was went the other way, from language to psychology. We know now, for example, about the social motivation of sound changes, thanks to Labov, but those motivations arise out of processes of social identification (conformity and differentiation) and can do no more than adjust the relative placement of sounds within patternings of contrast whose basis is informational rather than psychological.

25. *Pace* the efforts of Whorf to intrigue students and draw them into the field with the so-called Whorf-Sapir hypothesis (which continues to be remarkably effective in that function even today). The complaint goes rather the other way, e.g. from Dell Hymes and his students, that anthropology and ethnology have had too little influence on linguistics. See e.g. Hymes (1971).

funding increased for useful descriptions of languages deemed critical for the war effort. Bloomfield, Sapir, and their students, like Boas before them, had abundant direct experience of the propensity of experienced researchers, and not just new students, to project onto an unknown language expectations derived from their native tongue, overlaid with whatever they had been taught in school about alphabets, grammatical categories, paradigms, declensions, and the like. They well understood the need to take each language on its own terms, and Sapir sent his students questing after the unique ‘genius’ of each language. They also recognized genetic, areal, and typological commonalities across languages, which in some ways simplified, but in other respects complicated and confused their task. Sapir conveyed his methods inductively in sink-or-swim seminars, disclosing the patterning in great arrays of linguistic forms. Pike, Nida, and others developed training materials for student missionaries. With the wartime demand for competent linguists and scientific descriptions of languages there emerged with even more urgency a critical need for guidance how to proceed when confronted with an unknown language in the field. It was in response to this, for example, that Bloomfield wrote his 1942 *Outline guide for the practical study of foreign languages*.²⁶

In the pages of the *International Journal of American Linguistics* (founded by Boas in 1917) and *Language* (founded in 1925), published analyses of languages accumulated, exemplifying the more scientific methods. But to report the diversity of exotic language structures, linguists were devising almost equally diverse and exotic modes of describing them. So at the same time, in an almost parallel stream of articles and monographs, they searched for common theoretical and methodological ground. Implicit in the background was the question: which of the manifest differences between their presentations of language structure genuinely reflected the essential character or ‘genius’ of each language, and which amounted to what we might now call ‘notational variants’?

26. Sapir and Bloomfield obviously were strong influences in linguistics and on Zellig Harris. However, it is a peculiar distortion of the polemics of the 1960s to consider him a Bloomfieldian. Sapir thought very highly of Zellig’s work, beginning with a very positive review of his *Phoenician Grammar* (Harris 1936, the published form of his 1934 Ph.D dissertation), and children of Sapir have recalled that he considered Zellig to be his intellectual heir. The esteem was mutual. Zellig also greatly admired Bloomfield’s work, and him as a person. He studied formally with neither. The term ‘neo-Bloomfieldian’ is a rhetorical device equivalent to talk in the 1960s of the mythical ‘hegemony’ of ‘taxonomic linguistics’. Once this way of framing discussion is accepted, it becomes difficult to perceive the diversity that in fact obtained in the post-Bloomfield/Sapir era, and the spirit of collegiality in the presence of diversity — see the letter to Goetze quoted below for an instance — is misperceived as mere conformity.

To address this question as part of the necessary spadework for establishing the mathematical foundations of the field, Zellig wrote a series of brief ‘structural restatements’ of language descriptions that had been published by his peers (drawing also on other materials that were available to him, both published and in manuscript), mapping each to a kind of normal form. He was careful to present these not as a prescriptive norm, but rather as a formal basis by which they might fairly be compared, and by which essential linguistic differences might be distinguished from accidental differences of organization and presentation.²⁷ It should be realized that this was merely a published portion of a lifelong practice of reviewing and carefully analyzing linguistic descriptions written by others. Aside from familiarizing himself with the structures of diverse languages, the primary purpose was to test his methods (and, later, his emerging theory) at every stage. The related benefit was to avoid being hedged in by his own preliminary conclusions.

4. Carnap’s rules of transformation

Noam’s divergence from Zellig can be seen at an early stage in their respective interpretations of Carnap. Zellig contrasted Carnap’s enterprise with that of linguistics. Quoting Carnap (1928), he says:

It is widely recognized that forbidding complexities would attend any attempt to construct in one science a detailed description and investigation of all the regularities of a language. Cf. Rudolf Carnap, *Logical Syntax of Language* 8: “Direct analysis of (languages) must fail just as a physicist would be frustrated were he from the outset to attempt to relate his laws to natural things — trees, etc. (He) relates his laws to the simplest of constructed forms — thin straight levers, punctiform mass, etc.” Linguists meet this problem differently than do Carnap and his school. Whereas the logicians have avoided the analysis of existing languages, linguists study them; but, instead of taking parts of the actual speech occurrences as their elements, they set up very simple elements which are merely associated with features of speech occurrences. (Harris 1951a: 16n17, italics added)

From the context of this footnote (which we will consider presently), we know that these “very simple elements” are “set up,” not in any arbitrary way, but in the very process of identifying points of contrast between utterances, as perceived by native speakers. It is these differential elements, the contrasts, which are “associated with

27. “The justification ... is... the testing and exploring of statements of morphological structure.[...] The present restriction to distributional relations carries no implication of the irrelevance or inutility of other relations of the linguistic elements” etc. (Harris 1947a: 47).

[phonetic] features of speech occurrences” and thereby given a convenient and manipulable representation.²⁸

This passage (and others like it) may provide a good example of confirmation bias. Experience has taught me that many readers will understand here what they expect to read, instead of what is written. I have actually had linguists read this or other passages quoted below and say “he didn’t really mean that.” The usual expectation is that a relationship obtains between entities that exist as such prior to and independent of that relationship. A book is on a shelf. The relationship of one being on the other does not affect the book-ness of the book or the shelf-ness of the shelf. In the case of the primitive elements of language, however, it is the relationship of contrast which confers upon certain phonetic features (or feature bundles, to use Bloomfield’s term) the status of being “what are called phonemes” (Harris 1951a: 72n28).

For Noam, however, the fundamental elements are not the contrasts of the given language, they are phonetic descriptors that apply universally to any language. His understanding of the procedures of (Harris 1951a) was that they “were essentially procedures of segmentation and classification ... designed to isolate classes of phones, sequences of these classes” etc. (Chomsky 1975: 29). Subsequently, he took the fundamental elements to be predefined phonetic descriptors.

The alphabet of primitive symbols is determined by general linguistic theory, in particular, by universal phonetics, which specifies the minimal elements available for any human language and provides some conditions on their choice and combination. (Chomsky 1975: 5)

In the Introduction to (Chomsky 1951), Noam aligns himself with Carnap (1928):

Thus Carnap in the *Aufbau*,** for example, begins with a primitive relation between slices of experience and attempts to construct, by a series of definitions, the concepts of quality class, quality, sensation, etc., i.e., he tries to construct concepts for the most general description of experience. Similarly, it can be shown that the theoretical part of descriptive linguistics, beginning with three 2-place predicates of individuals, and restricting its individuals to a tiny domain of experience (i.e. speech sounds*) can construct concepts such as ‘phoneme’, ‘morpheme’, etc., which are available for a general description of that part of experience called linguistic phenomena. (Chomsky 1951: 1–2)

(The first footnote gives the citation for Carnap 1934; the second reads: “Or, perhaps, segments of magnetic tape on which speech is recorded.”) Thus, Noam is committed to the usual view that speech sounds, not speaker judgments of contrasts, are the primitive

28. Adapting Bishop Berkeley’s familiar example, if a recording of an utterance in some language is played in a forest, and no user of that language is there to hear it, what is the status of the sounds from the recorder?

elements from which phonemes and higher-level elements are ‘constructed.’ He simply did not grasp this radical difference in Zellig’s formulation of the foundations of linguistics, nor understand its importance for liberating linguistics from the strictures in which the philosophers felt themselves bound.²⁹

5. The Morphophonemics of Modern Hebrew

Let us now return to the mid 1940s to consider again the situation as Noam was entering the field as an undergraduate four or five years before writing this. Zellig had just completed the manuscript of his first book (Harris 1951a), and had just published (1945) or was then completing (1947a, 1947b) his brief series of structural restatements of descriptions by other linguists. This, then, was for young Noam the paradigm of linguistics — on the one hand using anecdotal examples and fragments of language data to demonstrate methodological and theoretical points, and on the other hand reframing someone else’s description of a language as a way of substantiating such points. Noam’s Master’s thesis, *Morphophonemics of Modern Hebrew* (Chomsky 1951) is a ‘structural restatement’ (without citation) of materials that Zellig gave him,³⁰

29. Lest it be thought that the pair test is an instance of a “2-place predicate of individuals,” note that Carnap’s notion of ‘primitive dyadic predicate’ is satisfied if two individuals ‘resemble’ each other, but in the judgment of a native speaker, a repetition is categorially identical, not merely resemblant.

30. Barsky (2007: 148) quotes a letter

that Harris wrote to Bernard Bloch ... on December 19, 1950...: “A student of mine, A. N. Chomsky has been doing a great deal of work in formulation of linguistic procedures and has also done considerable work with Goodman and Martin. Last year I [gave] him the morphological and morphophonemic material which I had here...”

Noam obviously also had access to Zellig’s description of the classical language (Harris 1941a), and his dialect study (Harris 1939), which would have been useful for the historical validation of morphophonemics. Zellig was a fluent speaker and writer of Modern Hebrew, a lifelong resident part of the year at kibbutz Mishmar Ha’emek in Israel. His continued work on the linguistics of Modern Hebrew during Noam’s studentship at Penn is testified by (Harris 1948, 1951b). The closest Noam comes to acknowledging his dependency on Zellig’s prior descriptive analyses is in Section 1.3, “The present paper will confine itself to step two in the description of Modern Hebrew.” Step one he identifies as the ‘discovery’ of elements and “the determination of the relevant sequences, classes, sequences and classes of classes, etc., of these elements.” Step two, then, is “the construction of a descriptive statement based on the results of this process of discovery” (sic). Of course to ‘discover’ and ‘determine’ all this is precisely to formulate a descriptive statement, since the very definitions (or determinations, if you will) of all of the relevant elements, sequences, classes, etc. are statements of their interrelations. This therefore amounts to an acknowledgement of relying upon a prior descriptive organization

and *LSLT* (Chomsky 1955a, 1956a, 1975), from which his (1955b) Ph.D dissertation was extracted, employs the same sort of organization and use of data as does Harris (1951a). And this of course has been the paradigm of Generative linguistics from its beginnings to the present.³¹

As to transformations, Noam only experienced them as means for regularizing the successive periods of a discourse. “As a student of Harris’s, I participated in seminars on discourse analysis from the outset until leaving for Harvard in 1951, along with Fred Lukoff, A.F. Brown, and a few others” (Chomsky 1975: 53n77). Noam’s Master’s thesis has nothing of syntactic transformations in it. He applies the terms ‘transform,’ ‘transformation,’ etc. to morphophonemic statements, in the sense of ‘rule of transformation’ borrowed evidently from Carnap.³²

of the data of the language (step one) and restating it (step two). So far as I am aware, Noam never returned to the grammar of Modern Hebrew in any subsequent writing, not even to cite examples. Much later, Noam proposed (Chomsky 1969: 33) that progress in linguistics in the 1950s depended on someone coming along who was familiar both with the mathematical work on recursive systems and with the tradition of historical linguistics. This describes Zellig far better than it does Noam, whose work demonstrates neither interest nor competence in historical linguistics, but who we may suppose was familiar with Zellig’s historical Semitic writings, and presumably also his own father’s. Murray (1994: 228n3) reports that Noam dismissed his proposal that he was referring to himself as this person, calling this a “malicious distortion” but evidently without clarifying who he in fact did mean.

31. The literature of generative grammar relies upon anecdotal examples taken in isolation, and in particular, in discussions of ‘taxonomic linguistics’ from the 1960s onward, Noam has repeatedly (one might say obdurately) selected distributional regularities that do not generalize as purported demonstrations that distributional methods are inadequate. His talk about having to choose between a large number of alternative grammars is supported only by such fragmentary examples. The choices are greatly reduced when full-coverage descriptions are considered, and dwindle even more under the requirement for a ‘least grammar’ articulated in Zellig’s last publications. Throughout his career, Zellig always aimed for and worked within a broad-coverage grammar. *Methods* (Harris 1951a), his most prominent work when Noam was a student at Penn, appears to have many exceptions to this generality, as it employs anecdotal examples to demonstrate particular points, but even there considerations such as “simplicity of statement” — the very considerations that have been most vexing to Noam — boil down to asking what works out best in the grammar as a whole. The basic principle of Optimality Theory captures the spirit of this.

32. In the published Introduction to *LSLT* (Chomsky 1975: 37) Noam says that (Chomsky 1951) is a nontransformational grammar. Its morphophonemic statements are ‘rules of formation’ and ‘rules of transformation’ in the sense of Carnap (1934). There is no evidence that Noam ever understood Zellig’s origination of the notion of grammatical transformation in algebra. Zellig’s transformations are a property of language, Noam’s are a formal device for representing that property by ‘enriching’ the rules of a phrase-structure grammar. Rules of grammar may be widely variant in form, as a matter of notation and system, but transformations in the algebraic sense are variable only insofar as language varies, and changes, and

Here it will be shown how sequences of morphemes (of [word classes] M and U, the basic words) are transformed by the morphophonemic statement into their constituent phonemes. Adjunction of this set of examples to the previous pair gives a complete exemplification of the transformation of all possible sentences into phonemic sequences. (Chomsky 1951a: 59)³³

And on p. 22:

The statements have the form of rules of transformation. Given a sequence of a certain shape, they direct you to alter the shape in a specified way. If the directions are followed, any sequence of morphemes, properly selected from M and U, will be transformed step by step into a sequence of phonemes.

In fact, syntax is presented only in a very simple manner, no more than a concatenation of construction classes into larger construction classes, first elementary sentences ES, and then sentences, which can be either ES or (recursively) **Sentence + Conn + Sentence**.³⁴ The concatenation possibilities within ES are represented by a table of the sort seen e.g. in Harris (1951a: 153,353), a notational variant of a simple phrase-structure grammar. The discussion of sentence forms and construction forms is perfunctory, sufficient merely to provide context for the morphophonemics by sketching (with great optimism!) what a complete grammar might look like. The main focus, as the title indicates, is on a system of 'rules of transformation' specifying the morphophonemic alternations in Modern Hebrew. Some of these statements are also in

possibly evolves (or is modified) to develop new capacities. Zellig developed a description of language as a mathematical object, and of linguistic information as its interpretation; Noam developed a formal system, the procedural steps of which produce (many, by intention all) sentences of a language, and advanced the hypothesis (couched as a necessary presupposition) that this system describes or corresponds to the cognitive means by which speakers of the language produce those sentences.

33. The page references are to the manuscript, or, more exactly, to a photocopy of the original ms. made for me by the staff of the Rare Book and Manuscript Library in the Van Pelt Library at The University of Pennsylvania, catalogued as 378.748PoA / 1951.60 (RBC). This passage is found in Section 5 "Derivations".

34. The form class **Conn** is defined by a list of morphemes. Over 30 years later (Chomsky 1975: 169), Noam described these and the construction rules that follow "phrase structure rules supplemented by extensive use of long components". The use of the string-rewriting ('tag machine' or 'Post canonical system') notation that Noam adapted (without credit) from Post (1943) is limited to the morphophonemic statements. It is known that Zellig made use of Rosenbloom's *Elements of Mathematical Logic* (Rosenbloom 1950), Chapter 4 of which contains a nice summary of Post's work. Noam could have read either or both, and there were developments of it elsewhere. It was Post who generalized the notion of an algorithm from its classical expression in arithmetic.

table form; some are definitions of morpheme classes and subclasses; and the rest are ordered morphophonemic rules, as in Bloomfield's (1939) *Menomini Morphophonemics*, cited in Harris (1951a: 236) as an instance of 'descriptive order' (ordered rules).³⁵ The demarcation between syntax and semantics is avowedly arbitrary: "It is convenient to consider the morphophonemic statement as being initiated at this point, although there is no systematic break."³⁶

More pertinent to our present considerations is the discussion of criteria for selecting the 'correct' grammar. They have nothing to do with the psychological capacities of the language user. As we have seen, Noam distinguishes (1.3) a first step, determining the elements and their combinations by distributional analysis, from a "second step ... the construction of a descriptive statement based on the results of this process of discovery [sic]."

[T]he statement of the grammar, the presentation of the results of the completed distributional analysis, must meet wholly different criteria which involve, essentially, considerations of elegance and considerations of adequacy as determined by the particular purposes of the grammar. (Chomsky 1951: 4)

For these criteria of elegance, he takes his authority from Goodman (1943: 107): "The motives for seeking economy in the basis of a system are much the same as the motives for constructing the system itself" (Chomsky 1951: 5n). He elaborates on this theme in Section 4:

The fundamental question about this preceding grammatical statement, aside from the question of its adequacy in describing the facts, is: in accordance with what general considerations was it constructed the way it was, and, in particular, to what extent is an order imposed upon the statements by these considerations? It will now be shown that the statements are, to a large degree, ordered by the criteria of 'elegance'. [...] (Chomsky 1951: 47)

The general considerations which have been regarded as relevant criteria are as follows:

1. Simplicity of statements
2. Maximization of the number of derivations in which a statement will occur relevantly.

35. Chapter 9 of Koerner (2002b) questions Noam's claim that he did not know about Bloomfield's treatment of morphophonemics with ordered rules until much later.

