

REVIEW ARTICLE

The legacy of Zellig Harris: *Language and information into the 21st century*, vol. 1: Philosophy of science, syntax and semantics. Ed. by BRUCE NEVIN. Philadelphia: John Benjamins, 2002. ISBN 1588112462. \$150 (Hb).

Reviewed by JOHN GOLDSMITH, *University of Chicago**

1. INTRODUCTION. Zellig Harris (1909–1992) cast a long shadow across twentieth century linguistics. In mid-century, he was a leading figure in American linguistics, serving as president of the Linguistic Society of America in 1955, just a year before Roman Jakobson. It is fair to say that during that decade—the years just before generative grammar came on the scene—Zellig Harris and Charles Hockett were the two leading figures in the development of American linguistic theory. Today, I daresay Harris is remembered by most linguists as the mentor and advisor to Noam Chomsky at the University of Pennsylvania—and the originator of transformational analysis.¹

But Harris was an extraordinarily deep thinker about linguistic theory, and he made important contributions to many fields of linguistics. Though I never met him myself, I have often felt that a great deal of my own work was exploration of territory where he had already been and had left signs for later researchers—signs that noted where important problems were to be found and how they might best be treated. And three of my teachers worked closely with Harris (Lila Gleitman, Haj Ross, and Noam Chomsky) and were greatly influenced by him. Truth in advertising: Zellig Harris is a bit like a grandfather I never met.

Bruce Nevin, a student of Harris's, has now produced a two-volume tribute to Harris bringing together work by a range of researchers in linguistics and the other disciplines where Harris did significant work. All of the areas of Harris's linguistic work are covered in these volumes, and a careful reading of them leaves this reader with the conclusion that there is no way to understand American linguistic theory through the second half of the twentieth century without understanding Harris's thought. I attempt here to explain why this is so in this discussion of volume 1; volume 2 focuses on computational issues.² As Nevin notes, what he produced is neither a festschrift nor a memorial volume, and as my purpose here is to better understand Harris's role in the development of twentieth century linguistics and the relevance of his thought for linguists today, I have little to say about some of the contributions; contributors were apparently invited to this project and asked to offer a paper that clearly represented 'some relationship to Harris's work', but this relationship is not always evident. A synoptic paper by Harris himself is included at the beginning of the collection.

Harris's work must be situated in terms of the conflict between two visions of linguistic science: the MEDIATIONALIST view, which sees the goal of linguistic research as the discovery of the way in which natural languages link form and meaning, and the DISTRIBUTIONALIST view, which sees the goal as the fully explicit rendering of how the individual pieces of language (phoneme, syllable, morpheme, word, construction, etc.)

* I am grateful to Lila Gleitman, Bruce Nevin, Brian Joseph, Noam Chomsky, and Morris Halle for comments on a preliminary draft, which is not to say that they endorse in any fashion what I have written, but it is a better paper for their frank comments. I am also grateful to Jennifer Parham for useful improvements.

¹ See Harris's obituary offered by Matthews (1999).

² I am preparing a book notice on volume 2 that will appear in a future issue of *Language*.

connect to one another in the ways that define each individual language.³ The mediation-
 alist view lurks behind most conceptions of language study, formal and nonformal, but
 it was Harris's view that each successive improvement in linguistic theory took us a
 step further AWAY from the mediationalist view, much as advances in biology led
 scientists to understand that the study of living cells required no new forms of energy,
 structure, or organization in addition to those which were required to understand nonliv-
 ing matter. Harris had no use for mediationalist conceptions of linguistics. For linguists
 in 2004, steeped as we are in an atmosphere of linguistic mediationalists, this makes
 Harris quite difficult to understand at first.

Harris's goal was to show that all that was worthwhile in linguistic analysis could
 best be understood in terms of distribution of components at different hierarchical
 levels, because he understood—or at least he believed—that there was no other basis
 on which to establish a coherent and general linguistic theory. His genius lay in the
 construction of a conception of how such a vision could be put into place concretely.

Harris's view, from his earliest work through his final statements in the early 1990s,
 was that the best foundational chances for linguistics were to be found in establishing
 a science of EXTERNAL LINGUISTIC FACTS (such as corpora, though they would typically
 be augmented by other external facts, like speaker judgments), rather than a science
 of internalized speaker knowledge. Today, in mainstream linguistics, that decision is
 generally described as an EMPIRICIST leaning on Harris's part, in contrast to an alternative
 RATIONALIST leaning that lay at the heart of Chomsky's conception of generative gram-
 mar by the late 1950s. Harris's linguistics was a science of English, or of Hidatsa or
 of Hebrew, insofar as it was manifested in various actual linguistic creations in those
 languages, while Chomsky's linguistics was, or is, a science of the capacities of speakers
 who could utter or create sentences in those languages.

I believe that it is very important to understand the challenge that is implicit in
 developing these two perspectives. As Morris Halle has noted (p.c.), it was a growing
 consciousness of a commitment to a rationalist linguistics (in the sense just noted), and
 its difference from Harris's empiricist view, which led to the intellectual estrangement
 between Harris and Chomsky, in marked contrast to their close relationship up until
 around 1960.⁴ There are important subfields of linguistics that remain thoroughly com-
 mitted to an empiricist view of linguistics: all corpus linguistics, to be sure, is based
 on such a view,⁵ and most contemporary computational linguistics is concerned with
 solving various problems that arise out of dealing with English (Hidatsa, Somali, etc.)
 in the real world, rather than with providing a model of how human minds or brains
 might deal with such problems, and arguably many other subfields of linguistics hold
 to such a perspective, including historical linguistics and descriptive linguistics, among
 others.

³ These ideas are developed at length in Huck & Goldsmith 1996.

⁴ There is very little in this volume that explicitly compares generative grammar and Harrisian linguistics. Nevin, in his introduction, expresses regret that Noam Chomsky declined to contribute to the volume (xxvi). Noam Chomsky (p.c.) emphasizes that the growing consciousness of a conceptual difference (in my terms) was entirely on his side, in that he is 'sure that Harris never looked at my 1949, 1951 work on generative grammar', and that 'it's next to inconceivable, for example, that Harris looked at my Ph.D. dissertation or LSLT', and that Chomsky and Harris did not discuss this material during the 1950s. In nonlinguistic areas, their close relationship 'from the '40s continued without change . . . until the late 1960s, and ended for the usual reasons. There was no break of any kind'.

⁵ For a strong statement of this perspective, see Sampson 2003.

I myself am of the opinion that an updated version of Harris's conception does indeed provide the working linguist with a solid theoretical foundation for linguistics, arguably a more solid foundation than that provided by the current rationalist view, and I aim to fill in a few of the details on that score in this review.⁶ Harris's picture is one that provides linguistics, for better or worse, with a nearly radical insulation from interaction with psychology and biology, and such a move is one that will be (at least at first glance) anathema to many leading theoreticians in linguistics, ranging across a spectrum from, say Ronald Langacker and George Lakoff to Ray Jackendoff and, of course, Noam Chomsky—all of whom feel that the study of linguistic structures should meet up with the study of human cognitive structures before too long (and may already have done so).

