Morris Halle (1923–2018)

Paul Kiparsky*

In 1940 Morris Halle, 17 years old, managed to emigrate in the nick of time with his parents from his native Latvia to the United States. After two years of engineering study in City College, New York, he was drafted into the U.S. Army and sent to Europe, where he ended up participating in the liberation of Paris. Discharged in 1946, he began to study Slavic and general linguistics, first at the University of Chicago, and then at Columbia University with Roman Jakobson, whom he followed to Harvard in 1949. In 1951 he was hired at MIT to teach Russian and German in the department of Modern Languages, and to work on phonetics in the Research Laboratory of Electronics (RLE), the peacetime incarnation of the Radiation Laboratory where much of the early work on radar had been done. Shabbily housed in the legendary Building 20, where it would later be joined by the Linguistics Department, RLE was a rich intellectual environment that reflected the scientific ferment of the postwar era, with electrical engineers, mathematicians, biologists, psychologists, as well as researchers on language engaged in a variety of projects on machine translation, acoustics, and speech communication.

During this decade Halle steadily built up MIT’s linguistics course repertoire and piloted it into a full-fledged Ph.D. program, adroitly maneuvering in MIT’s intricate archipelago of departments and research laboratories. In 1953 he introduced a graduate course on “Hearing, Speech and Language” co-taught with Walter Rosenblith, a specialist in the electrophysiology of hearing, which was offered jointly by the Electrical Engineering Department and the Modern Language Department. His key move was to get Noam Chomsky hired in 1955, who offered two undergraduate courses in general linguistics a year, in addition to teaching French and German and working in the machine translation group at RLE. Halle himself began teaching a regular phonology course in 1957. In a style that was to become a hallmark of the department’s curriculum, they integrated their courses with their research as far as possible, presenting their work and inviting students to actively engage with it. Out of these courses came seminal publications such as Chomsky (1957) and Halle (1954, 1957, 1959).

* Author’s address: Paul Kiparsky, Anne T. and Robert M. Bass Professor in the School of Humanities and Sciences, office: 420-024, Margaret Jacks Hall, Building 460 Rm. 127, Stanford, CA 94305-2150, USA. E-mail: kiparsky@stanford.edu.

https://doi.org/10.12697/smp.2018.5.2.06
Halle’s next masterstroke was getting Roman Jakobson appointed in 1960 as Institute Professor to lecture annually on his current research. The topic varied from year to year, and in some years it was metrics and poetics. When the Ph.D. program was officially inaugurated in 1961, Halle recruited additional linguists from RLE’s machine translation project to complete its original faculty, among them G. H. Matthews and Edward Klima.¹

Halle served as the program’s director and would remain its guiding spirit long after he handed the reins to S. J. Keyser in 1977, when the program joined with Philosophy into a department. He was a brilliant administrator, and a fortunate one in two respects. Academic governance in those days was unimaginably light by present standards. Matters that would now require dossiers of paperwork wending their way through a bureaucratic maze could sometimes be settled by Halle with a phone call to Rosenblith (who had meanwhile become Provost of MIT). The other lucky break was that in the fifties the government had begun to allocate sufficient support for basic research, often presented as having immediate practical applications and funded via the armed forces, in order to foil demagogic politicians out to cut taxes. Halle was able to support the entire linguistics program on successive blanket five-year grants from these sources.²

As a result the program grew rapidly. By 1962 there were already 16 graduate students enrolled, a diverse group recruited largely by word of mouth. That year, in another inspired coup, Halle organized the 9th International Congress of Linguists at MIT, featuring Chomsky among the plenary speakers. No-one who saw Chomsky calmly dominate such luminaries as Benveniste and Kuryłowicz in a dramatic debate on the foundations of linguistics could miss that they were witnessing a turning point in the field. Within a few years Halle and Chomsky had built the top linguistics program in the world, with an unmatched record of research and teaching that continues to this day.