36. Last paragraph of Section 2. The spelling of 'morphophonemic' here and in the title of Section 3 is inconsistent with the spelling of 'morphoneme', etc. elsewhere, with the additional '-pho' later inserted supra. The term seems to have originated with the Prague School. Anderson (1985: 113, in a section of Chapter 4 entitled "Morpho(pho)nology") attributes the shorter term to Henryk Ułaszyn (1874–1956), a student of Jan Niecisław Baudouin de Courtenay (1845–1929), citing Trubetzkoy (1934: 30), then proceeds to use the short form thenceforth.

3. Minimization of irrelevant applications.
4. Maximization of similarity among statements, and amalgamation of statements involving the same elements.

[...] 2 and 3 are applied only when 1 or 4 are not thereby violated. 1 is outweighed by 4 when the difference in complexity of the formulations under consideration is simply a matter of the addition of one or two symbols, with the structure of the statement remaining unaltered. (Chomsky 1951: 49–50)

This metagrammatical question of how to determine which of various alternative grammars is the ‘correct’ one is the central purpose motivating *LSLT* (Chomsky 1955a). The proposed means of adjudication is “a general theory of language structure, a metagrammar” (Ryckman 1986: 131). For Noam, theory construction is a prerequisite to linguistic analysis; for Zellig, his methods neither require nor constitute “a theory of the structural analyses which result from” their application (Harris 1951a: 1), and a theory of language or of grammar is an outcome which must not be leaped at too quickly, lest it prejudice analysis and obscure the real properties of language behind ‘realities’ of presupposed theory. Noam makes the point that no observation or analysis is theory free, and this is true, but Zellig keeps his theoretical assumptions to a very minimum.

I have quoted Zellig to the effect that he crafted his structural restatements not to tell other linguists what they should do but to demonstrate the interconvertibility of their alternative types of description, making it possible to peel away differences that are merely artifacts of notation or descriptive style, thereby more clearly disclosing what is essentially distinct in each language.³⁷ It was not in his purpose, in his interest, or in his nature to demand that others conform to one way of doing things, nor, in his view, was it in the interest of the developing field of linguistics to presume to have the final answer and attempt to squelch alternatives, because “the best way to get closer to the ‘truth’ — after I have figured out whatever I could — is to get the divergent opinions which arise from a different scientific analysis.” (letter to Goetze quoted above).

Noam, however, frames the problem in terms of competing alternatives, and restatements in the Generativist literature (almost always ‘aspects’ of a grammar, that is, grammar fragments on a smaller scale than Harris 1947a, 1947b, are made only for the purpose of asserting superiority of one view over another, and deciding what properly belongs in the metagrammar that should guide further ‘discovery’. We see a glimmering of this even in (Chomsky 1951: 49n):

Actually, if these considerations [1–4 enumerated above] are dropped, almost any order can be shown to function successfully, simply by enumerating in each statement the cases which lead to an incorrect result, and adjoining to each statement a correction for each such instance.

37. An obvious extension of Sapir’s interest in characterizing the unique ‘genius’ of each language.

The prospect of eliminating all criteria produces the appearance of a great number of alternative grammars. Then, instead of criteria for choosing one partial analysis over another within a single continuously refined description, as in (Harris 1951a), Noam frames the process as adjudicating between grammars and reducing a thus inflated set of possible grammars until (ideally) only one is left standing. Logically, they may be equivalent; practically, they are not.

From this and other discussions of that period and later, it is evident that Noam felt that where there are two alternative statements, one must be incorrect and the other correct. This is of course consistent with the difference of temperament posited at the beginning of the present paper. We noted earlier how the aim of the procedures in Harris (1951a) is to verify that conclusions, however reached, have a valid relationship to the data of the language. Chomsky (1955a: I–9) offers the peculiar interpretation (citing it as *Methods of [sic] Structural Linguistics*) that these procedures “provide a practical mechanical way of validating [a grammar], i.e., of showing that it is in fact the best grammar of the language.” In this, Noam appears to be attributing to Zellig his own desire for an algorithm, albeit at this stage no longer an algorithm of discovery, but rather of adjudication among alternative descriptions. As John Goldsmith says, “Generative grammar is, more than it is anything else, a plea for the case that an insightful theory of language can be based on algorithmic explanation” (Goldsmith 2004: 1), or indeed that it must be so based in order to be ‘interesting,’ ‘non-trivial,’ etc. Any non-algorithmic description is dismissed as ‘vague.’³⁸

You will look in vain for the above quotations in the published version of *The Morphophonemics of Modern Hebrew* (Chomsky 1979a). In that revision, Carnap disappears from sight. The growth in maturity of thought and expression, the greatly increased differentiation from the locutions and principles of (Harris 1951a), and indeed the sheer extent of these revisions, in six short months of 1951, is quite remarkable. Noam has said (1975) that he intended to include it as an appendix to *LSLT*. It seems extremely unlikely that it was never revised for that purpose between 1951 and 1956. Bearing in mind the extent of unacknowledged revisions differentiating (Chomsky 1975) from the several 1955–1956 versions of *LSLT*,³⁹ it seems at least possible that the published version may include late revisions for (Chomsky 1975), or even later. How likely is it that Noam would send to publication any but the most recent revision? No wonder that it “addresses the subject right from the start in a surprisingly self-assured manner”, as Koerner (2002b: 24) says. For a detailed analysis of the

38. On the algorithmic nature of generative grammar, and the pitfalls of an algorithmic approach, see also Gross (1979:882–883).

39. Murray (1999), Koerner & Tajima (1986: 3–5, 56), Ryckman (1986, chap.3, esp. pp. 143–147n).

differences between the original ms. and the publication of 1979, see the contribution by Peter Daniels to this volume.

6. Discovery and restatement

Noam was steeped in Hebrew language studies in his family, and he may have done some linguistic fieldwork with Hebrew-speaking informants in Philadelphia, but the ‘discovery’ phase of linguistic analysis of Hebrew had already been done in Zellig’s publications when he undertook the restatement in (Chomsky 1951). On that borrowed basis, this is the only work approaching a comprehensive grammar that he has ever done. All his subsequent work has been a succession of restatements of fragments of the grammar of English. This is deemed sufficient because his primary aim is a metagrammar of all languages (‘Universal Grammar’), and he proposes that the grammar of any one language cannot be written properly until that is achieved.

Did Zellig write about discovery procedures? Based on extensive fieldwork experience, both his own and that of others, eliciting information from informants about unknown languages, Zellig did not seek and did not expect to find any discovery procedures to simplify and short-cut the work of linguistic analysis.

These procedures are not a plan for obtaining data or for field work. [...] They also do not constitute a necessary laboratory schedule in the sense that each procedure should be completed before the next is entered upon. In practice, linguists take unnumbered short cuts and intuitive or heuristic guesses, and keep many problems about a particular language before them at the same time.[...] The chief usefulness of the procedures... is therefore as a reminder in the course of the original research, and as a form for checking or presenting the results, where it may be desirable to make sure that all the information called for in these procedures has been validly obtained. (Harris 1951a: 1–2)

Nor did he develop formal language-like structures such as the ramifications of PSG that so engrossed Noam. His interest was in “differences in how the language data responded to identical methods of arrangement” (Harris 1951a: 3) and in how diverse methods disclose different properties of language. The structures disclosed by his methods could be formalized as language-like systems (types of formal grammars), but this work was largely left to others.⁴⁰ Many years later, he noted this algorithmic

40. Principally Aravind Joshi and his students, who formalized string analysis as adjunction grammars and as the major component of tree-adjointing grammars. Naomi Sager and her colleagues developed the previously mentioned systems at NYU by extending the string parser developed at Penn. Stephen B. Johnson is working on the formalization of Operator

aspect of Noam's work in his remark that "the tree representation could be considered a representation not so much of the sources of the sentence as of the ordered choices to be made in that system for producing the given sentence" (Harris 2002[1990]: 6). As we have seen, his effort from an early stage was to identify what is essential in language, avoiding artifacts due to any particular mode of analysis or presentation.

[T]he data, when arranged according to these procedures, will show different structures for different languages. Furthermore, various languages described in terms of these procedures can be the more readily compared for structural differences, since any differences between their descriptions will not be due to differences in method used by the linguists, but to differences in how the language data responded to identical methods of arrangement. [...] The central position of descriptive linguistics in respect to the other linguistic disciplines and to the relationships between linguistics and other sciences, makes it important to have clear methods of work in this field, methods which will not impose a fixed system upon various languages.... (Harris 1951a: 3)

By the end of 1946 he had summarized this work, as far as it had then gone, in the hefty manuscript that was published five years later in 1951 as *Methods in Structural Linguistics*.⁴¹ As we have seen, the recurrent theme of this book is that linguists of course

Grammar, using the Lexicon Grammar framework of Maurice Gross for the system of reductions. (Due to an error by the publisher, his name as co-editor of Nevin & Johnson 2002 is given as Stephen M. Johnson because that is another author's name previously entered in the publisher's database.)

41. There can be no doubt that it was revised in the interim. In an undated letter to Bernard Bloch (probably early August 1949), he wrote "this g-d d----d (for the hyphens, substitute od, and amne) methods in descriptive linguistics book was revised by me for the (n + 1)th time last spring. Several people who saw it said I ought to try a couple of commercial publishers before sending it to you (even though, as I explained to them, I had an informal understanding about it with you), in order to see if a wider audience (of non-linguists) could be reached." The title in 1946 and as late as 1949 was *Methods in Descriptive Linguistics*. Zellig told me he was not certain who made the change, he or the publisher, and I have supposed that the publisher substituted a word with more marketing sizzle. However, Joos (1957: 96) says:

An older term for the new trend in linguistics was 'structural'. It is not idle to consider how the term 'descriptive' now [1942] came to replace it, even if not all the reasons can be identified. The Sapir way of doing things could be called structural, but the term was more often used for the stimulating new ideas that were coming out of Europe, specifically from the Cercle Linguistique de Prague.

So, although it may have been a restoration of an older term, the publisher's motivation may have been to align with the sales appeal of European linguistics. The notion that 'neo-Bloomfieldians' rejected European ideas seems ill supported.

have freedom to work in diverse ways (Zellig was after all an anarchist), employing all kinds of heuristics and hunches as they wrestle their data into a form that makes useful sense, but as scientists aware of their all too human proclivity to perceive what they expect to perceive, and as linguists aware of the all too human suggestibility of informants, especially in politically and economically subordinated cultures, it was incumbent upon them to verify that their conclusions continue to bear a valid relationship to their data. It was Noam, and not Zellig, who by his own account was avid for discovery procedures.⁴²

7. Fundamental data

Judgments of meaning are among the heuristics that linguists use (Harris 1951a: 365n6). Indeed, Zellig plainly indicates that among all the possible distributional regularities that may be found, just those are sought which establish “elements which will correlate with meanings” (Harris 1951a: 188). A late statement is found in Zellig’s last book, *A theory of language and information*, in Section 2.4, entitled “Can meaning be utilized?”: “All this is not to deny the usefulness of considering meaning in formal investigation...” etc. (Harris 1991: 42). Not only is it incorrect, therefore, to say that he avoided meaning, as is commonly alleged; we must rather acknowledge that identifying the fundamental data of linguistics depends upon “the meaning-like distinction

42. This is not to say that Noam was interested in discovering data about languages. I see no evidence of that. Rather, he seems to have had the conviction that Zellig’s “presentation of the methods ... of research in descriptive, or, more exactly, structural, linguistics... in procedural form and order” (Harris 1951a: 1) should be, despite disclaimers on the page just cited and elsewhere, algorithmic procedures corresponding to the process by which children learn language. Noam’s lifelong penchant for abstract analysis in preference to breadth of data, increasingly prominent in his later career, seems to have been the dominant factor in this early work “devoted to the problem of revising and extending procedures of analysis so as to overcome difficulties that arose when they were strictly applied” (Chomsky 1975: 30). The procedure described in (Harris 1955, 1967) for identifying morpheme boundaries (or rather, the boundaries of morpheme alternants, or allomorphs) is the closest thing to a discovery algorithm in his publications; it depends upon a prior representation of the contrasts (standard orthography suffices for an approximation) and requires subsequent morphophonemic and morphemic analysis to establish the morphemes of a language. Note that (Harris 1955: n1) acknowledges Noam’s input. The Turing-machine-like system for sentence recognition (Harris 1966) was presented as a demonstration of distributional principles, not as a discovery procedure. “The detailed problem here is ... the word-representation methods which made it possible to apply so simple a device. Indeed ... it can be studied as a notation for the modalities of requirement and permission [which a given word has to all envioning words].”

between utterances which are not repetitions of each other” (1951a: 363).⁴³ We will return to the nature and role of meanings again farther on.

This brings us to a subtle point, central to Zellig’s work, namely that the fundamental data of linguistics are not phonetic records or notations, but rather speaker judgments of what is different and what is a repetition. Phenomena cannot be repeated exactly; categorical perceptions of those phenomena can. In consequence, these most basic linguistic elements, the phonemic contrasts, are not segments or features of sound as discriminated by phoneticians, they are perceptual distinctions made by native speakers, which we keep track of by associating them with such phonetic segments or features.⁴⁴ Sounds are associated with the contrasts by the very process of the substitution tests that identify them, a process which requires segmenting utterances such that one segment can be substituted for another.⁴⁵ But the contrasts are primary, and can be represented by segmenting the speech stream in alternative ways. This crucial point is stated in numerous places, for example in the extended statement at (1951a: 16–21), and in the following:⁴⁶

Since the representation of an utterance or its parts is based on a comparison of utterances, it is really a representation of distinctions. It is this representation of differences which gives us discrete combinatorial elements (each representing a minimal difference). (1951a: 367)

Yet, despite the frequent reiteration with different words in a variety of contexts, so far as I know, none of his contemporaries recognized this.⁴⁷ Bernard Bloch, for example,

43. Likewise in many other places, e.g. (Harris 1942), where is cited also Bloomfield (1933: 161) to the same effect. See also (Hiž 1979: 344), quoted farther on below.

44. This is consistent with the development of logic and mathematics from the making of a distinction in a space, see Spencer-Brown (1969).

45. Ideally, but not necessarily, the pair test. The segments need not all be made by ‘vertical’ cuts of the phonetic record; the substitution of ‘horizontal’ features or simultaneous components is also possible.

46. See also (Harris 1941b), especially the third paragraph, beginning “It is in the second step, selection of the contrast-criterion....”

47. See e.g. Joos (1957: 108) “[Hockett’s] segmentation into ‘sounds’ is not so much logically justified as taken for granted. Today it is even considered possible to defend the thesis that such segmentation can’t be done strictly until after phonemic segmentation has been somehow (say distributionally) established.” Had he “today” (1956 or so) understood the pair test and its consequences as to the ontology of phonological contrast, he would have understood that an initial segmentation is inseparable from, and in a sense is effected by, the determination of contrasts. In particular, he could not have employed that last phrase, “somehow (say distributionally) established,” for the non-substitutability of contrasting segments in the pair test (or more generally in substitution tests) is precisely a distributional establishment of

assumed that phonemic distinctions can only be derived from distributional analysis of phonetic data (Bloch 1948).⁴⁸ Certainly it is no part of phonology today, and, although in several places Noam recognizes the utility of the pair test (e.g. 1975: 93, 147) for determining the phonemes without appeal to meanings, there is absolutely no evidence that he understood the consequences for the ontological status of phonemic distinctions and of linguistic elements established on that basis. Everything in his phonological work is grounded in a universal alphabet of descriptors or features, each of which is defined by some phonetic characteristic.⁴⁹

Those descriptors are considered to be universal phonological elements. The effort to achieve universality for them has turned out to be difficult through many successive revisions, and remains inconclusive. The features are considered to be phonological rather than merely phonetic because they cohere in a system; that is, their interrelations constitute an *a priori* abstract system which is taken to be more important phonologically than the equally or even more systematic considerations of their distribution in utterances.

them. See also the comments earlier on confirmation bias. Harris's writing is always careful and exceptionally clear, and (unless the matter itself is exceptionally complex) difficulties are usually where what he says is contrary to the reader's expectation. As Jane Robinson once observed (pc): "If I have an idea what he's talking about, I can understand him. As someone said of Quine, once you've understood what he means, you realize he couldn't have said it any other way. Harris is that way for me. It's just that what he's trying to say is difficult."

48. The critique of Bloch and others and the rejection of distributional methods in (Chomsky 1964) appears to be intended to apply, incorrectly, to Harris (1951); see Nevin 1993a, 1999 for discussion. Some readers have been confused by e.g. "This procedure takes the segmental elements of Chapter 5 ... and groups them into phonemes on the basis of complementary distribution" (Harris 1951: 59). What is overlooked, leading to this misunderstanding, is that "the segmental elements of Chapter 5" are not 'phones', that is, elements of a phonetic transcription; rather, they already constitute a preliminary phonemic representation, since they indicate speakers' phonemic distinctions, albeit inefficiently, and it is those inefficiencies or redundancies that distributional analysis removes.

49. Some work on a hierarchy of features that enter progressively, as it were, into the phonemic contrasts of a language (e.g. Dresher 2003) amounts to a statement of affordances for contrast (if I may borrow Gibson's (1977, 1979) terminology without importing with it his conclusions about the environmental causation of behavior). This principle is most clearly seen where in certain articulatory regions differences of articulation have little acoustic effect, and in adjacent regions slight differences of articulation result in a change in categorical perception. An example is the so-called quantal vowels (Stevens 1972). It remains that the contrasts of a language cannot be predicted from phonetic data, and the crucial factor is still human perception. See also Kirchner (1995).