2. HARRIS'S BIG IDEA. Harris did not appear to make a great effort to make his conclusions easily accessible to the reader. And yet once his ideas are understood, it is hard to deny that his way of stating them is direct, elegant, and striking. Let us approach the central idea of all of Harris's work, as summarized by Harris himself in his introductory paper to this volume:

[T]he structure of language can be found only from the non-equiprobability of combination of parts. This means that the description of a language is the description of contributory departures from equiprobability, and the least statement of such contributions (constraints) that is adequate to describe the sentences and discourses of the language is the most revealing. (9)

Picking this apart into pieces:

1. Linguistic analysis consists of building a representation out of a finite number of formal objects.
2. The essence of any given language is the restrictions, or constraints, that it places on how the pieces may be put together—these may be phonemes, morphemes, constituents, what have you. If there were no structure, then pieces could be put together any which way; structure MEANS—it is nothing more or less than—restrictions on how pieces can be put together.
3. These restrictions may be absolute ('no ⟨pk⟩ clusters are permitted in this language') or, much more likely, they are statements of distribution, best expressed in the mathematics of probability. A crude reformulation of this would be in the language of markedness, which is arguably an informal way of talking about

⁶ To offer a concrete formulation, consider Calabrese's (2004) formulation of realism:

This realistic view of language is based on two indisputable facts: given an utterance, 1) there must be a long term representation of the elements intervening in it; 2) there is an articulatory representation of it before actual muscular implementation. Phonology investigates the system of knowledge that allows the concrete occurrence of the real time computational steps that convert the mnemonic representation of the utterance into the articulatory representation of it. This knowledge involves representations and computations that have a concrete spatio-temporal occurrence and allows the production of concrete articulatory events, and are part of the workings of an actual brain with all its limitations.

Needless to say, perhaps, someone who did not agree with this position is not likely to want to call his position 'nonrealist'; but regardless of what names we apply to the positions, a person who, like Harris, felt that phonology need not view articulatory representations as the foundations of phonological theory might reply that it might be an 'indisputable fact' that basic research in particle physics is carried out only with massive financial support from a central government, but that does not tell us where the foundations of the scientific research lie, or that it might be an 'indisputable fact' that linguistic utterances in the real world are speech acts of one sort or another, but that does not guarantee (far from it!) that speech act theory will be critical to the development of a theory of phonology.

135 distributional frequencies. A better way is to use the mathematical vocabulary of
136 distributions, which is to say, probability theory.

137 4. A formal system can be described formally in a multitude of ways. These are not
138 equivalent: there is a priority among them based on their formal length. In general,
139 one will be significantly shorter than the others, and knowing its length is impor-
140 tant.

141 It is probably impossible to understand the intellectual pull of this research program
142 if one does not appreciate the revolutionary character (and the perceived success) of
143 the phoneme. If the phoneme today seems passé, the discarded error of an earlier
144 generation, then today's linguist should think of its descendant—for most of us, the
145 idea of an underlying segment; I return to this in a moment. By the mid twentieth
146 century, linguistics was arguably the most successful of the social sciences in dealing
147 with the major challenge faced by anyone trying to bring the words 'social' and 'science'
148 together: how to recognize and acknowledge the fact of human choice and action in
149 the social world (we say what we want to, not what we are obliged to) and at the same
150 time to apply a rigorous methodology that, at the end of the day, provides a revealing
151 analysis. Phonemic analysis of natural language established once and for all that utter-
152 ances in all the world's languages were not (to paraphrase Elbert Hubbard) just se-
153 quences of one damn sound after another, but were rather predictable realizations of
154 a language's choice of phonemes, a limited inventory chosen by a language with an
155 eye to symmetry and perhaps even a concern for its history—for one thing that was
156 certain was that the history of a language was not just a change of pronunciations, but
157 a change in the organized structure of the sound system of the language, as manifested
158 in its inventory of phonemes.⁷

159 Today we are likely to think that classical phonemic analysis focused on finding a
160 way to carve sound space into a language-particular set of boundaries; a vowel on one
161 side of a particular boundary in English is an /e/, and a vowel on the other side is an
162 /ɛ/. Learning the phonemes of a language (if we accept that view) amounts to learning
163 a language-particular sense of the distance between sounds (or phones): two phones
164 not separated by any phoneme boundaries in a language are perceived by speakers of
165 the language as identical, or nearly so, while sounds separated by a phonemic boundary
166 are perceived as quite distant (hence, distinct). It is easy to think that this is the classical
167 phonemic analysis, because this is what many phonemicists said, in so many words.

168 But in some respects, this view does not do justice to the notion of the phoneme,
169 nor to phonology. Some sound-relationships are more complex than what we can get
170 from establishing boundaries among the sounds of a language. Consider the flap of
171 American English: it is systematically related to the released voiceless [t] and to the
172 unreleased glottalized [tʔ], but the relationship is complex. Within words, the flap
173 appears obligatorily when a stressed vowel precedes (with only the possibility of an
174 /t/ interceding) and an unstressed vowel follows, and optionally when surrounded by
175 unstressed vowels; when word-final, a flap appears optionally when the next word
176 begins with a vowel, regardless of stress level. Now the idea that THIS kind of depth
177 and detail could govern the behavior of normal human beings in their everyday life

840 ⁷ One referee suggested that the formulation I offer anthropomorphizes language: it is PEOPLE who choose,
841 after all, not LANGUAGES. Yet we do say of languages that they borrow and lend words, they offer possibilities
842 to their speakers, they distinguish between sounds and between categories, they drop final vowels, and so
843 on; and we do ascribe voluntary acts to groups as well as to individuals (an individual VOTES FOR a candidate,
844 but only the entire electorate ELECTS a candidate).

178
179
180
181
182
183
184
185
186
187

was radical, and phonemics offered a method for discovering these relationships and for expressing the results in a simple format: there is a phoneme /t/, and its allophones are such-and-such, and each appears (obligatorily or optionally) in such-and-such contexts.

And if we no longer call this level of analysis PHONEMICS—if we call it an analysis of the underlying inventory of phonological segments, or something else—we have not lost sight of the crucial insight. Harris had certainly not lost sight of the crucial insight; his genius (or one of them) was to understand what the methodology was that lay behind phonemic analysis, to lay it bare, and to extend it to other aspects of human language.

188
189
190
191
192
193
194
195
196
197
198

It was his view that the important relationship between sounds lay not in their phonetics, but in their DISTRIBUTION, along the lines sketched two paragraphs above for the flap. What tells us that the flap and the other *t*'s of English are realizations of a single phoneme /t/ is not the similarity of sound, but the complementarity and predictability of the distribution. How can we extend this style of analysis to morphology, to syntax, and beyond? THAT was Harris's program. This point is clearly noted by TOM RYCKMAN ('Method and theory in Harris's grammar of information', 19–37): the central role played by the phoneme (25), and its distributional definition, in the development of Harris's method and general understanding of linguistics; he also notes that in Harris's work 'grammatical transformations are developed as a kind of "extended morphophonemics"', more powerful regularizing methods that enable even the derivation of tense and affixes' (32).

199
200

Harris used the opportunity of a review of Trubetzkoy in 1941 to articulate his methodological disdain for phonetics in a well-known passage (Harris 1941:346):

201
202
203
204
205
206
207
208
209
210
211

It is pointless to mix phonetic and distributional contrasts. If phonemes which are phonetically similar are also similar in their distribution, that is a result which must be independently proved. For the crux of the matter is that phonetic and distributional contrasts are methodologically different, and that only distributional contrasts are relevant while phonetic contrasts are irrelevant.

This becomes clear as soon as we consider what is the scientific operation of working out the phonemic pattern. For phonemes are in the first instance determined on the basis of distribution. Two positional variants may be considered one phoneme if they are in complementary distribution; never otherwise. In identical environment (distribution) two sounds are assigned to two phonemes if their difference distinguishes one morpheme from another; in complementary distribution this test cannot be applied. . . . [T]he distributional analysis is simply the unfolding of the criterion used for the original classification. If it yields a patterned arrangement of phonemes, that is an interesting result for linguistic structure.

212
213
214
215
216
217
218

From a broader perspective, there is an irony in choosing Trubetzkoy as the butt of this criticism: it was Trubetzkoy who, in his *Grundzüge der Phonologie* (1939), famously quoted Jakobson as saying that phonetics was to phonology as numismatics was to economics (in the context of a long discussion of the evolution of linguistic thought advancing towards a distinction between these sciences: an evolution which has continued FAR more in North America than it has in Europe, interestingly). But the review was an opportunity for Harris to express his own point of view.