A world-class pedagogue, Halle structured graduate study as an apprenticeship – then a radical break with tradition, now emulated throughout the field. Students get offices, and are on their way to independent research within a year of enrolling, always working closely with the instructors – not only with

¹ For Halle’s early years in the program see Harris, Harris 1974 and Halle 2011.
² When one of my first articles, on a point of Classical Greek phonology, was published in a German learned journal, I had to ask its incredulous editor to have a large number of extra reprints made for the various sponsoring branches of the armed forces (400 of them, if memory serves, which for all I know was more than the entire circulation of the journal). They were duly delivered to them, and politely returned with the note that no further documentation of my research would be required.
their formal advisers but in principle with the entire faculty. The faculty might be assigned to qualifying committees regardless of their area of specialization. Everyone is encouraged to be present on a regular basis, to attend the colloquia, to sit in on courses, and to be generally accessible. Faculty cannot live in another city and fly in once a week for their classes, students cannot disappear and then show up one day with a thesis to defend. The sustained mutual interaction of students and faculty fostered in this way builds an esprit de corps that sees students through the rigors of graduate study. In such an environment students learn at least as much from each other as they do in their coursework. In Halle’s words, learning happens best “in conversation between people who [are] interested in the problems” (Halle 2011).

Harder to replicate is Halle’s ability to develop a personal relationship with each student and engage them in the personality-appropriate way. He would patrol the corridors and drop in on the students’ offices, dispensing solace, cheer, insights, admonishment, and blunt reprimands as needed. His gruffness could be off-putting if you didn’t understand the respect and deep humanity it came from. Many times a student in trouble was helped by his personal kindnesses. From the shrinking violet to the oversized ego, from the “smartest boy in Pinsk” to the truly exceptional individual, all did their best under his astute tutelage. He took pride in his ability to manage the occasional problem student – even in one extreme case a violent psychopath who later self-destructed. Only rarely did he have to resort to the ultima ratio of a consilium abeundi (one of his favorite Latin expressions, uttered with a chortle). He turned his students not only into productive contributors to the field but into better and wiser persons.

Halle’s own research has transformed linguistics in many ways. It is the outcome of several concurrent lifetime projects, initiated at different times in his career but never terminated. The first of them concerned the theory of the binary distinctive features that according to Jakobson (1938, 1942) underlay the phonemic contrasts of the world’s languages. Jakobson had given some good phonological evidence for them, but their phonetic definitions remained sketchy, and they had therefore not yet been widely accepted at the time. Advances in electrical engineering and information science during the war had made new speech technology available, including the sound spectrograph, and provided the new mathematical framework of information theory. They fueled a boom in research on acoustics, of which MIT’s RLE became an important center, with notable figures such as Leo Beranek. Halle exploited the new tools to solidify the phonetic foundations of the features, in collaboration with several engineers and phoneticians, notably George Hughes, Kenneth Stevens, and Gunnar Fant. The results were presented in the classic
Paul Kiparsky  

Preliminaries to Speech Analysis (Jakobson, Fant, and Halle 1952), followed by the information-theoretically oriented Cherry, Halle and Jakobson (1953), the less technical Fundamentals of Language (Jakobson and Halle 1956), Halle and Hughes (1956), Halle, Hughes, and Radley (1957), and Halle’s *The sound pattern of Russian* (1959). Soon the quantal theory of Halle’s RLE collaborator Kenneth Stevens would provide independent articulatory phonetic support for the distinctive features by showing that they correspond to particularly stable configurations of the vocal tract, at which small displacements of the articulators have little effect on the sounds produced, making consistent speech production possible.\(^3\)

With feature theory on a more solid footing, Halle updated it in the following decades with a series of revisions and extensions in response to a flood of new descriptive and theoretical work that he himself had done much to stimulate. In Chomsky and Halle (1968) he added the feature [coronal], which together with high and back replaces Jakobson’s [compact/diffuse] and [grave/acute], and separate features for lip rounding and pharyngealization to replace Jakobson’s [flat]. Later he added a pair of antagonistic features which define tongue root position (Halle, Stevens 1969), laryngeal features which characterize voicing, aspiration, and pitch (Stevens, Halle 1971), and features for clicks (Halle, Vaux, Wolfe 2000). He also weighed in on the neuro-anatomy of the speech mechanism, to the point of associating the feature values with specific muscular gestures in the vocal tract (Halle 1983), and on the grouping of features into hierarchically organized bundles that can function as units in phonological rules (feature geometry, Halle 1995).

In five decades Halle had shaped feature theory into a form that since then has been widely adopted in phonetics and phonology and hardly even seriously challenged. If this was an area in which he carried Jakobson’s earlier brilliant initiatives to a successful conclusion, his other work was a series of his own brilliant initiatives, usually in collaboration with colleagues or former students, in which a somnolent field would be rudely kicked awake with well-reasoned unorthodoxies, usually too radical to gain universal assent.