8. The attack on distributional analysis

At the end of Chapter X of (Chomsky 1956a), we read the following approbation of distributionalism:

Within the general limits of formal distributional analysis, there are many avenues of investigation beyond those that have been followed here, notably, the whole study of the statistical structure of language. ... In short, there seems to be considerable unexplored territory within the boundaries of formal distributional analysis, and the possibilities and potentialities of such analysis show no sign of having been exhausted. It is surely premature to insist that the basis of linguistic theory be extended to include obscure and intuition-bound concepts, on the grounds that the clear notions of formal analysis are too weak and restricted to lead to interesting and illuminating results. (Chomsky 1956a: X751–752)

This Chapter X became⁵⁰ the Summary Chapter 1 of (Chomsky 1975), but when we consult the latter, such language has entirely disappeared. What happened?

In his plenary presentation at the 1962 Ninth International Congress of Linguists, as later expanded to (Chomsky 1964), several aspects of the divergence that we are discussing are in sharp relief. A full analysis of this *tour de force* of reframing⁵¹ would require treatment of at least equal length, but a few exemplary extracts will help illustrate the extent and nature of the divergence in views at this stage, beginning here with the question of phonetics and phonological contrast. Most of what Noam says

50. “In the spring of 1956, I began to revise the manuscript for publication. In the original, the tenth and final chapter was a summary. In the version that I was preparing for publication, I placed the summary chapter first, otherwise leaving the chapter order unaltered. During that year I did manage to rewrite the new summary chapter (Chapter I) and the first five chapters of the original (Chapters II–VI). The manuscript as published here consists of a preface written in 1956 and the chapters (here numbered I–VI) of this edited version.” (Chomsky 1975: 3) However, Chomsky (1956a) shows the summary chapter (from which I have quoted here) in its original position at the end as Chapter X. Ryckman (1986) has shown that extensive revisions in addition to the rewriting in 1956 intervened before publication of (Chomsky 1975).

51. I use the term ‘reframing’ in the sense of Lakoff (1987, 2002). (It is likely that he adopted the term from Bandler 1982. See the Wikipedia article on reframing for an account of the origination and earlier usage of the term.) Reframing shifts the terms of debate in presuppositional ways. In this instance, the diversity of descriptive linguistics was reframed as a unified, hegemonic ‘taxonomic linguistics’ straw man. This collapsing of diversity into a convenient form for argument is admitted e.g. at (Chomsky 1964:75): “Though modern phonologists have not achieved anything like unanimity, a body of doctrine has emerged to all or part of which a great many linguists would subscribe” etc. Developments in phonology that are too diverse to be made to fit are dismissed as inexplicit and vague (*ibid.* fn. 13), as measured against a standard of suitability for formalization in an algorithm.

is perfectly compatible with Zellig's views. He has told me (letter of September 18, 1995) that this apparent compatibility is only because Zellig is 'vague', but it is more accurate to say that Zellig did not demand that only one of several alternative formulations must be true. For Zellig, these are not rules in an algorithm, they are tools of analysis. It is true, as Noam says, that he held that these tools are useful for different purposes, yielding results which in the very nature of science are provisional. But more than that, each type of analysis (constituency, expansions, string analysis, transformations, least differences, discourse analysis, sublanguage analysis, etc.) reveals a different aspect of "what gives" in language. Where Noam looks for rules, Zellig offers methods. For example

It should be clear that while the method of 7.3 [grouping segments having complementary distribution] is essential to *what are called phonemes*, the criteria of 7.4 [phonetic and environmental symmetry] are not essential 'rules' for phonemicization, nor do they determine what a phoneme is. (Harris 1951a: 72n28, italics added)

Zellig developed diverse methods of analysis, and showed how each method discloses some of the properties of language well, and others not so well. His approach was to honor the method, apply it with neutral care, and see what emerges.

Zellig Harris's work in linguistics placed great emphasis on methods of analysis. His theoretical results were the product of prodigious amounts of work on the data of language, in which the economy of description was a major criterion. He kept the introduction of constructs to the minimum necessary to bring together the elements of description into a system. His own role, he said, was simply to be the agent in bringing data in relation to data.... [I]t was not false modesty that made Harris downplay his particular role in bringing about results, so much as a fundamental belief in the objectivity of the methods employed. Language could only be described in terms of the placings of words next to words. There was nothing else, no external metalanguage. The question was how these placings worked themselves into a vehicle for carrying the 'semantic burden' of language.... His commitment to methods was such that it would be fair to say that the methods were the leader and he the follower. His genius was to see at various crucial points where the methods were leading and to do the analytic work that was necessary to bring them to a new result. (Sager & Ngô 2002: 79)

Noam understands matters in terms of argument, of which there must be a winner. These are for him not tools of analysis, but theories of grammar which are either correct or incorrect. For example, central to the project of Generative Phonology, as we have seen, is the stipulation that the phonological features constitute a system which is independent of any particular language in which those features might be employed to make the contrasts between different utterances. Noam (1964: 77) sets this against the position that "languages could differ from each other without limit and in unpredictable

ways” (attributed to Boas by Joos 1957: 96).⁵² In his (1941b) review of Trubetzkoy, Zellig does not reject what Noam (op. cit.) calls ‘systematic phonetics’ (also ‘structural phonetics’); although he says “it is pointless to mix phonetic and distributional contrasts” [sic], and he asserts that the former are secondary to the latter:

[I]n order to study the relations between phonemic contrasts one must first have selected what kind of contrast to investigate. Those which Trubetzkoy studies are the phonetic contrasts. He does not say that he is intentionally selecting these rather than any other. He merely uses them as though they were the natural and necessary ones to consider. [...] But *there are other criteria in terms of which one may study the contrasts....* Chief among these is the positional distribution.... Trubetzkoy was quite aware of this....he discusses the importance of considering [this]... and... modifies the patterning of the phonetic contrasts by some results from distributional contrasts. However, it is pointless to mix phonetic and distributional contrasts... [and phonetic parallels among the phonemic contrasts] must be independently proved. (Harris 1941b.348, italics added)

This passage invites the quoting of sound bites (or ‘text bites’) out of context. Compare the more careful formulation of his (1951a) in which speaker perceptions of contrast are the primary data which are distributionally associated with discrete features of phonetic data by substitution tests, and then these associations are wrestled by various criteria into one of the several possible more efficient and useful representations “that are called phonemes”.⁵³ The difficulty is that he is using the term ‘contrast’ ambivalently

52. Noam quotes only this phrase, and attributes it solely to Joos, who in fact attributed it to Boas and the tradition following him: “Trubetzkoy phonology tried to explain everything from articulatory acoustics and a minimum set of phonological laws taken as essentially valid for all languages alike, flatly contradicting the American (Boas) tradition that languages could differ from each other without limit and in unpredictable ways, and offering too much of a phonological *explanation* where a sober *taxonomy* would serve as well. Children want explanations, and there is a child in each of us; descriptivism makes a virtue of not pampering that child.” One cannot but wonder if Noam took this last as being aimed at himself, and rankled at it. But see also Thomas (2002) on the now institutionalized misinterpretation of this quotation from Joos.

53. “At a time when phonemic operations were less frequently and less explicitly carried out, there was discussion as to what had to be done in order to arrive at ‘the phonemes’ and how one could discover ‘the phonemes’ of a language. Today we can say that any grouping of complementary segments may be called phonemic. [Bear in mind that the ‘segments’ here are not ‘phones’, having been established in substitution tests that identified the phonemic contrasts. — BN] As phonemic problems in various languages came to be worked out, and possibilities of alternative analysis were revealed, it became clear that the ultimate elements of the phonology of a language, upon which all linguists analyzing that language could be expected to agree, were the distinct (contrasting) segments (positional variants, or allophones) rather than the phonemes. The phonemes resulted from a classification of complementary segmental elements, and this could be carried out in various ways. For a particular language, one

in this passage. The phrase which I have italicized above, “other criteria in terms of which one may study the contrasts,” distinguishes the “contrasts” from two kinds of criteria for studying them, (a) “other [i.e. distributional] criteria” on the one hand and (b) the previously discussed phonetic criteria that produce what he equivocally called Trubetzkoy’s “phonetic contrasts,” on the other.

The thrust of this part of the review is that “phonemes are in the first instance determined on the basis of distribution” and phonetic criteria are secondary. This is because the substitution tests (idealized as the pair test) establish the “first instance” by identifying which parts of the speech stream may be substituted for each other without contrast, and which may not, and it is the substitution within an environment which makes the very establishing of them distributional in character.

Noam (1964:77) continues quoting Joos (1957: 228) as saying

that “distinctive features are established ad hoc for each language or even dialect”, and that “no universal theory of segments can be called upon to settle the moot points” (228). Similarly, Hjelmslev appears to deny the relevance of phonetic substance to phonological representation.

What Joos is discussing here, with reference to Hockett (1947), is the variety of phonetically defined distinctive feature systems that had emerged within ‘taxonomic linguistics’. Consequently, it is neither a “rejection of the level of structural phonetics” *qua* level nor of “the relevance of phonetic substance,” as Noam says. Rather, it is a rejection of applying any *a priori* solution universally. It is unclear whether or not Joos understood Zellig’s reason for such rejection. This reason, as we have seen, was that the terms of contrast must be determined for each language on the basis of distribution (substitution tests identifying speaker perceptions of contrast and thereby associating phonetic features with them), and that phonetic criteria are therefore necessarily secondary.

The difficulty is that the universality of any system of distinctive features had (and has) yet to be demonstrated, and to presume it in advance, begging the question, would actually hinder that demonstration. As an illustration, at a certain stage in the long and complicated (and still unfinished) history of this universal alphabet, so-called ‘pharyngeal’ or ‘faucal’ sounds in languages like Arabic and Achumawi have been represented by a retracted tongue root (RTR) feature. The actual articulatory gesture is

phonemic arrangement may be more convenient, in terms of particular criteria, than other arrangements. The linguistic requisite is not that a particular arrangement be presented, but that the criteria which determine the arrangement be explicit.” (Harris 1951: 72n28) The fact that the allophones are contrasting affirms that the contrasts are yet more basic, although of course they (the contrasts) are established simultaneously with the allophones by the same substitution tests.

epiglottideal (Laufer & Conday, 1979), with secondary effects on the tongue, pharynx, and larynx. In Achumawi, certain phenomena are simpler to describe if the epiglottideal aspirate (like that of Arabic) is treated in parallel with the glottalized consonants of the language, to which no one would apply this RTR feature. Yet convenience for the ‘universal’ feature hierarchy dictated the existence of RTR and that it should be used to distinguish this aspirate from the ordinary one, complicating the description of those laryngeal phenomena (Nevin 1998).

Noam continues (1964: 77):

Nevertheless, it seems to me correct to regard modern taxonomic phonemics, of all varieties, as resting squarely on assumptions concerning a universal phonetic theory of the sort described above. Analysis of actual practice shows no exceptions to the reliance on phonetic universals. No procedure has been offered to show why, for example, initial [p^h] should be identified with final [p] rather than final [t] in English, that does not rely essentially on the assumption that the familiar phonetic properties ... are the “natural” ones. Harris might be interpreted as suggesting that a non-phonetic principle can replace reliance on [sic] absolute phonetic properties when he concludes (1951a, 66) that “simplicity of statement, as well as phonetic similarity, decide in favor of the p–p^h grouping”; but this implication, if intended, is surely false. The correct analysis is simpler only if we utilize the familiar phonetic properties for phonetic specification. With freedom of choice of features, any arbitrary grouping may be made simpler.

There are several equivocations here. First, what is at stake is not a universal theory of phonetic description, but a universal theory of phonetically defined phonological features — in other words, of all possible phonetic descriptors, just that subset which is presumed to account for all contrasts in all languages, in such a way as to support the most perspicuous account of alternations in all languages.

Now, place what Zellig said about “simplicity of statement” in context:

It is also convenient to have the relation among segment definitions within one phoneme identical with the relation in other phonemes. This requires that the segments be grouped into phonemes in such a way that several phonemes have correspondingly differing allophones ... in corresponding environments.[...] We could have grouped [p] and [t^h] together, since they are complementary. But the above criterion directs us (barring other relevant relations) to group [p] with [p^h]... For if we do so, we can say that the /#__V/ member of all these phonemes is virtually identical with the /s___/ member except that [h] is added; such a simple general statement would not have been possible if we had grouped the segments differently. (Harris 1951a, 66)

It is in a footnote that Zellig says we could in principle violate this, combining e.g. [p] with [t^h], etc., but that doing so would complicate the description. The definition of phonemes would have to say that in addition to adding [h] the point of articulation

changes. This would also complicate any statement (rule) applying to phonemes defined in this way. Therefore, “simplicity of statement, as well as phonetic similarity, decide in favor of the [p] – [p^h] grouping.”

What Zellig says here cannot coherently be construed as “replacing” a reliance on phonetic features, since the statements or rules in question cannot be compared as to their simplicity except by reference to phonetic features that [p], [p^h], [t^h], etc. do or do not have in common. This is of course what Noam paraphrases in the quoted passage at (1964: 77), but it is so obvious that we have to wonder why he bothered to flog the point.

In the larger context of the section quoted above, what Zellig is saying is that at this (hypothetical!) early point in phonemic analysis many items in the preliminary segmentation are in complementary distribution, so complementarity is not always a sufficient criterion for associating two segments with the same speaker-identified contrast. By a second criterion, we may group segments so that “several phonemes have correspondingly differing allophones” — in other words, so that phonological rules can refer to such phonemes together as a set, or alternatively can refer to the features that they have in common. This is what he calls “simplicity of statement.” It is difficult to imagine what Noam could fault about making a phonological rule more general.

To suggest that the criterion of “simplicity of statement” somehow *replaces* “reliance on absolute phonetic properties” is so curiously obtuse that we must wonder at its possible motivation. In this context, consider how peculiar it is for Noam to wonder what Zellig meant here. After all, he could just ask him, if he had not already done so in the preceding 15 years or so of their association. The explanation evidently is that it was rhetorically useful for his argument.

What Noam is gearing up for here is a rejection of distribution as a criterion. This is necessary in order to supplant the distributional identification of contrasts (by substitution tests) with a universal alphabet of phonetically defined terms of contrast. He does not come right out and assert his own exclusive reliance on phonetic descriptors which are claimed to be ‘absolute’ universal phonological features. Perhaps the reason he does not say this is because it is convenient for his argument at this point to assert as a straw man that Zellig, by using other criteria in addition to shared phonetic properties, is doing the converse of what Noam is in fact doing — that is, rejecting the criterion of shared phonetic properties — a classic inoculation move in polemic argument. This need not have been nefarious abuse of rhetorical prowess. It has been suggested (e.g. R. Harris 1998) that it is difficult for Noam to perceive what someone is saying in any terms other than his own. This is of course a cognitive trap to which we are all subject, but is perhaps exacerbated by the characteristic cognitive style that I have attributed to him.

The Noam of (1951a, 1955a) was entirely in accord with Zellig on this point, placing “simplicity of statement” first in rank among criteria for choosing the ‘correct’

description (1951a: 49–50); he of (1964) rejects it, giving paramount importance instead to an a priori system of phonetic descriptors (their systematicity apparently that which makes them ‘phonological’), with a parallel shift respecting syntax. This shift, and the ratcheting up of polemical rhetoric, coincides with the emergence of ‘universal grammar’ in Noam’s doctrine.

9. Sounds and contrasts

Noam has acknowledged (letter of 9/18/1995 to the author) that his (1964) critique of linearity, invariance, biuniqueness, and local determinacy does not apply to Zellig’s work. In this, he may have backed off from the claim (Chomsky 1964: 98) that “The only general condition that they [Zellig’s procedures] must meet is the biuniqueness condition, which is not justified on any external count, but simply is taken as defining the subject.”⁵⁴ But here again they are not referring to the same thing. Noam is talking about the correspondence of ‘taxonomic phonemes’ to the ‘phones’ of a phonetic transcription. Zellig is talking about the correspondence between two phonemic representations of the contrasts attested by speaker judgments: a preliminary and relatively disorganized representation achieved by associating phonetic descriptors with the contrasts, and a more efficient representation achieved by distributional analysis yielding “the scientific arrangement” (Harris 1941b: 345) of those phonetic data. Both “arrangements of data” depend upon the segmentation of the utterance by those acts of substitution that identified the contrasts (that is, the speaker judgments of what is repetition and what is not) and which thereby associated phonetic data with them. The distributional analysis does not create or discover the phonemic contrasts; it merely maps a complex and redundant representation of the contrasts to “what are called phonemes,” a clearer and more useful association of phonetic data to the contrasts.

Zellig usually referred to this mapping as a ‘one-one correspondence,’ but introduced the term ‘bi-unique relation’ in his 1944 paper on phonemic long components:

54. Zellig admits a one-many correspondence as follows: “In general, the representation is in one-one correspondence with each occurrence of the represented speech. In the case of intermittently present distinctions, however, it is in one-one correspondence only with a set of repetitions of the represented speeches” (Harris 1951: 364n5). The “external [ac]count” justifying this correspondence depends upon the purpose to which the representation is put. Typically it is important to have a well-defined mapping from speech to representation so that the recording of speech for the sake of further analysis is straightforward; and since utterances predicted by the analysis must be tested with native speakers, it is obviously important to be able to reconstruct from the representation something that they recognize as utterances in their language.

Finally, if we are ready to admit partial overlapping among phonemes, we may agree to have different components in different environments represent the same phonetic value. So long as we do not have a component in one environment represent two phonetic values which are not freely interchangeable, or two components or component-combinations in the same environment represent the same phonetic value, we are preserving the bi-unique one-to-one correspondence of phonemic writing. (The term bi-unique implies that the one-to-one correspondence is valid whether we start from the sounds or from the symbols: for each sound one symbol, for each symbol one sound.) (1944a. 187–188, footnote suppressed)

This sounds like a correspondence of phones with phonemes. However, not any ‘sound’ will do, but only those that are associated with the contrasts by virtue of the segmentation established in substitution tests with informants to determine speaker judgments as to “phonemic distinctions, these being the ultimate and necessarily discrete primitives of the language structure. Phonemes were defined as a convenient arrangement of how these phonemic distinctions *appeared* in utterances” (Harris 2002.2, italics added).