219
220
221
222
223
224
225
226

3. DISCOVERY PROCEDURES. All linguists are aware today that a major difference between Chomsky's generative grammar and Harris's earlier distribution-based grammar lay in the role they assigned to what Chomsky (but not Harris) called DISCOVERY PROCEDURES, a term coined by Chomsky to contrast with EVALUATION PROCEDURES, a notion central to generative grammar (up until Chomsky 1981, when it was replaced by principles and parameters). Understanding this difference is both difficult and important.

The distinction transparently alludes to an issue in the philosophy of science that was raging in the 1950s, one that had its roots in the 1930s, and which has not yet

died down. In the philosophy of science, it referred to the distinction between the CONTEXT OF DISCOVERY and the CONTEXT OF JUSTIFICATION, terms coined by Hans Reichenbach (1938). The context of discovery of a scientific theory concerns history and personal psychology; the context of justification concerns logic and the relationship between a scientific statement and observation statements about events in space and time. A legitimate positivist should not confuse the two, and ultimately the scientist-positivist cares only about the context of justification, not that of discovery.

What is the relationship of linguists' work to their data? Are the data merely entries in their diary that provide some understanding of how they got to their eventual insight, or are they the stuff on which the theory stands or falls? This is the question an exegesis of Harris's work must answer. The rationalist today understands the linguist's data as being entries in the linguist's diary, of interest to historians of linguistics; the real subject of the work is the theory and the models. The empiricist sees the data which is collected by the larger community of linguistic researchers as an integral part of the work of the scientific community; no data, no science. It's not that it's simply unlikely the scientist will stumble upon the right theory without looking carefully at the data; there is no right theory to speak of except insofar as theory is united with data.

If we accept that perspective, then we can better understand what Harris's procedures were for and were about. They were proposed as methods which were systematic, which could be defined so as to be used in identical ways by different analysts, and which would provide insights into a wide range of human languages in ways that other methods would not. Developing such a method was understood as the primary goal of the scientific methodologist.⁸

In 'On discovery procedures' (69–86), FRANCIS Y. LIN focuses on the status of discovery procedures in Harris's work. He offers the following as a typical example:

[W]e take a form A in an environment C—D and then substitute another form B in the place of A. If, after such substitution, we still have an expression which occurs in the language concerned, i.e., if not only CAD but also CBD occurs, we say that A and B are members of the same substitution-class, or that both A and B fill the position C—D, or the like. (Harris 1946:102, cited by Lin, p. 72)

Lin argues that the entire panoply of discovery procedures, such as those Harris presented, should be understood as a scientific theory of how language is learned by the child. Lin himself feels that we are astride a dilemma: the discovery procedures can be conceived of either as a set of procedures that a linguist (either in theory or in practice) is to follow, or else as a set of procedures that the child language-learner is to follow. But Harris, says Lin, could not have meant the procedures to apply to the linguist: the linguist knows the phonemes, the morphemes, the transformations of the language before even starting the procedures, after all, and 'no serious scientists would ever think of such an idea' (80) as that of constructing an algorithmic procedure to analyze a set of data which the scientist him/herself is capable of analyzing by hand and for which no known algorithm that can replace the scientist exists at present; in this, it appears he is accepting, and following, Chomsky's suggestion, which Lin calls the 'oddity' argument, that 'there are few areas of science in which one would seriously consider the possibility of developing a general, practical, mechanical method for choosing among several theories, each compatible with available data' (81).

I'm afraid Lin's interpretation is unconvincing. There is no place in Harris's writing which Lin can cite that supports this view, for starters. More to the point, this interpreta-

⁸ Ryckman notes in his contribution (19–37): 'Once in conversation, when I had referred to him as a "linguist", Harris demurred, disclaiming any title as linguist, and said he thought of himself as a "methodologist".' (22)

273
274
275
276
277
278

tion is the Chomskyan interpretation: it is Chomsky who by the mid 1960s was pushing linguistic theory as a theory of language acquisition by human children. Harris had every opportunity to make explicit that this had been his intention too—but it had not been, as I read Harris (which is to say, conservatively). Just as importantly, Harris, I think, would never have agreed to let linguistics become just a branch of psychology, which is to say, a cognitive science.⁹

279
280
281
282
283
284
285
286
287
288
289

Now this, I am sure, may strike many a reader as quite, quite odd. There is a belief widely held among linguists along the following lines: to be a science, linguistics must be ABOUT SOMETHING, something in the real world; the only thing in the world it can be about is human brains; therefore, linguistics is a science that deals with biological (or psychological) facts about the brain or it is no science at all. Harris, I believe, would have rejected this logic, probably on the grounds that the scientific status of linguistics derives from its adherence to scientific METHOD, not to whether there was a physical THING, one that we could point to, that was what linguistics was ABOUT. It was the development of a METHOD that was Zellig Harris's lifework. Harris's framework of procedures, those called discovery procedures by Chomsky and others, was the heart of the method he was developing.

290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306

4. THE AUTONOMY OF LINGUISTICS. This view of Harris was one facet of a radically autonomous conception of linguistics. If Harris was anticognitivist (less anachronistically, antipsychologist) and antilogicist (for he rejected all explanation of language on semantic grounds) and antiphoneticist (as noted above), it was, at least at one level, because he was bound and determined to see linguistics become and remain an autonomous science, one whose conclusions were not dependent upon methods and conclusions of any other discipline, whether that discipline be psychology, logic, acoustics, or biology. It seems to me—and I cannot underscore this enough—that Harris took a clear and defensible position on this, one of the most important foundational questions in our field, and we have been rather lax in coming to grips with these questions, by and large. The challenges come up time and again: in the late 1960s, psychologists took seriously generative grammar's claim to be a psychological model, and then they were disappointed, and perhaps even shocked, to find that linguists were not interested in changing their minds about linguistic analysis when psychological evidence could be brought to bear on selecting grammatical models. Linguists continued to act as if their linguistic methods were independent of psychologists' criticisms, just as Harris would tell them to.

307
308
309
310
311
312
313
314
315

Harris adopted what seems to me to be the most honest and perhaps the most legitimate perspective on the relationship of linguistics to its neighboring fields: there is an interface between linguistics proper and such neighboring fields as phonetics, semantics, and psychology, and where there is disagreement, the differences have to be hashed out as best they can without a prior guarantee that both sides will come to an amicable agreement. When (distributional) linguistics tells us something about the sound structure of a language, such as that [ŋ] does not appear after silence in English (for this example, see Harris 1951:21), we may observe that 'English speakers will in general find difficulty in pronouncing them'. But 'all such predictions are outside the techniques and

848
849
850
851
852

⁹ Space does not permit a discussion of the history of psychologism and its rejection by logicians and philosophers since the late nineteenth century. Psychologism is the drive to reduce various fields of study (e.g. mathematics, linguistics, logic) to the study of how the human mind deals with the objects of those fields (numbers, sentences, deductions); it has waxed and waned several times in the last 150 years. A very brief remark along these lines can be found in Shaumyan 1996.

316 scope of descriptive linguistics. Linguistics offers no way of quantifying them. Never-
 317 theless, taking the linguistic representation as a clear and systemic model of selected
 318 features of speech, we may find that this model correlates with other observations about
 319 the people who do the speaking' (ibid.). Harris then directs the reader to Edward Sapir's
 320 classic observations about speaker intuitions in his 'La réalité psychologique des pho-
 321 nèmes' (1933). It is no accident, I would suggest, that this is precisely the same perspec-
 322 tive that Chomsky took on the relationship of semantics to syntax during the 1960s, a
 323 subject discussed in detail in Huck & Goldsmith 1996: meaning is not part of the
 324 business of syntax or grammar, but a well carried-out program of grammatical analysis
 325 is likely to have much to offer to neighboring disciplines that care about meaning and
 326 logical structure.¹⁰

327 ROBERT LONGACRE, in his contribution, 'Some implications of Zellig Harris's dis-
 328 course analysis' (117–35), remarks on Harris's unwillingness to see distribution-based
 329 analysis of discourse as yielding 'information as to its meaning', and one gets the
 330 impression that Longacre finds him being either overly modest or perverse (or perhaps
 331 ornery). Longacre rebuffs him in a thoroughly friendly way, saying, 'Surely [if we
 332 follow Harris's procedure] we may conclude that the result of [discourse analysis] is
 333 insight into the MEANING of the articles so analyzed' (119). I am sure that if Harris
 334 were to agree to that statement, it would be with the proviso that insights are outside
 335 of the scientific content of the theory proper. Longacre then provides a useful compari-
 336 son of Harris's analysis of a modern fable by James Thurber and his own analysis.

337 Tom Ryckman, in his contribution to this volume cited above, discusses the philo-
 338 sophical context in which Harris embraces grammar and rejects logic: as Ryckman
 339 puts it, the 'entry point . . . is the methodological postulate that there is no standpoint
 340 outside the data of language from which to advance theoretical inquiry' (21)—and this
 341 statement is paraphrased as the title of MAURICE GROSS's brief contribution, 'Conse-
 342 quences of the metalanguage being included in the language' (57–67).

343 Ryckman offers a rather long description of a parallel between the challenge of
 344 finding a foundation for logic, on the one hand, and finding a foundation for grammar,
 345 on the other. In the case of logic, once understood as the analysis of sound inference
 346 in thought, all reference to psychological foundations has been abandoned since the
 347 discoveries of the early twentieth century in favor of a purely formal account of how
 348 logic operates. How does this evolution bear on linguistics?¹¹ Ryckman offers some
 349 possibilities: a behaviorist, Leonard Bloomfield is antimentalist, hence uninterested in
 350 an explanation founded on logic; Franz Boas and Sapir were suspicious of categories
 351 historically derived from Indo-European languages. And Harris? What about Harris?

853 ¹⁰ Noam Chomsky (p. c.) notes that I misunderstand his position in saying this; cf. Huck & Goldsmith
 854 1996 for my reasons for suggesting this.

855 ¹¹ In Bar-Hillel's article (1964), he suggests that Rudolf Carnap and Leonard Bloomfield were doing much
 856 the same thing, in trying to do their job without fuzzy semantics: 'It is an interesting fact, deserving the
 857 attention of sociologists of science, that at approximately the same time, but in complete independence of
 858 each other, Bloomfield and Carnap were fighting the psychologism that dominated their respective fields,
 859 linguistics and logic. They both deplored the mentalistic mud into which the study of meaning had fallen,
 860 and tried to reconstruct their fields on a purely formal-structural basis. I think it is correct to say that the
 861 difference between the structural linguist and the formal logician is one of stress and degree rather than of
 862 kind' (Bar-Hillel 1964:43). Complete independence seems to this writer unlikely: Carnap and Bloomfield
 863 were, after all, colleagues in the Humanities Division at the University of Chicago from 1936 (when Carnap
 864 arrived) to 1940 (when Bloomfield left for Yale); Bar-Hillel later visited Carnap in Chicago when Bloomfield
 865 was no longer there.

352 Ryckman suggests that Harris cared more about patterns in language than in translation
 353 of language into another language, a logical language, and this is certainly true. But
 354 let's not forget that all academic conversation is dialogue between opposing points of
 355 view. I think Ryckman's point can be strengthened by taking note of the context in
 356 which Harris was working in the 1940s and 1950s.

357 Readers of *Language* fifty years ago were offered an explanation of this context by
 358 Yehoshua Bar-Hillel (1954), who began by pointing out to linguists that Rudolf Carnap
 359 (whom Bar-Hillel regarded as 'one of the greatest philosophers of all time' (1964:4))
 360 had twenty years earlier, in *The logical syntax of language* (published in 1934 under
 361 the title *Logische Syntax der Sprache*), argued for a coming together of formal syntax
 362 and formal logic: by formal, he meant analysis ignoring meaning and considering only
 363 categories and combinations of symbols; by syntax, the rules by which items are com-
 364 bined to form expressions (sentences); and by logic, the rules by which valid inferences
 365 from one sentence to another can be made. The contrast between syntax and logic
 366 was dubbed by Carnap (in English) as the difference between FORMATION rules and
 367 TRANSFORMATION rules, an interesting terminological suggestion and one that may have
 368 later influenced Harris. Bar-Hillel's paper starts—quite literally from the very first
 369 sentence—as an attack on Harris's claims about the sufficiency of distributional meth-
 370 ods for discovering the rules of both formation rules and transformation rules—this
 371 from someone who had once been quite sympathetic to Harris's project. Regardless of
 372 whether we are convinced or not by Bar-Hillel's arguments today, Harris clearly was
 373 not. Bar-Hillel thought that the innovations in logic brought about in the early 1930s
 374 by the Warsaw-Lwów school (Tadeusz Kotarbinski, Alfred Tarski, Kazimierz Adjukie-
 375 wicz) were sufficient to dispel Bloomfieldian doubts about the dangers inherent in
 376 basing grammatical analysis on formal logic; Harris did NOT agree. PAUL MATICK, in
 377 his contribution, 'Some implications of Zellig Harris's work for the philosophy of
 378 science' (39–55), notes that the studies that had given rise to Bar-Hillel's optimism
 379 were quickly subject to withering criticism, and that Harris did in later years allude to
 380 his reasons for not taking the Carnapian approach seriously: 'Not a few of the difficulties
 381 in the philosophy of language and in neighboring areas of philosophy arise from starting
 382 with the equipment which had been developed for truth systems and using it to analyze
 383 the information system that language represents' (Harris 1976; cited on p. 43). Mattick,
 384 in another context, summarizes one of the conclusions of Harris's later work on scien-
 385 tific language and notes that 'although logical structures are to be identified in natural
 386 language discourse, as in all cognitive activity, natural language is a much richer system
 387 of representation than logical calculi' (51).

388 Returning to Ryckman's development of Harris's overall program, Ryckman notes
 389 that there is a natural connection between the emphasis in the late 1940s and 1950s
 390 on information theory, on the one hand, and Harris's central view of language as a
 391 system of many levels, on the other, in which items at each level are combined according
 392 to their local principles of combination, and in which relatively simple cross-level
 393 principles could be established. What Ryckman does not say, but what is very clear
 394 from the quotation from Harris cited near the beginning of this review, is that Harris's
 395 method naturally leads one to a probabilistic grammar: the description of a language
 396 is the description of contributory departures from equiprobability, says Harris; hence,
 397 with a bit of mathematical work, to be sure, any utterance can be assigned a probability,
 398 and it can thus be assigned an information content.¹² Ryckman then goes on to discuss

866 ¹² This is accomplished through the basic definition that information content is the base 2 logarithm of
 867 the reciprocal of the probability of the message; this is a well-known expression in information theory.

Harris's idea of a 'least grammar', and he makes the connection to the ideas of computational complexity developed by Andrei Kolmogorov, Ray Solomonoff, and others. Here I think Ryckman is right in broad outline but off the mark in the details.¹³ He suggests that since by definition language structure is a structure of the restrictions on combinations of elements recognized by the speakers of the language in question, it is imperative that the characterization of this structure (in a 'grammar') not contribute to the redundancies of combinations, the bearers of information in the language, to be described. This is not just the general methodological virtue of economy of means. If language structure is informational structure, the requirement of a 'least grammar' is not a nicety, it is a necessity (31).