Those speaker judgments necessarily pertain to the given language, and are not themselves universal. There may indeed be some set of universal ‘affordances’ (Gibson 1977) that people preferentially, perhaps even necessarily, use to make the contrasts of their respective languages. There is evidence of this in so-called quantal vowels and in phonological universals. From Zellig’s point of view, that is an interesting and valuable result, but it cannot be presumed in advance. In any case, speaker judgments are the primary data.

Zellig concurs in the value of simultaneous features (e.g. 1951a: 64–65) as a representation of contrasts that may in some respects be more convenient than phonemic segments or phonemic long components, or combinations of these. Noam argues that phonological features are the *only* legitimate representation — when alternatives are offered, there can only be one winner. The argument has two aspects, firstly their convenience for stating phonological rules, and secondly their purported universality. Claims of their universality rest overtly on arguments for innateness. In this, Noam’s earlier wish for an algorithmic solution to the problem of language description has been reframed as a biological solution to the problem of child language.

The case for exclusive use of features is the core argument against ‘taxonomic phonemics’ in (Chomsky 1964). Split and merger processes in the historical development of languages can produce an ‘asymmetry’ such as that reported for Russian in (Halle 1959), where voicing is contrastive for stops but not for the affricates /c, č/ and fricative /x/, yet in all cases voicing is assimilated from a following obstruent. In a ‘taxonomic’ segmental representation, this must be described in the same way for all cases, but with two kinds of rules, morphophonemic rules for stops and allophonic rules for the affricates and fricative. “Simplicity of statement” requires a single rule

stated in terms of features. Zellig of course had advocated “simplicity of statement” long before, but had greater variety of means to that end at his disposal. Noam does not consider alternative formulations that are possible using phonemic long components instead of universal features.⁵⁵

10. The relative nature of elements

As we have seen, the substitution tests which establish native speaker judgments of contrast are inherently distributional in character, and establish phonemes *relative* to one another. This is why “phonemes are in the first instance determined on the basis of distribution” (Harris 1941b: 348). Given the relative nature of the contrasts, all the phonemic, morphophonemic, morphemic, constructional, and other elements are necessarily also relative in nature. For most practical purposes, these elements are treated as though they are records of things which happen to be different, but in fact each element is no more than a marker of differences.

Since each element is identified relatively to the other elements at its level, and in terms of particular elements at a lower level, our elements are merely symbols of particular conjunctions of relations: particular privileges of occurrence and particular relations to all other elements. It is therefore possible to consider the symbols as representing not the particular observable elements which occupy an environment but rather the environment itself, and its relation to other environments occupied by the element which occupies it. We may therefore speak of inter-environment relations, or of occupypings of positions, as being our fundamental elements. (Harris 1951a: 370–371)

Thus, when Zellig says things like the following, it is not ‘hocus-pocus’ or ‘game playing,’ he is talking about something that is essential in the nature of language:

...morphemes may be regarded either as expressions of the limitations of distribution of phonemes, or (what ultimately amounts to the same thing) as elements selected in such a way that when utterances are described in terms of them, many utterances are seen to have similar structures. (1951a: 367)⁵⁶

55. In 1964, Zellig had not concerned himself with phonology for probably 20 years. Any representation of the contrasts, even conventional English orthography, is adequate for his investigations into the informational properties of language. Phonology is encapsulated relative to syntax and semantics in Generativist writings as well, for the same reasons.

56. The locution “as expressions of the limitations of distribution of phonemes” in this passage shows that Zellig was aware of the next-successor stochastic determination of morpheme boundaries in the 1940s, long before writing (Harris 1955, 1967). This is of course

Thus, while intuitions of the meanings of morphemes (or lexical items) are suggestive, regularity and generality of structure in the description *as a whole* are criterial for determining by which distributional regularities to define the elements (phonemes, morphemes, etc.). Noam, however, sees here only a vague statement by a man who “did not believe there was any truth to the matter to be discovered” (letter to the author, 9/8/1995). It is not that Zellig did not believe that there was any truth to the matter to be discovered, but rather he did not believe that any of the diverse possible representations should be taken to *be* that which they represent. We have already noted that the findings of science may legitimately make no claims (or covert presuppositions) of absolute truth. See for example Hawking, who affirms (1988: 9) that a scientific theory “exists only in our minds and does not have any other reality (whatever that might mean).”

Zellig’s recognition of the relative character of all linguistic elements, and his facility in shifting perspective between an element and its environment, conferred great freedom on his investigations of language data and his formulation of methods and, later, of theory. For Noam, the purpose of analyzing language data is to specify universal elements and metagrammatical principles when then can be known in advance of analyzing language data.

11. The role of meaning

In the 1975 Preface of *LSLT*,⁵⁷ Noam clearly declared the autonomy of syntax from semantics:

Syntax is the study of linguistic form. Its ... primary concern is to determine the grammatical sentences of any given language and to bring to light their underlying formal structure. [...] Semantics, on the other hand, is [...] the study of how this instrument, whose formal structure and potentialities of expression are the subject of syntactic investigation, is actually put to use in a speech community. [...] We

related to information theory and to linguistic information, that aspect of meaning which is specific to language. Note also that the decade or so of incubation between the writing of this brief allusion in the 1940s and the publication of a scientific report in 1955 is typical of his way of working.

57. Although there is no Preface to the 1956 revision of the June 1955 ms. as posted at <http://alpha-leonis.lids.mit.edu/chomsky/>, and this passage does not appear as such in that ms., we do read on page 2 that “no reliance is placed on the meaning of linguistic expression in this study, in part, because it is felt that the theory of meaning fails to meet certain minimum requirements of objectivity and operational verifiability, but more importantly, because semantic notions, if taken seriously, appear to be quite irrelevant to the problems being investigated here.”

shall study [syntax] as an independent aspect of linguistic theory. [...] [S]emantic notions ... appear to assist in no way in the solution of the problems that we will be investigating. We will see, however, that syntactic study has considerable import for semantics. This is not surprising. Any reasonable study of the way language is actually put to work will have to be based on a clear understanding of the nature of the syntactic devices which are available for the organization and expression of content. (Chomsky 1975: 57)

Although the conflation of semantics and pragmatics that is evident here has been somewhat clarified since this was written, Noam's view of semantics as the interpretation of the productions of syntax has persisted in generativist proposals of semantic features, a semantic component parallel to a phonological component, and so forth.

We may contrast

...the attitude of Harris toward semantics. In his work, he uses semantics both explicitly and in a syntactic guise. Contrary to general belief, in his early work (1950), Harris does not eliminate explicit semantics. Rather, semantics is reduced to a single, simple, and testable question: Are two utterances repetitions of each other, or do they contrast? *Book* and *hook* contrast and this is a fundamental semantic datum. From it, by formal manipulations, the linguist reconstructs two different phones [*sic*] /b/ and /h/. To contrast means to not be a repetition: to say or to mean something else. This rudimentary, explicit semantic element was never eliminated from Harris's work and — I may add — is assumed in all linguistic efforts, even if that fact is not always recognized. It is present in phonetics, in syntax, in discourse analysis, in field methods, and in comparative studies. In addition, Harris uses semantics implicitly. The entire effect of Harris's syntax ... is oriented toward rendering semantic differences by syntactic means. His syntax is always semantically motivated, in spite of the changes in form through the years. It is not that the result of the syntax — the derived sentences — will later receive semantic interpretation, but that each syntactic step reflects or records a semantic property, a paraphrase being a particular case of such a property. (Hiž 1979: 344)

Here, semantics (meaning, information) is not a projective interpretation of syntax (form), but rather form and information are two views of the same thing, and Zellig's work was to make the correspondence of these two faces of the coin explicit and obvious by distinguishing the information- content aspect of language from the presentational aspects.⁵⁸

58. The separation of content from its presentation has become familiar to a new generation of computer scientists with the development of (crudely) semantic markup 'languages' for texts, especially SGML and XML.

Although the identification of phonological elements is semantic (as well as distributional) in nature, it is semantic to only a ‘rudimentary’ extent. They, or rather the contrasts that they represent, do not themselves correlate with meanings:

If in all the occurrences of a word the phonemes were replaced by others, we would simply have a variant form of the same word. But if we replaced some or all of the occurrences of a word by a word which had different selection, i.e. whose normal occurrences were different, we would have a different meaning. (Harris 1991: 322)

Zellig’s recognition that “inter-environment relations, or ... occupiyings of positions, [are the] fundamental elements” (Harris 1951a: 371, quoted above) has ramifications for semantics. If we accept the customary assumption that the meaning of a word determines in what environments it may occur (the word *owl* is only used in combination with those words with which it ‘makes sense’), then we must accept the converse: that the distribution of an element is a formal specification of its meaning, or rather, of that part of its meaning with which we are most concerned in linguistics.

As Leonard Bloomfield pointed out, it frequently happens that when we do not rest with the explanation that something is due to meaning, we discover that it has a formal regularity or ‘explanation’. It may still be ‘due to meaning’ in one sense, but it accords with a distributional regularity. (Harris 1954: 157)

If one wishes to speak of language as existing in some sense on two planes — of form and meaning — we can at least say that the structures of the two are not identical, though they will be found similar in various respects. (Harris 1954: 152)

Language is clearly and above all a bearer of meaning. Not, however, of all meaning. [...] Meaning itself is a concept of no clear definition, and the parts of it that can be expressed in language are not an otherwise characterized subset of meaning. Indeed, since language-borne meaning cannot be organized or structured except in respect to words and sentences of language ..., we cannot correlate the structure and meanings of language with any independently known catalogue or structure of meaning. (Harris 1991: 321)

In Zellig’s original form of discourse analysis there is a transform of the sentences of a discourse that regularizes its structure so that all repeated phrases and their ‘local synonyms’ can be placed in the same column of a table. Each row is a successive period of the discourse, each cell is a lexical item in the specialized lexicon of that discourse, and the column heads are the classifier vocabulary of that semantic domain. Zellig extended this in two ways. First, all discourses of a constrained subject-matter domain have discourse structures in common, exemplifying the sublanguage of that domain. They employ the same sublanguage lexicon and the same classifier vocabulary in a common specialized sublanguage grammar. This is especially clear in a technical

domain, such as the sublanguage of a science (Harris et al. 1989). Secondly, in the language as a whole there is a subset of sentences in which the correspondence of distribution with meaning is much more direct and transparent. This subset constitutes an informationally complete sublanguage with no ambiguity and no paraphrase (Harris 1969). The sentences of the language in which the correlation of form with meaning is less clear are paraphrases of sentences in this subset.

The correlation between distribution and meaning is sharpened progressively in Zellig's work, until with operator grammar we are concerned with the fuzzy acceptability gradation of operator-argument dependencies which, for a sublanguage, approach or reach a completely binary selection, where a given dependency is either fully acceptable or is rejected as nonsense (or non-science), and such fuzziness of selection as remains is a marker of controversy or change in the field (Harris et al. 1989).⁵⁹ From this work results a theory of linguistic information (Harris 1991).

This is a theory of the information that is 'in' language by virtue of its structure. Language users associate additional meanings with this information. "Correlations between the occurrence of linguistic forms and the occurrence of situations (features of situations) suffice to identify meanings; the term 'to signify' can be defined as the name of this relation." (Harris 1940: 228 [704 of reprint]).

12. The language, the whole language, and nothing but the language

To give more palpable substance to our understanding of how the work of Noam and Zellig has diverged, consider Noam's requirement (echoing Carnap) that a grammar is required to account for "all and only the sentences of the language." This is an idealization that omits many pertinent aspects of language structure, including on the one hand the sentence fragments that pepper ordinary discourse, and on the other hand most⁶⁰ restrictions that cross sentence bounds. For of course "Language does not occur in stray words or sentences but in connected discourse — from a one word utterance to a ten-volume work, from a monolog to a Union Square argument" (Harris 1952: 3). Even the conjunction of two sentences imposes word-sharing requirements. Consider this simple conjoining of sentences taken from two different books:

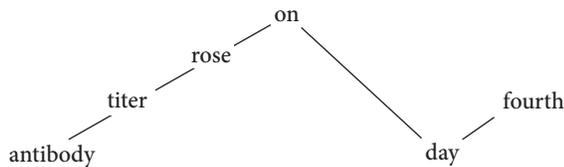
59. In some degree corroborating the so-called Whorf-Sapir hypothesis.

60. The treatment of pronouns and other referentials in a sentence grammar fudges this without addressing the larger questions of inter-sentential regularities.

*Let us make the assumption that a string X of I which is n phonemes in length is carried by ΦP_m into a string Y which is n phones in length and other views of Heidelberg can be seen in emblems, here reproduced for the first time.⁶¹

There is nothing wrong with this example from the point of view of a sentence-grammar of 'the language,' because such grammars cannot discriminate different domains of discourse. By that I do not refer to matters of presentation such as dialect, register, and style, but rather to subject-matter domain or (equivalently) the sublanguage of such a domain. The regularities that are found, and the elements and operations upon them adduced in a grammar, depend upon the domain or sublanguage being considered as well as the scope under consideration.

For instance the sentence *The antibody titer rose on the fourth day* (as may occur in a text of cellular immunology) can be represented in various ways by different sentence grammars: as having the phrase structure (S ($_{NP}$ *the antibody titer*) ($_{VP}$ ($_V$ *rose*) ($_{PP}$ ($_P$ *on*) ($_{NP}$ *the fourth day*)))) or a partially-ordered word dependence structure [as in Operator Grammar (Harris 1982, 1991)], representable as a semilattice



As occurring in a text in a sublanguage of cellular immunology, the sentence might be reconstructed, showing its similarities with other sublanguage sentences, as: *On the fourth day following the reinjection of viral antigen into the footpad of rabbits of the same strain, the titer of antibody rose in the lymph follicles*, which can be represented as an instance of the word class sequence GJB : AVT, with : = *on the fourth day following*; G = *viral antigen*; J = *(re)injection*; B = *rabbits of the same strain*; A = *antibody*; V = *titer present in*; T = *the lymph follicles*. (Ryckman 1986:246n)

This last representation is underwritten by the analysis and grammar of an immunology sublanguage as detailed in Harris et al. 1989. Each of these three representations

61. The first sentence is from (Chomsky 1975: 160), and the second is from (Yates 1972: 59), Ryckman (1986: 247) offers this illustration:

**Manifestation is a relation of a whole and its parts and it was all characteristically Teutonic, and, critically examined, not very tactful; but tact was never Wagner's strong suit when trying to convince the world that its only hope of salvation lay in hitching itself to the German chariot.*

of sentence structure enables one to reconstruct the well-formedness of the example sentence in accord with the sentence-defining elements and operations of a certain type of grammar. The elements and operations specified in each such grammar capture regularities shared by sentences that occur normally together in a discourse (or set of discourses in a sublanguage domain). But the project of “generating all and only the sentences of the language” entails an idealization that can capture only sufficient structure to assure that the sentence is sayable, regardless of whether it expresses any coherent meaning or has any likelihood of ever being actually said. This has artificially limited the notion of ‘grammaticality’ so that it contrasts with acceptability or meaningfulness.⁶² However, language users recognize structure (or “have intuitions of grammaticality”) much more extensively in respect to sentences as they occur together in a discourse.

Consider the utterance *It doesn't me either* (spoken in my hearing recently by my brother). Immediately prior context was a series of sentences of this sort: *He might have wanted to X. Or it might have been because of Y.* I had to remember four or five sentences prior, where his wife said *It doesn't make sense to me.* It is only in that distant context that the words *make sense to* are redundant so that zero allomorphs of them can occur in *It doesn't me either*. A classical Generativist gapping operation on a PSG-based tree structure is quite useless here.

Example sentences isolated in the metalinguistic context of grammatical discussion are not an idealization so much as a willful ignorance, which can warrant only very limited conclusions about the nature of language, and which gives rise to pseudoproblems.

62. Whereas for Zellig these attributes concur (grammaticality, acceptability, meaningfulness), Noam relegates acceptability to ‘performance’ in his revisionist Introduction to *LSLT* (Chomsky 1975: 7). He claims that Section 100.2 is an instance of making this distinction, although that section merely identifies a phenomenon that he does not know how to explain (and which not all informants recognize as an actual phenomenon, eerily suggestive of the disputes about Generative Semantics). In Section 11.1 (Chomsky 1975: 81–82) he rejects “informant response tests to determine the degree of acceptability or evocability of sequences” in favor of “a revealing general theory of which all [grammars thus approved] are exemplifications.” He may be obliquely rejecting the criterion for transformation that Zellig discussed in 1955 or earlier but only later in (Harris 1965) advanced as a replacement for ungraded selection-preservation (the foremost criterion in Harris 1957). However in correspondence with the author he has disclaimed any familiarity with Zellig’s writings since about 1955 (although clearly from the comment (Chomsky 1975: 38 and fn. 70) he had read (Harris 1965)). The ‘general theory’ or metagrammar that he posits as a desideratum in the above quotation is what he elsewhere reframes as a biologically innate Universal Grammar.

13. Pseudoproblems

As an example of such pseudoproblems, consider the notion that language users have an intuition of ambiguity, *e.g.* that a sentence such as *flying planes can be dangerous* is ambiguous. When such a sentence is no longer artificially isolated there is of course no ambiguity. Such context is called ‘disambiguating’, but it is rather the case that the isolation of sentences by linguists is ‘ambiguating’, and what are called intuitions of ambiguity are no more than the ability to imagine more than one context in which the isolated sentence might naturally occur. Noam has rejected this view, however:

Apparently many linguists hold that if a context can be constructed in which an interpretation can be imposed on an utterance, then it follows that this utterance is not to be distinguished, for the purposes of study of grammar, from perfectly normal sentences [...], though the distinction can clearly be both stated and motivated on syntactic grounds. [...] This decision seems to me no more defensible than a decision to restrict the study of language structure to phonetic patterning. (Chomsky 1964: 7n2)

There is a distinction, surely, but the proposal that it is a matter of grammaticality is coherent only given a prior commitment to restricting ‘grammar’ to a grammar of isolated sentences (with anaphora and other cross-reference a marginal problem).