And here Ryckman cites (31, N. 18) Harris himself: 'the grammatical description [must be kept] as unredundant as possible so that the essential redundancy of language, as an information-bearing system . . . not be masked by further redundancy in the description itself' (Harris 1982:10–11). Sympathetic though I am to the enterprise, this seems to me to be incorrect. One can have two grammars, each complete and each assigning exactly the same probabilities to exactly the same set of sentences, and at the same time, one grammar can be more redundant than the other, however one chooses to define redundancy ('wordier', longer in its formulations). For example, one of the two may distinguish between a generalization holding in main clauses and the same generalization holding in embedded clauses, while the other states the generalization once for all clauses. From an information-theoretic point of view, both grammars perform precisely the same function as long as they assign the same probabilities, and neither Ryckman nor Harris explain why the shorter, less redundant version is superior to the longer. Additional considerations need to be brought to bear here, it seems to me, of the sort offered by minimum description length analysis (Rissanen 1989) and others; I have tried to address this question in a Bayesian fashion in Goldsmith 2005, where I suggest that the motivation for formal accounts of empirical data is based implicitly on a prior distribution over formal accounts.¹⁴

5. LEARNING. LILA GLEITMAN ('Verbs of a feather flock together II', 209–29) explores an exegesis of Harrisian doctrine from a related perspective, but rather than telling us what Harris actually meant, she offers a personal overview of what we have learned in the last several decades about child language learning, and she highlights the many aspects of that learning which have turned out to be DISTRIBUTION LEARNING in important ways. There are several ways that this could be understood: one is that this exegesis is what Harris really meant, and he was right when he meant it this way, and if he didn't tell the whole story, at least he got some significant part of it right (this interpretation would be a continuation of Lin's paper, and it's not what Gleitman intends); a second interpretation would be that Harris's ideas were important in a closely related field that he didn't happen to be all that interested in (which is closer to what Gleitman means, I think); and a third interpretation is that Harris's ideas, grounded as they were in a deep intuition about the nature of language, gave a picture of any POSSIBLE

¹³ I have tried to deal with this subject in Goldsmith 2005 and also in a particular context, that of morphology, in Goldsmith 2001.

¹⁴ In my opinion, there is another important idea here that neither Ryckman nor Harris have exploited, but which is the heart of minimum description length analysis, and that is that both information content and grammar conciseness are notions that can be quantified in the same unit, the (information-theoretic) bit. One can thus coherently ask of an analysis that it MINIMIZE the sum of the grammar length (in bits) and the adequacy of the data modeling (also measured in bits).

theory of language analysis, so it's not surprising that ONE class of language-learners, children, follow many of his proposals. I would opt for the third interpretation of why children, follow many of his proposals. I would opt for the third interpretation of why Harris's ideas are as relevant as they are to the psycholinguistics of language acquisition: Harris was interested in determining what procedures IN PRINCIPLE could lead to a deep understanding of a natural language system, so it shouldn't be surprising that the one existing system that actually acquires a natural language should display a set of behaviors that resemble in interesting ways a Harrisian system.

We can go a step further and inquire (so to speak) of Gleitman what aspects of contemporary linguistics are of the most value in understanding child language acquisition. Some mainstream theorists have opted in strong and even strident terms for the position that language acquisition is not learning at all, but is rather akin to the maturation of a biological organ like a liver or an eye. It seems to me that at a point like this, there must be an amicable parting of the ways between such theorizing and the study of child language acquisition: such linguistic theory has nothing to say of interest to psycholinguists. Gleitman goes on to give an overview of some of the most interesting results in language acquisition of the last decade, and not surprisingly, they involve, well, learning: real learning, from evidence, data in the language at hand.

She points to work done during the 1990s by Patricia Kuhl, Peter Jusczyk, and others demonstrating the early acquisition by children of knowledge of the distribution of sounds in acoustic space. This work suggests that in some fashion children recognize that the linguistic sounds they hear are not uniformly distributed over acoustic space, but rather are distributed in a more finely articulated way—in a distribution that is the sum of several gaussians (several normal distributions), perhaps, where each gaussian represents the sound realization of a phoneme, or rather something at a slightly finer level of structure; armed with such a parametric model of sound distribution, children deform the metric of acoustic space in such a way that two sounds close to the center of the same gaussian are perceived as being extremely close to one another. If this view is correct, the learning that is involved is based on distribution, not the distribution of a structuralist linguist such as Harris, but rather of the statistician who enforces a parametric distribution on the data.

Gleitman points as well to the influential and widely cited Saffran, et al. 1996, which provides evidence that children learn something like words even without knowing that they are symbolic—that is, that they are associated in any sense with meaning. Experiments have been constructed in which within chunks of word-size length, the mutual information between successive segments is high, while across these chunks the mutual information is zero ('mutual information' here measures the degree to which there is a statistical connection between the values taken on by successive phonemes, crudely put). Mutual information can be understood as a way of modeling what linguists would call phonotactics, and the laboratory evidence suggests that learners are capable of using this information in identifying language chunks.

How do children learn about categories of words, such as noun and verb? Gleitman notes:

The most familiar modern statement of a distributional learning algorithm for grammatical categories is from Maratsos & Chalkley (1980) whose schema directly follows linguists including Bloomfield (1933) and Harris (1951). They proposed that children could sort words into grammatical categories by noting their co-occurrences with other morphemes and their privileges of occurrences in sentences. Thus *-ed* is (probabilistically speaking) a verb-follower and *the* is a noun preceder. That these analyses are carried out by children even where the supportive semantic evidence is absent . . . makes a strong

489 case for the continuing power of distributional analysis at higher linguistic levels Just as Harris
 490 proposed for linguistics-internal descriptive purposes, there is by now abundant evidence that probabilis-
 491 tic distributional analyses operate again and again to build up phonetic, syllabic, word, and sentence-
 492 like units and—by virtue of the same procedures—to grasp the distributional properties within each
 493 level. (214)

494 Here I have some trouble with the direction of the discussion. As far as I can see,
 495 psycholinguists do not have the tools to develop what is or isn't a good 'noun preceder',
 496 and if Bloomfield or Harris thought that *the* is a good noun preceder, it may mean that
 497 they didn't look at what words actually follow *the* in English. Most of the highest
 498 frequency words following both *the* and *a* are actually not nouns, but adjectives and
 499 adverbs. In the Brown Corpus, for example, the five most popular words after *the* are
 500 *first, same, most, other, and new*; after *a*, the top ten are: *few, little, new, man, good,*
 501 *small, great, very, long, and number* (all in these orders). Computational linguists are
 502 working these days on developing perspectives that will categorize words into a pattern
 503 that resembles traditional linguistic labeling; the results are either promising or disap-
 504 pointing, depending on your expectations (I lean towards the former).

505 But the heart of Gleitman's argument deals with a point that speaks to one of the
 506 central points in Harris's program. It involves a skill that English speakers have that
 507 Gleitman has been interested in since she was a student of Harris's (and I remember
 508 her presenting it to my introductory linguistics class in 1969 when she called it 'the
 509 great verb game'). Tell a group of linguistically sophisticated people that you are
 510 thinking of a verb, and present them with one syntactic subcategorization frame (e.g.
 511 NP ___ NP *from* NP). Rarely will that be enough information to guess the verb. Give
 512 them another frame in which the same verb appears, and then another, and generally
 513 someone will guess the right verb after three or four such clues. Furthermore, the set
 514 of verbs that people guess become more and more semantically limited and homogene-
 515 ous as they are given more syntactic frames. Why is this? What does this tell us about
 516 the importance of distributional characteristics of words in language? And can it be
 517 linked to other things that we know about the acquisition of language, such as the
 518 robust and highly significant finding that in early language acquisition, nouns dominate
 519 the child vocabulary, a result 'robust across individuals and across languages' (216)?

520 Gleitman's Harrisian idea (which she calls SYNTACTIC BOOTSTRAPPING) is that knowl-
 521 edge of verbs is based largely on the syntactic environments in which they appear:
 522 such knowledge is a prime component of the relevant input to the learning procedure
 523 needed to learn verbs, in any language. Knowledge of nouns, by way of contrast, is much
 524 less dependent on syntactic context, and knowledge of the environment of language use
 525 can play a much more robust role in learning the meaning and function of nouns, at
 526 least, of concrete nouns. And, in her words, 'the required "sophisticated" linguistic
 527 representations [necessary to learn verbs] must themselves be constructed by using
 528 lower-level representations as the scaffolding' (217).