A more extended class of pseudoproblems are the so-called island phenomena. These are problematic for grammatical operations on abstract phrase-markers derived from phrase-structure grammar. When Haj Ross was visiting Joshi’s class at Penn in 1969, and spoke, as one might expect, about island phenomena,⁶³ I told him that they are not an issue in string adjunction grammars (Harris 1962a, Joshi et al. 1968, 1972a, 1972b). He asked, with evident concern, “Is that true, Aravind?” and Joshi replied “Yes, but they [adjunction grammars] have their problems, too,” whereupon he relaxed and the discussion moved on. What Joshi meant was that rewrite rules and adjunction rules are complementary: rewrite rules handle exocentric constructions well but not endocentric constructions, and the converse is true for adjunction rules, for which the small set of center strings and adjunct strings is simply specified in a list.⁶⁴ At the time, Joshi

63. The subject of his 1967 dissertation (Ross 1967).

64. “[E]ach style [of formal grammar] ... is well suited for characterizing certain aspects of natural language structure and is awkward for characterizing certain other aspects. The awkwardness can be due to either an inherent difficulty in characterizing a certain aspect (*e.g.*, the characterization of the notion of the ‘head’ of a constituent in a PSG) or actually a counter-intuitive characterization (*e.g.*, this often happens in a PSG, especially in the context of transformational grammars, because a PSG allows an ‘uncontrolled’ introduction of new ‘non-terminals’). ... The main purpose is to set up a class of grammars which has no more

and his students were working on grammars combining the two types of rules so as to exploit this complementarity.⁶⁵ This led to tree-adjoining grammars (TAGs, see e.g. Joshi 1985), in which the (exocentric) center strings and adjunct strings are generated by very simple rewrite rules rather than being listed, and adjunction rules then handle the much more complex endocentric structures of language.⁶⁶ There is a natural transition from an adjunction grammar to a transformational grammar, by deriving adjunct strings from sentences.⁶⁷ In an Operator Grammar (Harris 1982, 1991), island phenomena fall out naturally from the linearization and reduction constraints, calling for no particular special attention (Nevin 1989). The exocentric properties of language arise from the dependency of operators on their arguments which must have previously entered into the construction of a sentence; the endocentric properties arise from the ('extended morphophonemic') reductions of the overt forms of the resulting strings.

To substantiate in more detail the assertion that island phenomena are pseudo-problems arising from use of phrase structure rules in the base would take us far too far afield in this context.⁶⁸ It is well supported by the relevant literature. Two points are especially germane to our present narrative.

Firstly, adjunction grammars form a hierarchy distinct from the Chomsky hierarchy (more correctly, the Chomsky-Schützenberger hierarchy) of formal grammars, intersecting it in interesting ways. This is not widely understood. The Chomsky hierarchy is often thought of, at least by computer scientists of my acquaintance,⁶⁹ as

power than necessary and which also can characterize different aspects of natural language structure in a natural way" (Joshi et al. 1968). Compare also Zellig's discussion of the relation of the Morpheme to Utterance grammar of string expansions to string analysis and to transformations in (Harris 2002: 3–4), quoted later in the present paper, and in the first section of (Harris 1965).

65. This may be why he was so close-mouthed.

66. Zellig had developed string analysis (published after some delay as Harris 1962a) as an alternative to constituent structure, providing the basis for the 1959 computational analysis system at Penn (Joshi 1959, 2002).

67. This was noted at the very beginning, in the 1930s and 1940s (Harris 2002: 3–4). Naomi Sager and her group exploited this relationship in implementing the very successful Linguistic String Processor (LSP) and the Medical Language Processor (MLP) at NYU. The MLP system uses XML (mentioned earlier) in the data-organization part of its implementation (Sager & Ngô 2002). Most of the work is in refining the subclasses of the lexicon. The same may be said for the development of a sublanguage grammar and lexicon. This work is conveniently summarized in (Sager 1984; Sager & Ngô 2002). Successful computer implementation is a demonstration that speaks far more strongly than argument.

68. For a detailed study from this point of view see (Nevin 1989).

69. Or in their case even a restricted subset of grammars in that hierarchy. See (Manaster-Ramer & Kac 1990: 328–329): "[T]he term *phrase structure grammar* is often — in theoretical

comprehending or outlining the full range of formal grammars, although it concerns only rewrite grammars of the sort invented (as noted earlier) by Emil Post.⁷⁰ Yet PSG trees of abstract pre-terminal symbols have become in effect indicators of paradigm membership for writings on syntax.

Secondly, a principal reason that adjunction grammars, tree-adjointing grammars, and Operator Grammar are more suited to natural language than PSG-based formalisms is that semantic relations hold between adjacent morphemes and words.⁷¹ In adjunction grammars, the initial strings are simple enough to meet the contiguity requirement, and when adjunct strings that are inserted, creating discontiguities in the previously present host string, they are inserted next to their head word.⁷² There is no problem of constraints at a distance.

computer science always [...] — used to denote type-0 (unrestricted) grammars, which generate all the recursively enumerable languages, a proper superset of the context-sensitive languages. [etc.]”

70. The essential differences between language and language-like formal systems such as computer ‘languages’ should be borne in mind. Language subsists essentially in spoken form of which any graphic form is a representation for convenience. A formal language such as that generated by phrase-structure rules subsists essentially in graphic form, of which any spoken form is a representation for convenience. In his *Mathematical Structures of Language*, Zellig notes “certain apparently universal and essential properties of language, which are observable without any mathematical analysis, and which are such as to make possible a mathematical treatment” (1968: 6). These are that language elements are discrete, socially preset in speaker and hearer, and arbitrary, that combinations of these elements are linear and denumerable, that not all combinations constitute a discourse, that operations are contiguous, that the metalanguage is in the language, and that language changes. “Each of these properties has metatheoretical consequences for linguistics.” In a formal language, the elements are discrete, preset by explicit definition, contingent on the terms and conventions of definition, operations need not be contiguous, the metalanguage is in natural language, and the language does not change (changed terms, definitions, etc. produce a different language).

71. Word-order phenomena are also defined locally. This is important computationally. See Joshi & Rambow (2003).

72. Insertion of adjuncts accounts for almost all discontiguities. The adjunct that ends up most distant from the head is inserted first, and becomes separated from it by subsequent adjunctions. In string grammar, the head of an adjunct string may possibly be separated from the insertion point at the head word by other words in the adjunction string itself, but an alternate word order is available that permits contiguity. This is exploited by the extension to transformational grammar (and to operator grammar), in which e.g. an adjunction string in which a word is fronted may be a reduction from a sentence in which that word is fronted. The procedure of string analysis is to ask repeatedly what is a least part that can be excised from a sentence without destroying sentencehood. The result is to identify each least increment of information to the sentence.

This is easier to see in Operator Grammar, where the source of all adjuncts is the reduction of an interrupting parenthetical sentence. For example:

- a. ? *books* — *books (same as prior) I like especially well — are in this box.*
- b. *books which I like especially well are in this box.*

This example demonstrates that the required metalanguage assertions that two words have the same reference are available in the language itself. (The metalanguage is a sublanguage in the language.) Co-reference is a condition for the reduction to wh-pronoun. In the example, co-reference of the two occurrences of the word *books* is expressed here by the phrase *same as prior*.⁷³ If it were expressed by an index (e.g. *books_i — books_i I like especially well — are in this box*), that index marker is no more than a graphical representation of that metalanguage assertion in words. This is explicitly seen when the use of such index tags is explained or defined, as it must be.⁷⁴

However, the indexed terms must be contiguous in order for this notational convention of subscript indices to be interconvertible with metalanguage assertions that are intrinsically available within language.⁷⁵

Indices were invented precisely to address arbitrary points in a string, and points that are separated by arbitrarily long intervening strings cannot be addressed by stating a simple relationship in language (e.g. *prior*). Non-language indices are an unavoidable consequence of the use of rewrite rules in the base, because unbounded extraneous material may intervene between coreferential terms. But in Operator Grammar, as in adjunction grammar and tree-adjointing grammar, co-referential terms are contiguous at the time that the reduction takes place, becoming separated from each other only

73. For a more careful formulation, see Harris (1968, 1982, 1991). It might be objected that *same as prior* is itself an adjunct phrase requiring a sentential source and reduction under metalinguistic sameness. Note that lexical items in a sublanguage lexicon often are lexically complex when considered in a broad-coverage grammar or in the grammar of another sublanguage. For example, *has high fever* might be a single lexical item in the **Symptom** class in a medical sublanguage. Other examples were seen in the quotation above from Ryckman (1986), e.g. *rabbits of the same strain* as a single lexical item in the **B** lexical category of the immunology sublanguage.

74. And for comprehension such definitions in words must be in the hearer's context (tacit or overt) when reading examples that employ such index tags.

75. For some types of reductions, 'nearby' suffices. See (Harris 1982; Nevin 1984) for details. All anaphoric and epiphoric reference is handled by the same sort of zeroable metalinguistic sameness statement. So-called crossing coreference, as in *The man who shows he deserves it will get the prize he desires*, is not problematic for the reductions in Operator Grammar, since they are not defined in terms of constituent structure.

by the subsequent entry of additional sentences which may be reduced to adjuncts.⁷⁶ The same problem arises for semantic relations between words (selection restrictions), mandating in addition to indices an extra superstructure of semantic features or the like.

This illustrates a contrast between the two approaches to linguistics. Noam's construal of a particular 'tool of analysis' (immediate constituent analysis) as a theory of grammar (phrase-structure grammar enriched by transformation rules) and his commitment to it as the 'correct' theory until disproven and replaced by another, constrains what he takes to be worthy of consideration, and focuses attention on anecdotal examples set up as proof or disproof of one detail of theory or another. (Even after the radical reformulations in Government and Binding Theory and in Minimalism, the machinery of abstract phrase-markers has been retained.) The examples and counter-examples of one stage are only of historical interest at a later stage (rule ordering and island phenomena, for instance, are *passé*, no longer hot topics). Zellig used diverse 'tools of analysis' to disclose different properties of language (constituency is good at endocentric constructions, string analysis at exocentric constructions, etc.). "It is not that grammar is one or another of these analyses, but that sentences exhibit simultaneously all of these properties" (Harris 1965: 365). Properties of language, once identified, are not subsequently set aside or replaced, though they may be subject to further analysis (as in the factorization of transformations in Harris 1964, 1969). This freed him to develop broad-coverage grammars and enabled him to defer commitment to a theory of language and information until the relevant properties of language had become clear.

14. Grammatical machinery is expressed in a metalanguage

Zellig realized that⁷⁷

There is no way to define or describe the language and its occurrences except in such statements said in that same language or in another natural language. Even if the grammar of a language is stated largely in symbols, those symbols will have to be defined ultimately in a natural language (1991: 274).

This means that any apparatus of grammar is only an abbreviation (or representation by symbols) of that which can also be expressed by metalanguage assertions conjoined

76. Note that Noam concurs in the obvious fact that the language in which grammars are formulated is a metalanguage of the language which they describe (Chomsky 1975: 116).

77. I recommend to your attention the more careful yet concise account of the thinking that led to recognizing the role and status of the metalanguage as a sublanguage, given at (Harris 2004[1990]: 7–9).

to sentences, but which are not usually uttered in overt phonemic form because they are common knowledge. Examples of zero allomorphs like the zero form of the plural morpheme in *sheep* are of course well accepted. The zeroing of redundant or low-information morphemes in e.g. so-called gapping (*John plays piano and Mary violin*) is an obvious and uncontroversial reduction in Operator Grammar. The zeroing of common-knowledge conjuncts, such as the dictionary sentence in *It's raining [and an umbrella is for protection from rain or sun] so I'll take an umbrella* is a straightforward extension. It enables a simple statement of the word-sharing requirement for conjoined sentences illustrated above with a conjunction of sentences taken from two different books. Similar to this, then, is the zeroing of metalanguage assertions that constitute the machinery of grammar, e.g. for co-reference. (Noam's solution to this dilemma has been to assert that the machinery of grammar must be innate by "What else could it be?" arguments, as noted earlier.)

This also means that linguistics cannot rely upon any prior metatheory of language.

In the absence of an external metalanguage, the entities of each language can be identified only if the sounds, markers, or words of which they are composed do not occur randomly in utterances of the language. That is, the entities can be recognized only if not all combinations occur, or are equally probable. This condition is indeed satisfied by languages. A necessary step, then, toward understanding language structure is to distinguish the combinations of elements that occur in the utterances of a language from those that do not: that is, to characterize their departures from randomness. (Harris 1988: 3)

In addition, it imposes a requirement on the results of linguistic analysis:

This task entails an important demand: it calls for a least description, that is, for a characterization of the actually occurring combinations by means of the fewest and simplest entities and the fewest and simplest rules and conditions of their combination, and with no (or least) repetition. The reason for this demand is that every entity and rule, and every complexity and restriction of domains of a rule, states a departure from randomness in the language being described. Since what we have to describe is the restriction on combinations in the language, the description should not add restrictions of its own. If two descriptions, one more efficient than the other, characterize the same data, then the less efficient description must have overstated the actual restrictions in the language — by overstating and then withdrawing part, or by repeating a restriction, or whatever. (Harris 1988: 3–4)

[T]here is also a responsibility to formulate a theory based on self-organizing capacities: one that will present language as a system that can arise in a state in which language does not exist. This is so because of the unique status of language as a system which contains its own metalanguage. Any description we make of a

language can only be stated in a language, if even only in order to state that some items of the description are properties of some other items (i.e. how to read the table). We cannot describe a language without knowing how our description can in turn be described. And the only way to avoid an infinite regress is to know a self-organizing description, which necessarily holds also for the language we are describing even if we do not use this fact in our description. (Harris 2002[1990]: 10)

There are two important contrasts here with Noam's work. First, Noam has asserted from an early stage that linguistics cannot proceed unless informed by a prior metatheory of language.

It is elementary that theoretical investigation and collection of data are independent activities. One cannot describe a linguistic system in any meaningful way without some conception of what is the nature of such a system, and what are the properties and purposes of a grammatical description. (Chomsky 1956a: 2–3)

Where Zellig stipulates only the restrictions on method (distributionalism) and on formulation (least grammar) due to absence of an external metalanguage, Noam rejects distributionalism and asserts a need to know “what is the nature of ... a [linguistic] system, and ... the properties and purposes of a grammatical description” of such a system. He simply asserts, without argument, that no alternative is possible.

For this reason, it is important to develop a precisely formulated and conceptually complete construction of linguistic theory, based on the clearest possible elementary notions, even when the more elaborate constructions based upon these notions cannot, because of insufficient evidence, be empirically supported. The establishment of such a theory may be an essential step towards obtaining this evidence. (ibid.)

Here, in the 1956 draft of *LSLT*, this amounts to a leap of faith. Formulating one's conclusions in advance may be an essential step towards obtaining evidence for those conclusions — or maybe not. Subsequently, as we know, that faith was placed in the innate biological endowment of humans. That is the second point of contrast, his attribution of metalanguage functions to a biologically innate Language Organ or (more recently) Language Function.⁷⁸ In the event, what was ambitiously presented in 1956 as a sequence (first theory, then evidence) was later reverted to a bootstrapping operation

78. We would be remiss not to recognize the reduction of the Language Function in its narrow sense (exclusive to human language) to just the property of recursion, and the ensuing disputation. See Hauser, Chomsky, & Fitch (2002); Pinker & Jackendoff (2005); Fitch, Hauser & Chomsky (2005); Jackendoff & Pinker (2005). Zellig's discussion of the nature and probable evolutionary origin of language (Chapter 4 of Harris 1988 and Chapter 12 of Harris 1991, esp. pp. 365–373) is much more simple and straightforward, as a natural consequence of accounting for information-bearing constraints in language rather than accounting for formal

in which theory guides investigation of examples and counterexamples which motivate revisions of theory, and so on, with no end in sight.

These matters constitute some of the *substance* of the divergence of views that we are considering, and they figure importantly in various miscommunications and misconstruals over the years, but lest we stray too far from our present theme, discussion elsewhere must suffice (Harris 1991, 2002[1990]; Nevin 1993b, 2002).

Noam has said (p.c.) that this talk of the metalanguage is not anything that he remembers hearing from Zellig. The importance and function of the intrinsic metalanguage is implicit in Zellig's writings of the 1940s and 1950s, being the essential motivation of distributionalism and what is commonly taken as "avoidance of meaning," although it may not have been clearly communicated until after Noam had (by his account) ceased to pay attention to what Zellig was saying and (by any account) was firmly established in his views on innateness. The crux of the matter for the present discussion is that Zellig recognized (and demonstrated) that the metagrammatical resources that are available within language itself suffice, and necessarily so; whereas Noam assumes that the requisite metagrammatical resources are necessarily external to and prior to language, given in the biological endowment of humans. Zellig's assumption freed him to employ diverse methods of analysis to disclose the properties of language, and at the end of his career he organized these properties into a theory of language and information (Harris 1988, 1991). Noam's assumption constrains him to treat each method as a theory of language, and to determine, by deploying examples as disproof, which one is correct.