529 Gleitman and her colleagues presented to subjects (college-age students) video clips
 530 of mother-child interactions without sound and told them that the mother uttered a
 531 common noun or verb (she told them which) at the moment that the subjects hear a
 532 beep along with the video. Thus the subjects have most of the pragmatic context that
 533 the child might use to infer the MEANING that the mother intended, at that moment; the
 534 subjects, furthermore, already KNOW the words and merely need to pick out the one
 535 they think the mother must have meant in that context.

536 The results from the experiment are very clear: the task is much easier when the
 537 subject is asked to guess which noun the mother used, compared to when the subject

538 must guess the verb. At the same time, further evidence strongly suggests that the most
 539 significant factor in accounting for this difference is not the grammatical category
 540 difference, but the difference in IMAGEABILITY, a gradient property on which verbs as
 541 a whole uniformly score worse than nouns as a whole (though there is additional
 542 refinement to the scale that goes well beyond grammatical category). Gleitman con-
 543 cludes that a person can learn a lot about noun-word meanings, given real world interac-
 544 tion but no knowledge of linguistic context, but a person cannot learn much about verb
 545 meanings without the linguistic context. Once a person knows enough about the lan-
 546 guage to be able to identify nouns in the speech stream, it becomes possible to learn
 547 the crucial properties of verbs, especially those that help identify many of the subtle
 548 semantic characteristics of a verb's meaning.

549 This is a subtle and extremely appealing vision of some of the critical aspects of
 550 basic language learning and has the merit of introducing a pair of highly correlated
 551 distinctions (imageability on the one hand, and noun/verb on the other) whose relevance
 552 to acquisition was not foreseen by linguistic theory (Ryckman presents Harris's imag-
 553 ined scenario for the way children project from reality to words (32ff.), and there is
 554 little hint that nouns and verbs function according to different principles).¹⁵ I can only
 555 conclude that this is an important consequence of Harrisian sensitivity to distributional
 556 analysis, one that Harris himself might never have countenanced. Gleitman ends her
 557 paper with a thought about what Harris's reaction to her paper might have been: 'I
 558 imagine that if Harris saw these findings (and if he could for a moment stifle a natural
 559 modesty) he would have responded "Obviously". And he would have been literally
 560 correct had he said "I told you so". Yes he did, clearly and in detail'. (227).

561 **6. OTHER MATTERS: SYNTACTIC STRUCTURE, SUBLANGUAGES.** The late MAURICE GROSS
 562 ('Consequences of the metalanguage being included in the language', 57–67) suggested
 563 that 'Harris demonstrated exceptional intellectual courage in abandoning the notion [of
 564 SUBJECT] and adopting for the description of sentences the general schema: $N_0 V W$
 565 (60–61). He goes on to explore a few of the nonsentential structures found in English
 566 (*Good night, Merry Christmas, The hell with that!*), and the difficulty of establishing
 567 generalizations without significant exceptions.

568 On quite another hand, PIETER A. M. SEUREN ('Pseudoarguments and pseudocomple-
 569 ments', 179–207) discusses the difference between arguments and adjuncts, exploring a
 570 number of related questions, such as differences between pairs of sentences given in 1.

- 571 (1) a. She gave her sister a hug.
 572 b. *She gave a hug to her sister.
 573 c. She wrote a letter to the Pope.

875 ¹⁵ Needless to say, there are enormous additional steps that need to be taken, including one place where
 876 I think the present paper understates the level of the challenge for language learning; Gleitman notes that
 877 'a correlated language-internal cue to subjecthood is that different categories of nouns probabilistically
 878 perform different thematic roles. . . . Specifically, animate nouns are vastly more likely than inanimates to
 879 be the causal agents in events. Once the position of the sentence subject is derived by matching up the
 880 observed agent with its known noun label, the young learner has a pretty good handle on the clause-level
 881 phrase structure of the exposure language' (223). Actually, I don't think this could be called a good handle
 882 on the data needed to infer phrase structure. To make this point in detail would be long, but imagine how
 883 difficult the task would be of writing an automatic algorithm that would take in an unlabeled and unanalyzed
 884 text in an unknown language and infer the phrase structure rules of the language. Hard way past what we
 885 know how to do at present. Now imagine the same text, but the phrases corresponding to the most animate
 886 referents in the discourse have been marked as such (by some oracle, so to speak). Does that make the task
 887 any easier? Not in any way we know today.

581
584
587

- d. She wrote the Pope a letter.
- e. John loves Africa.
- f. Africa is loved by John.

590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617

But it is not hard to find many sentences of the sort exemplified by 1b on the internet (taking it as typical of searchable collections of textual data), such as *The other day, I gave a hug to a vendor named Jerome in an awkward moment; Howie's mother gave a hug to the girl walking out of the kitchen; or Even more strangely I noticed that Sharon gave a hug to the head of the Psychology department, who would eventually be my professor that second semester*; and so on, to the point where one wonders why someone might have starred 1b. Seuren suggests that 1d, unlike 1c, evokes a relation of familiarity between the Pope and the letter writer, yet we find many examples of the following sort on the internet, presupposing no such familiarity: *An American delegation arranged to go to see the Pope. I couldn't go, but organizers said that if I wrote the Pope a letter, they would deliver it personally.* Or again: *With a quick Google search we found the Pope's address in Vatican City. We wrote the Pope a letter and we are now anxiously awaiting his reply.* Seuren offers an analysis which provides for a deep syntactic distinction (with important semantic consequences) between the location of NPs complement to prepositions and those that are syntactic arguments of the verb; I, at least, would have been more convinced by an analysis which predicted more robust grammatical differences. Indeed, without wanting to get deeply into a question of the appropriate origins of data for grammatical analysis, I have already indicated that one major difference between Zellig Harris's conception of linguistic analysis and that of generative grammar's conception is that Harris took the task to be analysis of the totality of intersubjective linguistic events (which includes, to be sure, much of what the internet has to offer), something similar to what Chomsky has called E-language,¹⁶ while generative grammar takes the task to be the description of I-language, the internalized knowledge of grammar possessed by humans. The Harrisian E-view has been criticized for being, in some critical cases, too sparse: where, oh where, are the examples of parasitic gaps on the internet? But by the same token, we are becoming increasingly sensitive to the fallibility of data on which the Chomskyan I-view is based (for a recent discussion, see Bresnan & Nikitina 2004).

618
619
620
621
622
623
624
625
626
627
628

MORRIS SALKOFF ('Some new results on transfer grammar', 167–78) presents several fascinating cases that arise from a systematic study of the patterns of translation between French and English, work done in the context of a machine translation project. Salkoff extends some of Harris's suggestions on the general theme with striking illustrations. Particularly interesting are the cases where two languages (here, French and English) have a semantic contrast that is in rough correspondence, such as the French *être en train de* and the English progressive *be + ing*. It will be true for most English verbs that they can appear in the English progressive form if and only if their best French translation can appear after *être en train de*. And yet there are exceptions at the edges: if *Max est en train de comprendre le problème* is fine ('Max is beginning to understand the problem'), we do not find *??Max is understanding the problem*.¹⁷ And, we might

888
889
890
891
892

¹⁶ Chomsky (1986:20) defines E-language as a genus of technical concepts regarding language with the characteristics that the 'construct is understood independently of the properties of mind/brain'.