15. Generalization is not the same as abstraction

As we have seen, Zellig's concept of transformation came directly from linear algebra. The discussion in (Harris 2002.3–4) is instructive:

The relevance of the hierarchy of word expansions [Harris 1946] ... was not simply in providing a direct procedure that yielded the structure of a sentence in terms of its words, but in opening a general method for the decomposition of sentences into elementary sentences, and thus for a transformational decomposition system. This unexpected result comes about because, first, the small sentence which is at the base of the expansions is recognizable as the grammatical core sentence of the given sentence, and, second, each expansion around a particular word can be seen to be a reduction or deformation of a component sentence within the given one. The status of expansions as component sentences was visible from the beginning: when the expansion method was presented at the Linguistic Institute, a question was raised as to how the method would distinguish the two meanings

properties of language-like mathematical systems that are proposed to be (or describe) the means by which language users produce and understand language.

of *She made him N* in *She made him a good husband because she made him a good wife*; the answer was in showing that two different expansions obtained from two different component sentences yielded here the same word sequence (sec. 7.9 in [Harris 1946]). The expansion analysis was formulated later as a decomposition of the given sentence into word strings[...]79

While the machinery for transformations was provided by the “Morpheme to Utterance” equivalences, the motivation for developing transformations as a separate grammatical system was furthered by the paraphrastic variation in sentences that was found in discourses. In 1946, with the completion of *Methods in structural linguistics*, the structure of a sentence as restrictions on the combination of its component parts seemed to have gone as far as it could, with the sentence boundaries within an utterance being the bounds of almost all restrictions on word combination. I then tried to see if one could find restrictions of some different kind which would operate between the sentences of an utterance, constraining something in one sentence on the basis of something in another. It was found that while the grammatical structure of any one sentence in a discourse was in general independent of its neighbors, the word choices were not. [This is related to the word-sharing requirement noted earlier. — BN]

In a discourse, the component sentences revealed by the Morpheme to Utterance expansions were often the same sentence appearing in different paraphrastic forms in neighboring sentences. The use of reductions and deformations of sentences

79. That is, string analysis is a later formulation of the expansion analysis. This refutes the common mistake of judging the expansion grammar of (Harris 1946) to be a notational variant of phrase structure grammar. Just prior to the above quotation he writes:

[N]o a priori justifiable general method was found to reach the structure of a sentence (or an utterance) by a hierarchy of constituent word sequences, or other parital structures of words. The problem was finally resolved by a single general procedure of building, around certain words of a given sentence, graded expansions in such a way that the sentence was shown to be an expansion of a particular word sequence in it, this word sequence being itself a sentence. (Harris 2002: 3)

Where (Harris 1965: 363) contrasts string analysis with constituent analysis, he is not talking about the expansion grammar of 1946, but rather the immediate constituent analysis of Bloomfield. This distinction should be obvious upon reading even the first paragraphs of (Harris 1946). Constituent analysis depends upon native speaker judgments of where to bisect sentences and the resulting parts of sentences, on the assumption that such bisection is fundamental to human psychology. The expansion analysis, instead, depends upon judgments of the substitutability of strings of morphemes (with single morphemes being the limiting case). String analysis, conversely, depends upon judgments of successive least portions of a sentence that may be excised without destroying sentencehood. X-bar notation, head-driven PSG, and other innovations aim to capture the endocentric capacities of the expansion grammar. So-called bare phrase structure (BPS) proposed in Noam’s Minimalist Program appears to be a more direct attempt to capture Zellig’s original insight, to the extent that the notion of ‘projection’ corresponds to expansion from word-class to construction-class.

both to produce expansions and also to produce separate paraphrastic forms of a sentence motivated the formulation of a whole transformational system, and [Harris 1952] was presented to the Linguistic Society of America in 1950.⁸⁰... The transformational system...was presented to the Linguistic Institute at Indiana University in 1951–1952. The formal presentation, with detailed structural-linguistic evidence that the expansions were indeed transformed sentences, was given at the [LSA] meeting in 1955 ([Harris 1957]). (Harris 2002: 3–4, footnote added)

Taking the notion ‘transformation rule’ from Carnap, Noam thought of transformations as an algorithmic rule of a new type acting on the productions of the phrase-structure rules. In this we see a critical parting of ways in their treatment of language and their understanding of linguistics, and (quite apart from questions of temperament) the basis of many mischaracterizations of Zellig’s work.

Noam’s development of transformations as operations on abstract tree structures derived from PSG has led to an ever more abstract metatheory of grammar. Zellig’s work demonstrates that such abstract treatment is not required, and that properties of this grammatical machinery actually obscure the essential properties of language. Zellig recognized this:

There is an advance in generality as one proceeds through the successive stages of analysis above [from structural linguistics through transformational linguistics to operator grammar]. This does not mean increasingly abstract constructs; generality is not the same thing as abstraction. Rather, it means that the relation of a sentence to its parts is stated, for all sentences, in terms of fewer classes of parts and necessarily at the same time fewer ways (‘rules’) of combining the parts, i.e. fewer constraints on free combinability (roughly, on randomness). But at all stages the analysis of a sentence is in terms of its relation to its parts — words and word sequences — without intervening constructs. Because of this fact, and because the parts which are determined are such that their meanings are preserved under the sentence-making operations, the meaning of a sentence as a particular combination of particular words is obtained directly as that combination of the meanings of those words. (Harris 1981: v–vi)

The abstractness of Generative Grammar seems to arise primarily from the abstractness of phrase structure trees. For Noam, language is at root a logico-mathematical system, and transformations are rules that operate on tree structures generated by phrase-structure rules.⁸¹ In recent years, the rules have been factored into interactive

80. And prior to that “was given at an anthropological congress held in 1949” (Hymes & Fought 1975: 153).

81. I take the substantive differences to be in Zellig’s and Noam’s respective treatments of syntax and semantics. Setting aside for the present the claims of Generative Phonology, discussed earlier, we can see that they both start with the discrete elements of language, whether identified by language users (and by descriptive linguistics) in the speech stream or directly observable in

modules with cumulative effect, but they still operate on abstract structures of preterminal symbols.⁸² In addition, abstract treatment has been useful to Noam for controlling the terms for argument and debate, a labyrinthine playing field with which he alone is most familiar, all the more so to the extent that others have not been so free as he to redefine the field from time to time (as exemplified by R. Harris 1993).

For Zellig, transformations are algebraic mappings from subset to subset of the set of sentences. For convenience, each subset is represented by an n -tuple of morpheme classes and constants⁸³ (called a sentence form), but the relationship is between the actual correspondant sentence pairs that are the subset members. This illustrates what he means by generalization as distinct from abstraction. The set of transformations themselves turned out to be factorable into elementary sentence-differences (Harris 1964, 1969). These in turn were restated as the effect or increment of an operator word entering on the words in its argument, with reductions ('extended morphophonemics') being carried out at the time that word entry creates the requisite string conditions⁸⁴ (Operator Grammar, Harris 1982). Each tool of analysis that he developed and used along the way (word expansions, string structure, transformational structure, etc.) revealed or favored some of the properties of language and disfavored others.

It is not that grammar is one or another of these analyses, but that sentences exhibit simultaneously all of these properties. [...] Each of these properties can be used as the basis for a description of the whole language because the effects of the other properties can be brought in as restrictions on the chosen property. (Harris 1965: 364)

writing, and with the basics of morpheme classes and subclasses. Except for intonation, issues of phonology are encapsulated, that is to say, they do not enter into examples used in discussions of syntax and semantics, which even in an 'exotic' language are presented with nothing more phonologically sophisticated than a practical orthography.

82. Morphemes were initially understood to enter these structures by a subsequent step of lexical insertion at each terminal symbol. Subsequently, Noam introduced the Projection Principle to alleviate some of the problems that were encountered. Since these problems do not arise in Zellig's Operator Grammar or in sublanguage grammars, the Projection Principle is evidently an artifact of notation.

83. A single morpheme may function as a constant in a sentence form, as e.g. the *-ing* of the gerund.

84. The string conditions for reduction are created by the product of an operator word entering on the product of some prior operator(s) and the reductions previously carried out on them, or on primitive arguments (mostly concrete nouns). The linearization of operator and arguments is also determined at the time of operator entry into the construction of a sentence. See Harris (1982, 1991) for details.

In a footnote, he added: “The pitting of one linguistic tool against another has in it something of the absolutist postwar temper of social institutions, but is not required by the character and range of these tools of analysis.”⁸⁵

With Operator Grammar all the identified properties of language could be accounted for in a reasonable and comprehensive way, and in a way that supports a theory of linguistic information. The role of redundancy in this theory (and, crucially, in the linguistic analysis that led to it) is clearly related to mathematical information theory, but whereas information theory concerns only quantity of information in a channel, the theory of linguistic information concerns information content.

The problem here was not to find a broad mathematical system in which the structure of language could be included, but to find what relations, or rather relations among relations, were necessary and sufficient for language structure. Then anything which is an interpretation of the model can do the work of natural language. Modifications of the abstract system are then considered, which serve as models for language-like and information-bearing systems. (Harris 1968: v)

I commented to Zellig once that descriptions by Generativist linguists seemed unnecessarily complex. He only said “They do seem to be over-structured,” with a slight smile. Phrase-structure trees introduce layers of abstraction that are neither needed

85. In correspondence with the author (18 December 1999) Noam took this personally, as follows: “[H]e had no comment, no suggestions, and as far as I am aware no familiarity with or interest in anything that I did in generative grammar, from my undergraduate thesis through later years. The only comment I recall in print is a remark about how it reflected the Cold War atmosphere.” Noam has frequently disclaimed Zellig’s involvement or interest in his work, e.g. Barsky (1997: 53). However, he reported a very different memory in his 1975 Introduction to *LSLT*: “While working on *LSLT* I discussed all aspects of this material frequently and in great detail with Zellig Harris, whose influence is obvious throughout” (Chomsky 1975:4). Zellig’s student and friend Bill Evan has told me that on a visit to the Harrises at Princeton, where they lived during the time Zellig’s wife, the physicist Bruria Kaufmann, was assistant to Einstein, he found Zellig and Noam “going at it hammer and tongs” with the manuscript of *LSLT* (Chomsky 1955) spread out on the kitchen table. The only explanation I have for this self-contradictory denial of Harris’s help and influence is that he was never able to convince him. If Zellig said “That’s not what I’m doing, but you are welcome to try” (words very similar to what he did in fact say to me on one occasion), that wouldn’t be enough. Noam would feel that if Zellig only just really understood, he would necessarily agree. So he would keep arguing. But for Zellig such argument was beside the point. If Noam had a different way of working, fine. Zellig wanted to see what would result from doing the work. As he wrote to Goetze, in the letter quoted earlier, “I [do] not consider it ‘unpleasant’ but am glad of the controversy. No person, certainly not I, can be sure of his judgments as ‘always right’; the best way to get closer to the ‘truth’ — after I have figured out whatever I could — is to get the divergent opinions which arise from a different scientific analysis. The only fun in science is finding out what was actually there.”

nor desirable. Aravind Joshi has called this the pseudo-hierarchy of preterminal symbols. “Zellig Harris pursued the strategy of eschewing as much hierarchical structure as possible in describing sentence structure” (Joshi 2002: 121). In operator grammar, the hierarchy is of operators entering on their arguments; in tree-adjointing grammars (TAGs), Joshi has used rewrite rules because they are widely understood, but has also shown that dependency rules may be used for this limited hierarchy. The bulk of syntax is linear adjunction of words and strings of words (including morphophonemic reductions of words), requiring no hierarchy or nested bracketing.

In the comparison of theories, there is a diagnostic that to my knowledge has not been applied. If phenomena are reported as puzzles or issues by practitioners of one paradigm, but for another they are not problematic, or do not even merit status as phenomena, that disparity strongly suggests that what is involved is an artifact of method, of notation, or of proposed theory or metatheory. When we consider that issues with island phenomena, the Projection Principle, the Trace Erasure Principle, and so much else simply do not arise with Operator Grammar (nor with some other types of grammar, such as Stratificational Grammar), it is difficult to avoid the supposition that the Generative enterprise has been notationally entrapped.

16. Psychological realities

In Noam’s writings of the 1950s, there is no suggestion of psychological or biological considerations, either as an interpretation of an assumed metagrammar called UG, or as its hypothetical origin.⁸⁶ Only later is there any claim that ‘linguistic competence’ (particularly of children as language learners) derives from a biologically innate ‘language organ.’⁸⁷ In (Chomsky 1951: 4), as we saw near the beginning of this paper,

86. Revisions in (Chomsky 1975), and statements in the Introduction to that work, might suggest the contrary, but in e.g. (Chomsky & Ronat 1977: 123, Chomsky & Ronat 1979: 113) Noam affirms that “The psychological point of view did not begin to appear until the end of the fifties”. Just before this (Chomsky & Ronat 1979: 111), Ronat asks “When did you think for the first time of proposing an explanatory theory in linguistics?”. In English, Noam replies “That was what interested me about linguistics in the first place”. In the earlier French publication, his reply is given as “En fait, c’est une préoccupation qui est apparue très tôt, dès mon travail sur la morpho-phonologie de l’hébreu moderne.”

87. In passing, we should note that Bloomfield’s ‘failure’ to propose an innate language organ was not because he was a behaviorist, but because that was (and is) an undisprovable explanatory principle, and in any case is not required for linguistics — as Zellig has decisively demonstrated. All that was disproved in Hauser et al. 2001 was the purported language specificity and human specificity of the cognitive capacities involved with language use (with recursion still hanging on by its fingernails). So innateness lives on, as explanatory principles must until

[T]he statement of the grammar, the presentation of the results of the completed distributional analysis, must meet ... criteria which involve, essentially, considerations of elegance and considerations of adequacy as determined by the particular purposes of the grammar.

Around 1960, instead of the “purposes of the grammar” it is the speaker’s linguistic competence that determines the relevant “considerations of adequacy”. This psychologization of generative grammar transpired after the commencement of Noam’s association with a cohort of young psychologists in Cambridge who were forging the new field of cognitive psychology in rebellion against the behaviorists. We may presume that in their company he became familiar with one of their touchstones, Lashley’s (1951) critique of behaviorism.⁸⁸ Seuren (2009) suggests that the section of Lashley’s paper that discusses language provides a kind of skeleton for much of the argument in Noam’s own now famous (1959) review of Skinner (1957), which indeed cites and quotes that paper.

A proper discussion of psychology and linguistics must be reserved for another place, but two aspects require brief notice here. Firstly, I must acknowledge that Noam’s algorithmic bent has a natural home in psychology. Whatever the structure and properties of language may be, humans are somehow able to use language so as to achieve ends which require the mutual understanding and cooperation of others. A generative model of that accomplishment should be able produce utterances as an individual human does, and to demonstrate understanding by acting appropriately upon hearing what one has said. Results so far have been limited — for example, question-answering systems correlate utterances in a highly restricted domain with database queries, so that it is a relational database system rather than linguistics that is doing the work.⁸⁹ The reason for this is the same as the reason that robot systems are operational only in artificially constrained environments or when augmented with human direction. That reason is that the environment is unpredictably variable. Algorithms that can replicate observed behavior in unconstrained environments involve continuous negative feedback control of perceptual inputs in accord with internally maintained reference values

alternatives are recognized — “the slaying of a beautiful hypothesis with an ugly fact”. It is unlikely that Zellig’s demonstration that we have no need of that hypothesis will affect matters. Null is hard to recognize as an alternative or as a fact.

88. Based on his presentation at the Hixon Symposium in 1948, see (Gardner 1985: 10–14).

89. In an analogous way, the metagrammatical systems that Noam has developed correlate sentences with ‘semantic representations’ of one kind or another in which meanings are supposed to be formally indicated. As discussed earlier, Zellig has shown that the information carried by language is immanent in language itself.

for those perceptions.⁹⁰ The explanation of this is beyond the scope of this paper; the interested reader is directed to Powers (1973, 1998, 2008); Marken (1992, 2002); Cziko (1995, 2000); and Runkel (1990, 2003).

Secondly, we must touch on Noam's convictions about innate ideas. Lin (1999) provides a useful survey of the issues. Innateness is a stool with three legs: (1) infants' limited cognitive abilities, (2) paucity of data, and (3) complexity of language. Ideas about infant cognition which were available to Noam in the late 1950s and early 1960s have long been superseded by research showing the remarkable learning capacities of infants, and statistical learning theory has vitiated Noam's early claims that empirical learning from experience is not possible (see e.g. Gleitman 2002; Aslin & Newport 2009; Gebhart, Newport & Aslin 2009; Reeder, Newport & Aslin, in press). One of the founders of Cognitive Psychology, Jerome Bruner, demonstrated (as others have elaborated since) the supportive social scaffolding within which children learn language (Bruner 1985), challenging the paucity of data thesis.⁹¹ And Zellig has shown that the structure of language is actually remarkably simple.⁹²

90. In engineering control theory, these 'setpoints' are set by an agent or operator external to the 'plant'. In hierarchical perceptual control theory, or HPCT, they are set by the outputs of control loops at a higher level of the perceptual hierarchy. Error in control of biologically innate 'intrinsic values' necessary for survival triggers reorganization within the hierarchy, so infinite regress in a finite perceptual hierarchy is not an issue. By changing synaptic connections, reorganization can change any phase of the control loop, including the reference signal input. Error in control of other variables can also trigger reorganization. It is this that behaviorism studied as a theory of learning. Obviously, at any given time many perceptions are not controlled, and just as one may be aware of any given perception or not, so also may control of a perception be conscious or unconscious.

91. Bruner also opposed himself to Noam's philosophical Realism, holding that, rather than discovering the organization of phenomena, science invents the "scientific arrangement" of them (Bruner et al. 1962: 7), reminding us of that same locution in Zellig's review of Trubetzkoy (Harris 1941b).

92. "[T]he very simplicity of this system, which surprisingly enough seems to suffice for language, makes it clear that no matter how interdependent language and thought may be, they cannot be identical. It is not reasonable to believe that thought has the structural simplicity and the recursive enumerability which we see in language. So that language structure appears rather as a particular system, satisfying the conditions of Chapter 2 and perhaps also bound by a history, which may evolve not only in time but also by specialization in science languages, and which is undoubtedly necessary for any thoughts other than simple or impressionistic ones, but which may in part be a rather rigid channel for thought." (Harris 1968: 216) The reference to Chapter 2 is to the enumeration of essential properties of language summarized earlier in this paper, properties that make language amenable to mathematical treatment.

Just in order to see what there is in language, and whether it is unique, we can even in principle count the demands (the departures from randomness) in language. We can count the demands that suffice to enable a person to speak a given language. [...] Nobody will do this counting, but we can see that there is nothing magical about how much, and what, is needed in order to speak. Finally, and this is perhaps more important, we can see roughly what kind of mental capacity is involved in knowing each contribution to the structure — in knowing phonemic distinctions; in knowing the phonemic composition of words; in knowing the requirement status of words, i.e., their dependence on the occurrence of other words; in knowing the (mostly pairwise) likelihoods of operator-argument choice and the rough meanings attached to each word; and finally, in knowing the reductions in phonemic shape of given words in operator-argument situations. The kind of knowing that is needed here is not as unique as language seems to be, and not as ungraspable in amount. (Harris 1988: 111–113)

Given the observation that language contains its own metalanguage, Zellig specified methods which themselves require no additional metalinguistic or metagrammatical assumptions, and followed where those methods led, disclosing essential properties of language, which he then formulated as a theory. Noam begins with the contrary assumption that research is impossible without a prior theory of language, and that this prior theory is justified by complex metalinguistic and metagrammatical resources provided by a biologically innate Universal Grammar existing prior to and external to any particular language. His arguments then are of the “What else could it be?” sort, in which you argue against not-X in order to prove X. But because science does not prove anything, *tertium non datur* (the law of the excluded middle) is not an appropriate argument form for science. Universals of language, for example, can exist without being genetically inherited, as for example quantal vowels and the universal preference for certain means of making contrasts have an acoustical basis. Zellig has shown how other universal properties of language have an informational basis, such that without them utterances could not embody and ‘convey’ the information that they do. To show this, he has not had need of Noam’s hypothesis of an innate Universal Grammar.

17. Revolutionaries

We come now to the role of academic politics and the question of what can possibly be meant by ‘revolution’ as applied to Noam’s career. Noam’s portrayal of himself as isolated and disregarded in the 1950s because no one was interested in what he was doing, and that the old guard of ‘taxonomic linguistics’ opposed the rise of Generative Grammar, has been thoroughly debunked by (Murray 1994, Chapter 9). After signing off on his dissertation, Nelson Goodman and Zellig helped Noam to get a Research

Fellowship at Harvard, which he held from 1951 through 1955.⁹³ In 1956, Noam was one of the speakers at the Symposium on Information Theory held at M.I.T. (Chomsky 1956b), an occasion which the participants mark as the birth of Cognitive Psychology.⁹⁴ In 1959, when Archibald Hill invited Noam to speak at a conference in Texas, he delivered an early version of the caustic attack on ‘taxonomic linguistics’ which was renewed at the Ninth International Congress of Linguists, held near M.I.T. in Cambridge in 1962. I have been told that Zellig had been invited as a plenary speaker, but was unable to attend, and promoted the acceptance of Noam to speak in his stead. One of the organizers was Noam’s colleague and, as it were, co-conspirator, Morris Halle (a friendship dating from 1951 — Chomsky 1975: 30). The plenary session at which Noam spoke was devoted to “The logical basis of linguistic theory,” practically identical to the title of *LSLT* (Chomsky 1955a, 1956a, 1975).

[O]rganizational leadership [of Generative Linguistics was provided] by Halle, who helped Chomsky obtain a position at M.I.T., instituted first a degree program and then a department which he headed..., connected Chomsky with the publisher Mouton (and with its *raison d’être*, Roman Jakobson), and was involved in getting Chomsky into the spotlight in the 1962 international congress. (Murray 1994: 243)

He alone of all presenters at the International Congress was given opportunity and a free hand to revise and expand his remarks for the proceedings. His presentation in this plenary session was the attack on ‘taxonomic linguistics’ which was later expanded in *CILT* (Chomsky 1964 and its successive revisions), as discussed earlier.

It has been amply documented (*e.g.* Murray 1994 and elsewhere) that the rise and rapid expansion of Generative Linguistics in the 1960s was made possible by a sharp increase in the availability of government funding, analogous to but much greater than that during World War II.

It is tempting to speculate on the speed with which transformational [generative] grammar would have won general acceptance had Chomsky’s and Halle’s students had to contend with today’s [late 1970s] more austere conditions, in which not just military, but ALL sources of funding have been sharply curtailed, and the number of new positions has been declining yearly (Newmeyer 1980: 52n8; other references at Murray *loc. cit.*).

93. Noam was also invited to Columbia for the 1957–58 year, with expectations of a permanent position (Murray 1994: 246). It seems likely that Zellig’s connections in New York and at Columbia were instrumental in this. Zellig’s friend Seymour Melman was on the faculty from 1949 until his retirement in 2003, and he himself repaired to Columbia when he left Penn.

94. See (Gardner 1985: 28–29).

This was “striking...support for adherents of a supposedly persecuted perspective engaged in making a ‘scientific revolution’”(Murray 1994: 242). In the extraordinarily polemical climate of the 1960s, prevalent in society at large as well as being cultivated in linguistics, there was talk of overthrowing the ‘hegemony of taxonomic linguistics’. There were claims, baseless in fact, of being misunderstood and rejected by that supposed hegemony.⁹⁵ A course about structural linguistics at M.I.T. was called ‘The Bad Guys’.⁹⁶ In the event, the outcome manifestly was a hegemony of Generative Linguistics, and to suppose that this was not also the aim requires extraordinary credulity.

Chomsky was able to define linguistics as whatever he and his associates did. Success in this ‘definition of the situation’ was facilitated by the tenuousness of institutionalized linguistics and the paucity of neo-Bloomfieldians relative to the hordes of students. While there was a faculty-based core with claims to forming an elite specialty, it was the exponential increase in the number of linguists and M.I.T.’s military funding that diluted the strength of opposition by those trained earlier. (Murray 1994: 244)

The rhetoric of revolution does not fit the facts in crucial ways. Noam’s work and that of his followers was

...published, sought out and taken seriously by major neo-Bloomfieldians. The most central of them actively fostered Chomsky’s and Lees’s careers. In terms of aggression, the Chomskians struck first. Their revolutionary rhetoric was not a reaction to the incomprehension of the ‘establishment,’ nor a defense against neophobia or persecution by angry elders. (Murray 1994: 244)

On the other hand, the rhetoric of revolution was extremely effective marketing.⁹⁷

95. “The active contributors to linguistics in the period were themselves far from agreement on these points, and there is some reason to think that the program was being superseded, or at least significantly extended, by the beginning of the 1950s...” (Hymes & Fought 1975: 122) “It would be quite mistaken to regard the period as a fortress of fixed opinions. ... What would have been insisted upon was the necessity of the *kind* of work — development of explicit, abstract, wholly general models of the nature of linguistic structure.” (*ibid*: 126).

96. In correspondence with me, Noam has rejected any responsibility to counter misrepresentations that his students or followers have made of Zellig or his work. To some extent, and perhaps a great extent, he may be genuinely innocent of their excesses, in the sense that he has been swept along by social phenomena that he neither created nor really guided, but was able to ride willy-nilly. But neither has he done anything to discourage or correct those who look to him for an insider’s knowledge in these matters, and in that, certainly, my notions of responsibility differ from his.

97. It may be relevant that advertising and marketing texts were (perhaps pointedly) among those subjected to discourse analysis in seminars and discussions in which Noam was

Revolutionary rhetoric seems to appeal to the enthusiasms of a new generation (with no stake in the old perspective and little knowledge of it) for meaningful and original work. (Murray 1994: 245)

As Generative Linguistics achieved the perception in other fields and in the public as ‘mainstream linguistics’, the new departments that were being formed fell in line, eager to share in the funding, and pushed by administrators anxious to maintain the reputations of their institutions in the eyes of parents, students, and alumni. Though peer review was freely available — and extraordinarily forbearing, given the ill-founded viciousness of many attacks — Generativists sidestepped regular publication channels in favor of circulating mimeographed papers. New journals were founded — *Foundations of Language* (1965), *Papers in Linguistics* (1969), and especially *Linguistic Inquiry* (1970) at M.I.T. — whose closed editorial policies were in sharp contrast to the openness of the ‘establishment’ journals *Language* and *IJAL*.⁹⁸ Subsequently, despite the drying up of government largesse, and the consequent contraction or demise of many of the new Departments of Linguistics, these generativist benchmarks of establishment have all survived, except for one (*Papers in Linguistics*). While the status of Generative Linguistics as a scientific revolution is uncertain, there can be no doubt of its success as a revolution in academic politics.

Noam and his students had an open invitation to come down to Penn for collaboration. In the late 1980s, speaking with me of our student days twenty years earlier, Ellen Prince told me that the emissaries that were sent from M.I.T. were commissioned essentially to spy and take ideas back. “And Harris and Hiz *invited* them!” she exclaimed. “They were such *suckers!*” Recently, I connected this remark with certain sociological observations recounted on Ira Glass’s radio program, “This American Life”, about the Israeli concept of being a *freier*.⁹⁹ Among many striking examples illustrating freierism, there is an interview with Tom Segev, an Israeli journalist and historian, who says:

You constantly hear it, *constantly*: don’t be a *freier*. That is the worst thing for an Israeli to be, a *freier*, in his own eyes, and also in the eyes of other Israelis. So never ever be too generous, be always on guard. Somebody is out there to take what is due to you.

I think it would be impossible to understand Israelis without understanding the whole notion of freierism. It is at the heart of Israeli culture, affecting how

involved, as for example the “Millions Can’t be Wrong!” text analyzed in Section 2.2 of Harris (1952).

98. The sole exception in North America was *Word* under the editorship of Martinet.

99. The program “Episode 222: Suckers” may be heard at http://www.thislife.org/Radio_Episode.aspx?episode=222. There is more discussion at <http://www.balashon.com/2007/10/freier.html> and <http://www.wzo.org.il/en/resources/view.asp?id=2226> and elsewhere.

people work, how they shop, how they vote, how they think about themselves and the people around them.

From an Israeli point of view, Jews were suckers for 2000 years in exile, constantly being tricked and persecuted. The whole idea of Israel is to create a place where Jews were in control, where Jews would never again be freiers. And even though Israel is now a powerful state, the fear of being taken advantage of hasn't gone away.

Part of the 'revolutionary' conflict here may be that low-synergy¹⁰⁰ characteristics such as freierism became more prevalent during and after WWII, resulting in a generational difference in which Zellig and other elders of the field, as noted above, had the more forbearing nature of Sapir and Bloomfield, and Noam's already contentious temperament, joined with that of his new colleagues and students, was encouraged and amplified in an environment of "post-War absolutism" (Harris 1965: fn5).¹⁰¹ But while this sort of Zeitgeist effect may be part of the picture, it is no explanation. There were many students; there was only one Noam Chomsky.

For present purposes, however, these developments, by the same token, extend beyond the scope of this paper; for by the time the rising flood-tide of funding, the accretion of politically adept colleagues, and the peculiar institutional circumstances at M.I.T. came together in a kind of perfect storm, differences in temperament and predisposition had already set Noam on an increasingly divergent course from the family friend to whom his father had turned for help so many years before.

100. Borrowing the term from Ruth Benedict, as presented in Maslow & Honigmann 1970. Coming from her experience of culture contact with Native Americans and in the Pacific before and during the war, she sought means for determining which of two juxtaposed culture traits or complexes is to be preferred. Simplistically put, that culture is best which is best for those enacting it, a subtle and sophisticated relative of utilitarianism. In a culture or trait that is low in synergy, such as freierism, there is a strong conflict between selfishness and altruism, while in systems that are very high in synergy, one cannot benefit others without benefit to oneself, and the converse, so that the vocabulary and concepts for selfishness and altruism may not even exist in the language of such a people. Kindred ideas are often expressed in terms of zero-sum games, etc.

101. It may be objected that Zellig lived in Israel a lot of the time, but he lived on one of the dwindling number of left-idealistic kibbutzim. This is not to say that he was naively insulated from these corrosive developments. The comment in Harris (1965 fn. 5) indicates his awareness of this social change, and as well his detachment from it.

References

- Aslin, R.N. & E.L. Newport. 2009. "What statistical learning can and can't tell us about language acquisition". *Infant pathways to language : methods, models, and research directions*, ed. by John Colombo, Peggy D. McCardle & Lisa Freund. New York: Psychology Press.
- Bandler, Richard, John Grinder, Steve Andreas & Connirae Andreas. 1982. *Reframing : N L P and the transformation of meaning*. Moab, Utah: Real People Press.
- Barsky, Robert F. 1997. *Noam Chomsky : a life of dissent*. Toronto: ECW Press.
- Barsky, Robert F. 2007. *The Chomsky effect : a radical works beyond the ivory tower*. Cambridge, Mass.: M.I.T. Press.
- Barsky, Robert F. Forthcoming. "The Chomsky Effect: Episodes in Academic Activism (Chapter 6)". *Beyond the Ivory Tower: Public Intellectuals, Academia and the Media*, ed. by Saleem H. Ali & Robert F. Barsky. [ms. available at <http://www.mit.edu/~saleem/ivory/>]
- Bloch, Bernard. 1948. "A Set of Postulates for Phonemic Analysis". *Language: Journal of the Linguistic Society of America* 24: 1. 3–46.
- Bloomfield, Leonard. 1939. "Menomini Morphophonemics". *Travaux du Cercle Linguistique de Prague* 8. 105–115.
- Bruner, Jerome S., Jacqueline J. Goodnow & George A. Austin. 1962. *A study of thinking*. New York: Wiley.
- Bruner, Jerome S. & Rita Watson. 1985. *Child's talk : learning to use language*. Oxford: Oxford University Press.
- Butkus, Ben. 2009. "US Universities' Patent Policies Retard Plant Biology Research, Survey Suggests" [January 28, 2009]. *Genomeweb Biotech Transfer Week* <http://www.genomeweb.com/biotechtransferweek/us-universities-patent-policies-retard-plant-biology-research-survey-suggests>.
- Carnap, Rudolf. 1928. *Der logische Aufbau der Welt*. Berlin-Schlachtensee: Weltkreis-verlag.
- Carnap, Rudolf. 1934. *Logische Syntax der Sprache*. Wien: J. Springer.
- Chomsky, Noam. 1951. *Morphophonemics of Modern Hebrew*. [Unpublished masters Thesis catalogued as 378.748 PoA/1951. 60 (RBC) in the Rare Book and Manuscript Library in the van Pelt Library at the University of Pennsylvania.]
- Chomsky, Noam. 1955a. "The Logical Structure of Linguistic Theory". Ms., date June 1955. Available from Columbia University Psychology Library.
- Chomsky, Noam. 1955b. *Transformational Analysis*. Ph.D. dissertation, University of Pennsylvania.
- Chomsky, Noam. 1956a. *Logical Structure of Linguistic Theory* [Indicated to be a revision of Chomsky 1955a. Preserved on microfilm at the M.I.T. Humanities Library. PDF available for download at <http://alpha-leonis.lids.mit.edu/chomsky/>. Further revision published as Chomsky 1975.]
- Chomsky, Noam. 1956b. "Three Models for the Description of Language". *IRE(Institute of Radio Engineers) Transactions on Information Theory* 2. 113–124.
- Chomsky, Noam. 1959. "Review of Verbal Behaviour, by B.F. Skinner". *Language* 35: 1. 26–59.
- Chomsky, Noam. 1964. *Current issues in linguistic theory*. The Hague: Mouton.
- Chomsky, Noam. 1969. "Linguistics and Politics". *New Left Review* 57. 21–34.
- Chomsky, Noam. 1975. *The logical structure of linguistic theory*. New York: Plenum Press. [Revision of Chomsky 1956a.]
- Chomsky, Noam & Mitsou Ronat. 1977. *Dialogues avec Mitsou Ronat*. Paris: Flammarion. [English translation revised and expanded as Chomsky & Ronat 1979.]