¹⁷ Actually, we do find it. There are hundreds of examples (though not with the name Max) on the internet, most of them quite natural, such as *Throughout, Rebecca has been nodding or saying OK, letting the mediator know she is understanding.*

629 add, there are mismatches the other way: if McDonald's in English can declare *I'm*
630 *lovin' it*, McDo (in France) cannot translate this as *??Je suis en train de l'aimer*.¹⁸

631 There are two contributions to this volume, 'Grammatical specification of scientific
632 sublanguages' (89–101) by MICHAEL GOTTFRIED and 'Classifiers and reference'
633 (103–16) by JAMES MUNZ, which deal with Harris's treatment of sublanguages, espe-
634 cially as detailed in Harris et al. 1989. I found the papers hard going. If I understand
635 the challenge here, it is this: how can linguistic analysis shed light on a set of documents
636 which are all about a similar theme and part of a common style of discourse? This is
637 not a question linguists generally ask, but whether it is a reasonable question to ask
638 depends mainly on whether insightful or useful answers follow, and it is not easy for
639 me to say whether the results are revealing in the case of scientific sublanguages.

640 In the same section, CARLOTA SMITH ('Accounting for subjectivity (point of view)',
641 137–63) offers a number of insightful comments on indicators of perspective and point
642 of view, based in part on a number of examples from fiction. DAYTHAL L. KENDALL
643 ('Operator grammar and the poetic form of Takelma texts', 261–78) gives a fascinating
644 analysis of a Takelma text, 'Coyote's rock grandson', with a range of structures binding
645 the sections of the story. The late FRED LUKOFF ('A practical application of string
646 analysis', 279–304) discusses the usefulness of grammatical simplifications, identified
647 using string analysis, for learners of Korean.

648 **7. PHONOLOGY.** In 'On the bipartite distribution of phonemes' (241–58), FRANK
649 HARARY and STEPHEN HELMREICH return to the subject of an article by Harary and
650 Paper (1957) on the distributional properties of vowels and consonants. They explore
651 a radically distributionalist view of what it means to distinguish between vowels and
652 consonants, adumbrated in Harris 1951. This view says that the proper division of
653 segments in a language into vowels (V) and consonants (C) is that division for which
654 the largest number of transitions from one segment to the next (in the utterances of the
655 language) are transitions from vowel to consonant or consonant to vowel; alternatively
656 put, it is that division for which the smallest number of transitions are C to C or V to
657 V. Harary and Helmreich suggest associating with any such division a number that
658 they call the BIPARTNESS RATIO, the ratio of the transitions in a corpus that are from V
659 to C, or C to V, divided by the total number of transitions in the corpus; this ratio is
660 thus bounded from above by 1.0.

661 Digressing slightly from the paper, from a mathematical point of view, this is natu-
662 rally expressed in the following terms: the vowel/consonant distinction can be thought
663 of as a CHARACTERISTIC FUNCTION, a function whose domain is the phonemes and whose
664 range is the set {0,1}; phonemes mapping to 1 are vowels, and those mapping to 0 are
665 consonants. An analysis that looks for the function that maximizes a mapping from
666 functions to real numbers is referred to as a VARIATIONAL problem.

667 Harary and Helmreich's procedure is brute force, in the sense that they consider each
668 of the 2^{k-1} ways of dividing k phonemes into two groups. For a language with as
669 many as seventy-three phonemes, as is the case with Achumawi, one of the languages
670 they explore, it is not practical to search all of them. I return to an alternative strategy
671 below.

672 The authors explore Hawaiian, Esperanto, Spanish, and Achumawi, and obtain rea-
673 sonable results in each case, but the results depend rather heavily on whether the digraph

893 ¹⁸ In fact, McDo does not translate it for most of their publicity in France, leaving it in English; but when
894 they do translate it, the sentence becomes *C'est tout ce que j'aime*.

674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708

that links adjacent words is counted or not,¹⁹ and on whether the relative frequencies of transitions are counted (as I have suggested thus far they are) or not: in the latter case, a phoneme-to-phoneme transition is counted as occurring once whether it has occurred once or 10,000 times. The best analysis is one in which word-word transitions are ignored, and the frequencies of transitions are taken into consideration—not surprisingly, in my view.

What is the point of this? The authors suggest some practical applications and speculate on how these procedures could be extended to further phonological questions. But the point of this work within a book such as this one is that it serves to justify Harris's position of the radical autonomy of linguistics even *vis-à-vis* phonetics: the vowel/consonant distribution can be justified as the optimal solution of a problem conceived in purely linguistic terms, or rather, in terms that can be reduced to representations that consist of sequential and discrete sequences of symbols chosen from a phonetic alphabet, that is, phonetic transcriptions.

Harary and Helmreich employ the terminology of graph theory in their exposition, as the term 'bipartness' (above) suggests: the nodes of the graph in question are the segments of the language, the edges are the phonologically permitted transitions, and the bipartness ratio is a quantitative measure of the degree to which the graph satisfies the conditions for being a bipartite graph (one in which the nodes can be divided into two proper subsets in such a way that no edge has both vertices in the same subset). There have been developments in graph theory in the past two decades that provide what I think are rather superior techniques and results to those described in this paper, however. Spectral graph theory employs techniques of linear algebra, applied to matrices derived from graphs, and one of the most important basic results in this area is that the eigenvector corresponding to the second smallest eigenvalue of the laplacian of a graph satisfies the variational problem that Harary and Helmreich seek to solve (see Chung 1997 or Godsil & Royle 2001 for a gentler introduction to spectral graph theory).²⁰ Two conclusions follow from this last observation. First, brute-force total search methods are not needed; a very simple and quick eigenvector decomposition does the trick, from a computational point of view. Second, the results that are obtained distribute the segments over a continuum, with roughly half assigned negative numbers and roughly half assigned positive numbers.²¹ If we sort the phonemes of a language by their associated coordinate in this eigenvector, we get a graded statement of the cline from the most vowel-like to the most consonant-like phonemes of the language—a neat trick: in effect, a purely distributionalist account of the sonority hierarchy.

895
896
897
898
899

¹⁹ That is, since there is in most languages no local phonological principle that governs the transition from the last phoneme of a word to the first phoneme of the next, that transition is governed by word transition probabilities and the distribution of vowels vs. consonants in word-initial and word-final positions. Some languages, such as Spanish or Swahili, have strong biases towards vowels in word-final positions, while others, such as English, do not.

900
901

²⁰ I explore such an approach in the context of syntactic distribution in Belkin & Goldsmith 2002, and in the context of phonology in a paper in progress.

902
903
904
905
906
907
908
909

²¹ The result depends on being able to perform the eigenvector decomposition on a symmetric matrix, which requires some nonobvious manipulations. With appropriate manipulations we can restrict our attention to undirected graphs, hence symmetric matrices, hence everywhere real eigenvalues, and a situation in which all eigenvectors are orthogonal. The explicit coordinates of the eigenvector associated with the lowest eigenvalue takes on values that reflect the frequencies of the phonemes (that is, of the nodes), and these are all positive; since all other eigenvectors must be orthogonal to this first eigenvector, we know that the coordinates of the other eigenvectors distribute themselves with roughly half their values greater than zero, and roughly half less than zero, as noted in the text.

709 LEIGH LISKER ('The voiceless unaspirated stops of English', 233–40) presents experi-
 710 mental evidence regarding the phonetic realization of /b/ after /s/ ('this bout'), of /p/
 711 after /s/ ('the spout') if this is indeed a /p/, and a word-final /p/ before a vowel ('[dro]p
 712 out'). In the experiment he presented the phonetic material that immediately follows
 713 the labial closure in each case and asked the subjects to determine which of three
 714 natural-source sentences (e.g. Did he win this bout?') each sample came from. The
 715 result was that speakers could not distinguish the three, suggesting a single phonetic
 716 notation is appropriate for these three cases.