- Chomsky, Noam. 1979a. "Morphophonemics of Modern Hebrew". *Outstanding Dissertations in Linguistics, a Garland Series*, ed. by J. Hankamer. [Substantial revision of Chomsky 1951.].
- Chomsky, Noam & Mitsou Ronat. 1979. *Language and responsibility : based on conversations with Mitsou Ronat*. New York: Pantheon Books. [Revised and expanded translation of Chomsky & Ronat 1977.]
- Church, Alonzo. 1956. *Introduction to mathematical logic*. Princeton: Princeton University Press.
- Church, Alonzo. 1956. *Introduction to mathematical logic*. Princeton: Princeton University Press.
- Cziko, Gary. 1995. *Without miracles : universal selection theory and the second Darwinian revolution*. Cambridge, Mass.: M.I.T. Press.
- Cziko, Gary. 2000. *The things we do : using the lessons of Bernard and Darwin to understand the what, how, and why of our behavior*. Cambridge, Mass.: M.I.T. Press.
- Dresher, B.E. 2003. "The contrastive hierarchy in phonology". *Toronto Working Papers in Linguistics (Special Issue on Contrast in Phonology)* 20. 47–62.
- Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. "The Evolution of the Language Faculty: Clarifications and Implications". *Cognition: International Journal of Cognitive Science* 97: 2. 179–210.
- Gebhart, A.L., E.L. Newport & R.N. Aslin. 2009. "Statistical learning of adjacent and non-adjacent dependencies among non-linguistic sounds". *Psychonomic Bulletin & Review* 16. 486–490.
- Gibson, James J. 1977. "The Theory of Affordances", ed. by Robert E. Shaw & John Bransford. Hillsdale, N.J.; New York: Lawrence Erlbaum Associates & Halsted Press Division, Wiley.
- Gibson, James Jerome. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gleitman, Lila. 2002. "Verbs of a feather flock together II: The child's discovery of words and their meanings". *The legacy of Zellig Harris : language and information into the 21st century*, ed. by Bruce E. Nevin, 209–229. Amsterdam: John Benjamins.
- Goldsmith, John A. 2004. "From algorithms to generative grammar and back again". *Proceedings from the 40th Annual Meeting of the Chicago Linguistic Society* 243–259. [Available at <http://cls.metapress.com/index/P8JP462N387V142V.pdf> and <http://humanities.uchicago.edu/faculty/goldsmith/Papers/CLS2004Algorithms.pdf>.]
- Goodman, Nelson. 1943. "On the Simplicity of Ideas". *Journal of Symbolic Logic* 8. 107–121.
- Gross, Maurice. 1979. "On the Failure of Generative Grammar". *Language: Journal of the Linguistic Society of America* 55: 4. 859–885.
- Halle, Morris. 1959. *The sound pattern of Russian : a linguistic and acoustical investigation*. The Hague: Mouton.
- Harris, Randy Allen. 1993. *The linguistics wars*. New York: Oxford University Press.
- Harris, Randy Allen. 1998. "Review of *Noam Chomsky: A Life of Dissent*". *Books in Canada*: March 1998. [Available at <http://www.arts.uwaterloo.ca/~raha/reviews/Harris-Barsky.pdf>].
- Harris, Zellig S. 1936. *A grammar of the Phoenician language*. New Haven, Conn.: American Oriental society.
- Harris, Zellig S. 1939. *Development of the Canaanite dialects; an investigation in linguistic history*. New Haven: American oriental society.
- Harris, Zellig S. 1940. "Review of *Foundations of Language*, by L.H. Gray". *Language* 16. 216–230.
- Harris, Zellig S. 1941a. "Linguistic Structure of Hebrew". *Journal of the American Oriental Society* 61: 3. 143–167.
- Harris, Zellig S. 1941b. "Review of N[ikolaj] S[ergeevich] Trubetzkoy (1890–1938), *Grundzüge der Phonologie* (Prague: Cercle Linguistique de Prague, 1939)". *Language* 17. 345–349.

- Harris, Zellig S. 1942. "Morpheme Alternants in Linguistic Analysis". *Language* 18: 3. 169–180.
- Harris, Zellig S. 1945. "Navaho Phonology and [Harry] Hoijer's Analysis". *IJAL* 11: 4. 239–246.
- Harris, Zellig S. 1946. "From Morpheme to Utterance". *Language* 22: 3. 161–183.
- Harris, Zellig S. 1947a. "Structural Restatements I: Swadesh's Eskimo; Newman's Yawelmani". *IJAL* 13: 1. 47–58.
- Harris, Zellig S. 1947b. "Structural Restatements II: Voegelin's Delaware". *IJAL* 13: 3. 175–186.
- Harris, Zellig S. 1948. "Componential Analysis of a Hebrew Paradigm". *Language* 24: 1. 87–91.
- Harris, Zellig Sabbetai. 1951a. *Methods in structural linguistics*. Chicago The Univ. of Chicago Press. (Repr. as "Phoenix Books" P 52 with the title *Structural Linguistics*, 1960; 7th impression, 1966; 1984.) [Preface signed "Philadelphia, January 1947"].
- Harris, Zellig S. 1951b. "Ha-Safah ha-Ivrit l'or ha-balshanut ha-chadashah ["The Hebrew language in the light of modern linguistics"]". *Lashenanu* 17. 128–132.
- Harris, Zellig S. 1952. "Discourse Analysis". *Language* 28: 1. 1–30.
- Harris, Zellig S. 1954. "Distributional Structure". *Word* 10: 2. 146–162.
- Harris, Zellig S. 1955. "From Phoneme to Morpheme". *Language* 31: 2. 190–222.
- Harris, Zellig S. 1957. "Co-Occurrence and Transformation in Linguistic Structure". *Language* 33: 3. 283–340.
- Harris, Zellig S. 1962a. *String analysis of sentence structure (= Papers on Formal Linguistics, 1)*. The Hague: Mouton.
- Harris, Zellig S. 1962b. "A Language for International Cooperation". *Preventing World War III, some proposals*, ed. by Quincy Wright. New York: Simon & Schuster.
- Harris, Zellig S. 1964. *The elementary transformations. (Transformations and Discourse Analysis Papers, 54.)* Philadelphia: University of Pennsylvania.
- Harris, Zellig S. 1965. "Transformational Theory". *Language* 41: 3. 363–401.
- Harris, Zellig S. 1966. "A Cycling-Cancellation Automation for Sentence Well-Formedness". *International Computation Centre Bulletin* 5. 69–94.
- Harris, Zellig S. 1967. *Morpheme boundaries within words: report on a computer test. (Transformations and Discourse Analysis Papers, 73.)* Philadelphia: University of Pennsylvania.
- Harris, Zellig S. 1968. *Mathematical structures of language*. New York: Wiley Interscience.
- Harris, Zellig S. 1969. *The two systems of grammar: report and paraphrase. (Transformations and Discourse Analysis Papers, 79.)* Philadelphia: University of Pennsylvania.
- Harris, Zellig S. 1970. *Papers in structural and transformational linguistics*. Dordrecht: Reidel.
- Harris, Zellig Sabbetai. 1981. *Papers on syntax*. Dordrecht: Reidel.
- Harris, Zellig S. 1982. *A grammar of English on mathematical principles*. New York: Wiley.
- Harris, Zellig S. 1988. *Language and information*. New York: Columbia Univ. Press.
- Harris, Zellig S. 1989. *The form of information in science : analysis of an immunology sublanguage*. Dordrecht/Holland & Boston: Kluwer Academic Publishers.
- Harris, Zellig S. 1990. "La genèse de l'analyse des transformations et de la métalangue" [Sept. 1990]. *Langages* 99. 9–19. [Tr. of Harris 2003]
- Harris, Zellig S. 1991. *A theory of language and information : a mathematical approach*. Oxford; New York: Clarendon Press; Oxford University Press.
- Harris, Zellig S. 2002. "The background of transformational and metalanguage analysis". *The legacy of Zellig Harris : language and information into the 21st century*, ed. by Bruce E. Nevin. Amsterdam: John Benjamins.
- Hawking, Stephen William. 1988. *A brief history of time : from the big bang to black holes*. Toronto: Bantam Books.
- Hiž, Henry. 1979. "On Some General Principles of Semantics of a Natural Language". *Syntax and Semantics* 10. 343–352.

- Hiž, Henry. n.d. Linguistics at the University of Pennsylvania: Internal document of the Linguistics Department at the University of Pennsylvania.
- Hockett, Charles F. 1947. "Peiping Phonology". *Journal of the American Oriental Society* 67: 4. 253–267.
- Huck, Geoffrey J. & John A. Goldsmith. 1995. *Ideology and linguistic theory : Noam Chomsky and the deep structure debates*. London ; New York: Routledge.
- Hymes, Dell H. 1971. *On communicative competence*. Philadelphia: University of Pennsylvania Press.
- Hymes, Dell & John G. Fought. 1975. *American structuralism*. The Hague: Mouton.
- Jackendoff, Ray & Steven Pinker. 2005. "The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, & Chomsky)". *Cognition* 97: 2005. 211–225.
- Joos, Martin. 1957. *Readings in linguistics 1 : the development of descriptive linguistics in America, 1925–56*. Chicago: University of Chicago Press.
- Joshi, Aravind K. 1959. "Recognition of local substrings". (= *Transformations and Discourse Analysis Papers*, 18.) University of Pennsylvania.
- Joshi, Aravind K. 1985. "Tree-adjointing grammars: How much context sensitivity is required to provide reasonable structural descriptions?" *Natural Language Parsing*, ed. by D. Dowty, L. Karttunen & A. Zwicky, 206–250: Cambridge University Press.
- Joshi, Aravind K. 2002. "Hierarchical structure and sentence description". *The legacy of Zellig Harris : language and information into the 21st century, Vol. 2: Computability of language and computer applications*, ed. by Bruce E. Nevin & Stephen B. Johnson. Amsterdam: John Benjamins.
- Joshi, Aravind K., S.R. Kosaraju & H.M. Yamada. 1968. String adjunct grammars (= *Transformations and Discourse Analysis Papers*, 75.) University of Pennsylvania.
- Joshi, Aravind K., S.R. Kosaraju & H.M. Yamada. 1972a. "String adjunct grammars: I. Local and distributed adjunction". *Information and Control* 21: 2. 93–116.
- Joshi, Aravind K., S.R. Kosaraju & H.M. Yamada. 1972b. "String adjunct grammars: II. Equational representation, null symbols, and linguistic relevance". *Information and Control* 21: 3. 235–260.
- Joshi, Aravind K. & Owen Rambow. 2003. "A Formalism for Dependency Grammar Based on Tree Adjoining Grammar". *International Conference on Meaning Text Theory (MTT) 2003*. Paris.
- Kirchner, Robert. 1995. "Contrastiveness is an epiphenomenon of constraint ranking". Paper presented at the 21st Annual Meeting of the Berkeley Linguistics Society, University of California, Berkeley, February 1995.
- Kleene, Stephen Cole. 1952. *Introduction to metamathematics*. New York: Van Nostrand.
- Koerner, E.F.K. 2002. *Toward a history of American linguistics*. London ; New York: Routledge.
- Koerner, E.F.K., Matsuji Tajima & Carlos Peregrín Otero. 1986. *Noam Chomsky : a personal bibliography, 1951–1986*. Amsterdam/Philadelphia: John Benjamins.
- Lakoff, George. 1987. *Women, fire, and dangerous things : what categories reveal about the mind*. Chicago: University of Chicago Press.
- Lakoff, George. 2002. *Moral politics : how liberals and conservatives think*. Chicago: University of Chicago Press.
- Lashley, Karl S. 1951. "The Problem of Serial Order in Behavior". *Cerebral mechanisms in behavior : the Hixon Symposium*, ed. by Lloyd Alexander Jeffress. New York: Wiley.
- Laufer, Asher & I.D. Conday. 1979. "The epiglottis as an articulator". *UCLA Working Papers in Phonetics* 45. 60–83.

- Lentin, André 2002. "Reflections on references to mathematics in the work of Zellig Harris". *The legacy of Zellig Harris : language and information into the 21st century, Vol. 2: Computability of language and computer applications*, ed. by Bruce E. Nevin & Stephen B. Johnson, 1–9. Amsterdam: John Benjamins.
- Lin, Francis Y. 1999. "Chomsky on the 'ordinary Language' View of Language". *Synthese* 120. 151–192.
- MacFarquhar, Larissa. 2003. "The Devil's Accountant" [March 31, 2003]. *New Yorker* 79: 6. 64–79.
- Manaster-Ramer, Alexis & Michael B. Kac. 1990. "The Concept of Phrase Structure". *Linguistics and Philosophy: An International Journal* 13: 3. 325–362.
- Marken, Richard. 2002. *More mind readings: methods and models in the study of purpose*. St. Louis, Mo.: Newview.
- Maslow, Abraham H. & John J. Honigmann. 1970. "Synergy: Some Notes of Ruth Benedict". *American Anthropologist, N. S.* 72: 2. 320–333.
- Murray, Stephen O. 1994. *Theory groups and the study of language in North America : a social history*. Amsterdam/Philadelphia: John Benjamins.
- Murray, Stephen O. 1999. "More on Gatekeepers and Noam Chomsky's Writings of the 1950s". *Historiographia Linguistica* 26: 3. 343–353.
- Nevin, Bruce E. 1984. "Review of Harris 1982: A Grammar of English on Mathematical Principles". *Computational Linguistics* 10: 3–4. 203–211.
- Nevin, Bruce E. 1989. "Unbounded Dependencies in Constructive Grammar". Ms.
- Nevin, Bruce E. 1993a. "Harris the revolutionary: Phonemic theory". *History of Linguistics 1993, Papers from the Sixth International Conference on the History of the Language Sciences (ICHoLS VI)*, ed. by Kurt R. Jankowsky. Amsterdam/Philadelphia: John Benjamins.
- Nevin, Bruce E. 1993b. "A Minimalist Program for Linguistics: The Work of Zellig Harris on Meaning and Information". *Historiographia Linguistica: International Journal for the History of the Language Sciences/Revue Internationale pour l'Histoire* 20: 2–3. 355–398.
- Nevin, Bruce E. 1998. *Aspects of Pit River Phonology*, Unpublished Ph.D. dissertation, University of Pennsylvania, Available at <http://repository.upenn.edu/dissertations/AAI9913504/>.
- Nevin, Bruce E. 1999[1993]. "Harris the revolutionary: Phonemic theory". *History of Linguistics 1993, Papers from the Sixth International Conference on the History of the Language Sciences (ICHoLS VI)*, ed. by Kurt R. Jankowsky. Amsterdam/Philadelphia: John Benjamins.
- Nevin, Bruce E., ed. 2002a *The Legacy of Zellig Harris: Language and Information into the 21st Century. Volume 1: Philosophy of Science, Syntax and Semantics*. Amsterdam: John Benjamins.
- Nevin Bruce E. 2002b. "Foreword." In Nevin 2002a: ix–xxxiv.
- Nevin, Bruce E. & Stephen B. Johnson. 2002. *The legacy of Zellig Harris : language and information into the 21st century, Vol. 2: Computability of language and computer applications*. Amsterdam: John Benjamins.
- Newmeyer, Frederick J. 1980. *Linguistics in America: The first quarter-century of transformational-generative grammar*. New York: Academic Press.
- Overbye, Dennis. 2009. "Elevating Science, Elevating Democracy" [January 26, 2009]. *New York Times* Available at http://www.nytimes.com/2009/01/27/science/27essa.html?_r=2.
- Pinker, Steven & Ray Jackendoff. 2005. "The faculty of language: what's special about it?" *Cognition* 95. 201–236.
- Post, Emil L. 1943. "Formal Reductions of the General Combinatorial Decision Problem". *American Journal of Mathematics* 65: 2. 197–215.

- Powers, William T. 1973. *Behavior: the control of perception*. Chicago: Aldine Pub. Co.
- Powers, William T. 1998. *Making sense of behavior : the meaning of control*. New Canaan, Conn.: Benchmark Publications.
- Powers, William T., ed. 2008. *Perceptual Control Theory: Science & Applications — A Book of Readings*. Hayward, CA: Living Control Systems Publishing.
- Reeder, P.A., E.L. Newport & R.N. Aslin. In press (2009). "The role of distributional information in linguistic category formation". *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*, ed. by N. Taatgen, H. van Rijn, L. Schomaker & J. Nerbonne.
- Ross, John R. 1967. *Constraints on Variables in Syntax*, Ph.D. Dissertation, M.I.T.
- Runkel, Philip Julian. 1990. *Casting nets and testing specimens : two grand methods of psychology*. New York: Praeger.
- Runkel, Philip Julian. 2003. *People as living things : the psychology of perceptual control*. Hayward, CA: Living Control Systems.
- Russell, Bertrand [Arthur William]. 1905. "On Denoting" [Oct. 1905]. *Mind, New Series* 14: 56. 479–493.
- Ryckman, Thomas A. 1986. *Grammar and Information: An investigation in linguistic metatheory*. Unpublished Ph.D. dissertation, Columbia University.
- Sager, Naomi. 1984. *Natural language information processing : a computer grammar of English and its applications*. Reading, Mass.: Addison-Wesley.
- Sager, Naomi & Ngô Thanh Nhân. 2002. "The computability of strings, transformations, and sublanguage". *The legacy of Zellig Harris : language and information into the 21st century, Vol. 2: Computability of language and computer applications*, ed. by Bruce E. Nevin & Stephen B. Johnson, 79–120. Amsterdam: John Benjamins.
- Seuren, Pieter A.M. 2009. "Concerning the Roots of Transformational Generative Grammar". *Historiographia Linguistica* 36: 1. 97–115.
- Skinner, B.F. 1957. *Verbal behavior*. New York: Appleton-Century-Crofts.
- Spencer-Brown, G. 1969. *Laws of form*. London: Allen & Unwin.
- Stevens, Kenneth N. 1972. "The quantal nature of speech: Evidence from articulatory-acoustic data". *Human communication: a unified view*, ed. by Edward E. David & Peter B. Denes, 51–66. New York: McGraw-Hill.
- Tarski, Alfred. 1933. *Der Wahrheitsbegriff in den formalisierten Sprachen [Tr. as The Concept of Truth in Formalized Languages]*. Indianapolis, Ind.: 1983, Hackett Pub. Co.
- Thomas, Margaret. 2002. "Roger Bacon and Martin Joos: Generative Linguistics' reading of the past." *Historiographia Linguistica* 39(3). 339–378.
- Trubetzkoy, Nikolaj Sergeevi. 1939. "Grundzüge der Phonologie". *Travaux du Cercle Linguistique de Prague* 7.
- Weintraub, Pamela. 2008. *Cure unknown : inside the Lyme epidemic*. New York: St. Martin's Press.
- Wells, Rulon S. 1947. "Immediate Constituents". *Language: Journal of the Linguistic Society of America* 23: 2. 81–117.
- Yates, Frances Amelia. 1972. *The Rosicrucian enlightenment*. London; Boston: Routledge & Kegan Paul.