717 **8. FINALLY: WHY IS HARRIS'S WORK SO UNKNOWN?** Harris's work after *Structural*
 718 *linguistics* (1951) is generally little known, especially his work in the years after 1960.
 719 Why is this the case? Surely it is not because there was little of it, nor because it was
 720 of no inherent interest. It did not help that Harris did not attend linguistics conferences,
 721 but then again, neither has Chomsky. Perhaps it would have helped if Harris had come
 722 to LSA meetings. Chomsky's views have been consistently presented at linguistics
 723 conferences by his students, but Harris's views, by and large, have not been. Perhaps
 724 it would have helped if Harris's students had defended his ideas at LSA meetings. Even
 725 so, even if these things had occurred, it would have been necessary to go to Harris's
 726 writings, and they are neither easy to read nor captivating. Perhaps what they lack most
 727 of all is what Thomas Kuhn (1962) originally referred to as PARADIGMS (before the
 728 term took on additional freight): good examples of analyses that are striking and that
 729 make the reader want to copy them, in a creative sort of way. It is not that Harris was
 730 incapable of providing such paradigms; his idea in 'From phoneme to morpheme'
 731 (1955) about detecting morpheme boundaries by counting possible phoneme successors
 732 had just that degree of insight, cleverness, and elegance, and there are linguists who
 733 continue to find research inspiration in that suggestion.

734 So one possible answer lies in the failure of the Harrisian school to produce a cadre
 735 of dedicated young scholars willing to proselytize, while another lies in the failure to
 736 develop a product that could be mass-marketed. A third surely lies in the poor fit
 737 between Harris's style of linguistics and the superfield of cognitive science that held
 738 sway during the last quarter of the twentieth century.

739 Zellig Harris's work offered more opportunities to the style of work that was done
 740 not by linguists over the last three decades, but by computational linguistics over the
 741 last fifteen years. But that is a story for another time, and the focus of vol. 2 of Nevin's
 742 presentation of Harris's legacy.

743

REFERENCES

- 744 BAR-HILLEL, YEHOShUA. 1954. Logical syntax and semantics. *Language* 30.230–37.
 745 BAR-HILLEL, YEHOShUA. 1964. *Language and information: Selected essays on their theory and application*.
 746 Reading, MA: Addison-Wesley.
 747 BELKIN, MIKHAIL, and JOHN GOLDSMITH. 2002. Using eigenvectors of the bigram graph to infer morpheme
 748 identity. *Morphological and phonological learning: Proceedings of the sixth meeting of the ACL Special*
 749 *Interest Group in Computational Phonology (SIGPHON), Philadelphia, July 2002*, ed. by Michael
 750 Maxwell, 41–47. Philadelphia: Association of Computational Linguistics. Online: [http://acl.ldc.upenn](http://acl.ldc.upenn.edu/acl2002/MPL/contents.htm)
 751 [.edu/acl2002/MPL/contents.htm](http://acl2002/MPL/contents.htm).
 752 BLOOMFIELD, LEONARD. 1933. *Language*. Chicago: University of Chicago Press.
 753 BRESNAN, JOAN, and TATIANA NIKITINA. 2003. On the gradience of the dative alternation. Online: [http://](http://www-lfg.stanford.edu/bresnan/new-dative.pdf)
 754 www-lfg.stanford.edu/bresnan/new-dative.pdf.
 755 CALABRESE, ANDREA. 2004. Prolegomena to a realistic theory of phonology. Online: [http://web.gc.cuny.edu/](http://web.gc.cuny.edu/Linguistics/events/phonology_symposium/Calabrese_paper.doc)
 756 [Linguistics/events/phonology_symposium/Calabrese_paper.doc](http://web.gc.cuny.edu/Linguistics/events/phonology_symposium/Calabrese_paper.doc).
 757 CARNAP, RUDOLF. 1934. *Logische Syntax der Sprache*. Vienna: J. Springer.
 758 CHOMSKY, NOAM. 1981. *Lectures on government and binding*. Dordrecht: Foris.

- 759 CHOMSKY, NOAM. 1986. *Knowledge of language: Its nature, origin and use*. Westport, CT: Praeger.
- 760 CHUNG, FAN. 1997. *Spectral graph theory*. Providence, RI: American Mathematical Society.
- 761 GODSIL, CHRIS, and GORDON ROYLE. 2001. *Algebraic graph theory*. Berlin: Springer.
- 762 GOLDSMITH, JOHN. 2001. Unsupervised learning of the morphology of a natural language. *Computational*
- 763 *Linguistics* 27.153–98.
- 764 GOLDSMITH, JOHN. 2005. From algorithm to generative grammar and back again. *Papers from the fortieth*
- 765 *meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistics Society, to appear.
- 766 HARARY, FRANK, and HERBERT H. PAPER. 1957. Toward a general calculus of phonemic distribution. *Language*
- 767 33.143–69.
- 768 HARRIS, ZELIG S. 1941. Review of *Grundzüge der Phonologie*, by Nikolai Trubetzkoy. *Language* 17.345–49.
- 769 HARRIS, ZELIG S. 1946. From morpheme to utterance. *Language* 22.161–83.
- 770 HARRIS, ZELIG S. 1951. *Methods in structural linguistics*. Chicago: University of Chicago Press.
- 771 HARRIS, ZELIG S. 1955. From phoneme to morpheme. *Language* 31.190–222.
- 772 HARRIS, ZELIG S. 1976. A theory of language structure. *American Philosophical Quarterly* 13.237–55.
- 773 HARRIS, ZELIG S. 1982. *A grammar of English on mathematical principles*. New York: John Wiley and
- 774 Sons.
- 775 HARRIS, ZELIG; M. GOTTFRIED; THOMAS RYCKMAN; PAUL MATTICK, JR.; ANNE DALADIER; T. N. HARRIS;
- 776 and S. HARRIS. 1989. *The form of information in science: Analysis of an immunology sublanguage*.
- 777 Dordrecht: Kluwer.
- 778 HUCK, GEOFFREY, and JOHN GOLDSMITH. 1996. *Ideology and linguistic theory*. London: Routledge.
- 779 KUHN, THOMAS. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- 780 MARATSOS, MICHAEL, and MARY ANNE CHALKLEY. 1980. The internal language of children's syntax: The
- 781 ontogenesis and representation of syntactic categories. *Children's language*, vol. 2, ed. by Keith Nelson,
- 782 127–214. New York: Gardner Press.
- 783 MATTHEWS, PETER. 1999. Zellig Sabbetai Harris. *Language* 75.112–19.
- 784 REICHENBACH, HANS. 1938. *Experience and prediction: An analysis of the foundations and the structure of*
- 785 *knowledge*. Chicago: University of Chicago Press.
- 786 RISSANEN, JORMA. 1989. *Stochastic complexity in statistical inquiry*. Singapore: World Scientific.
- 787 SAFFRAN, JENNY R.; RICHARD N. ASLIN; and ELISSA L. NEWPORT. 1996. Statistical learning by 8-month old
- 788 infants. *Science* 274.1926–28.
- 789 SAMPSON, GEOFFREY. 2003. Are we nearly there yet, Mum? University of Sussex, ms. Online [http://](http://www.grsampson.net/Aawn.html)
- 790 www.grsampson.net/Aawn.html.
- 791 SAPIR, EDWARD. 1933. La réalité psychologique des phonèmes. *Journal de psychologie normale et patholog-*
- 792 *ique* 30.247–65. [Reprinted in translation in *Phonological theory: Evolution and current practice*, ed.
- 793 by Valerie Becker Makkai, 22–31. New York: Holt, Rinehart and Winston, 1972.]
- 794 SHAUMYAN, SEBASTIAN. 1996. On psychologism in linguistics. Online: [http://test.linguistlist.org/issues/7/](http://test.linguistlist.org/issues/7/7-1478.html)
- 795 [7-1478.html](http://test.linguistlist.org/issues/7/7-1478.html).
- 796 TRUBETZKOY, NIKOLAI. 1939. *Grundzüge der Phonologie*. (Travaux du Cercle linguistique de Prague 7.)
- 797 Prague: The Linguistic Circle of Copenhagen and the Ministry of Public Instruction of the Republic
- 798 of Czechoslovakia.

802 Department of Linguistics
 803 1010 E. 59th Street
 804 Chicago, IL 60637
 805 [ja-goldsmith@uchicago.edu]

[Received 2 September 2004;
 accepted 27 September 2004]