The first of these was phonological theory. In his earliest writings Halle still inhabited Jakobson’s structuralist world, where phonology was a system of contrastive segments bi-uniquely related to phonetics, which was strictly separated from morphophonology (Halle 1951, 1953). Through discussions with Chomsky he came to see phonology in a different way: as a component

---

\(^3\) Halle and Murray Eden also applied the idea of feature decomposition to cursive handwriting (Eden, Halle 1961).
of generative grammar that maps an input representation containing only phonological information to an output representation by a system of unidirectional ordered rules, subject to its own principles of locality and interfacing with other components including syntax.

The inaugural presentation of this new conception was Chomsky, Halle, and Lukoff (1956), which showed that the stress patterns of English sentences can be predicted from syntactic structure. Halle (1961, 1962) went on to argue that phonological descriptions are governed by the principle of simplicity, that this principle requires that phonological rules be formulated in terms of features, and that the rules apply in a strict order, in at least some cases cyclically from the innermost constituents outward. He further demonstrated that under these assumptions the phonemic level, in the sense of a representation of purely contrastive information, can only be derived at the cost of redundancy and loss of generalizations. From this he drew the conclusion that such a level is merely an artifact of structuralist methodology, and that the only significant levels in phonology are underlying representations and output (phonetic) representations. Halle’s formal argument was incontrovertible, and yet its conclusion was understandably controversial, since the phoneme seemed a useful intermediate level of abstraction for many purposes, including historical linguistics and the study of poetic form.4

The summation of this phase was Chomsky’s and Halle’s *Sound Pattern of English* (1968). It is a remarkable work for several reasons, each of which reflects a facet of Halle’s scientific principles and personality: for the care with which it lays out and motivates the formal principles of the theory, for the depth, detail, and insight of its treatment of English phonology, synchronic as well as historical, and for the self-critical final chapter that draws attention to a fundamental shortcoming of the theory and puts forward a solution for it (which unfortunately turned out to be stillborn). Halle would often return to aspects of English phonology (Halle, Mohanan 1985, Halle, Kenstowicz 1991, Halle 1997), and justifiably demand that alternative theoretical proposals should be as explicit and deal with the complexities of English as successfully as SPE does.

Halle was just as ready to criticize, modify and even abandon his own ideas as those of others. The ability to backtrack quickly out of a dead end was in fact one of the keys to his productivity and creativity as a scholar. But in retrospect it seems that he sometimes gave up too soon. The idea of contrastive feature

4 Stratal OT provides a resolution to this contradiction. The output of the lexical phonology is actually a better characterization of the relevant level than structuralist phonemic representations are (Kiparsky 2018).
hierarchies, which is prominent in Halle 1959, is absent in his subsequent work, and indeed largely missing in the phonological literature of the next forty years. Dresher (2009, 2018) has now revived it and made a good case for its relevance to understanding sound change. A more poignant example is Halle’s disavowal of probabilistic and information-theoretic approaches, in which he had invested a great deal of effort, as having been a waste of time (Halle 1975). Yet this too has begun to flourish again in the last decade, with intriguing results (e. g. Cohen Priva 2017). Like rivers, scientific ideas grow when new tributaries join them, but they can also dwindle or even go underground, to resurface as the terrain or climate changes.

Halle never swerved from the rule-based proceduralism of SPE, and had no use for Optimality Theory, which took phonology by storm in the nineties. He complained that OT proponents had only addressed the cases that yield easily to constraint-based analyses, and that the theory could not deal with the hard cases (Halle 2003: 10–13). This is the objection that every new theory has faced, including generative phonology itself. Theoretical innovations in OT, such Base/Output constraints and Paradigm Uniformity constraints on the one hand, and Stratal OT on the other, have taken much of the sting out of this objection. Trickier were the conceptual objections that Halle raised in collaboration with Sylvain Bromberger, which to my knowledge have not been explicitly addressed. Halle and Bromberger claimed that rule ordering in phonology (and its absence in syntax) is justified by an intrinsic asymmetry in phonological derivations which is absent in syntax: underlying representations are intrinsically prior to surface phonological representations because they are stored in memory, whereas surface phonological representations are generated in the course of a derivation (Bromberger, Halle 1989). The response to this argument would be that the asymmetry does not entail rule ordering, for it is equally compatible with constraint based formalisms such as OT. A second conceptual argument was based on the premise defended in Bromberger, Halle (1995) that phonological theory is not about abstract types, but about concrete intentional events and states that occur in the world, and enter into causal relations. Building on this claim, they then submit that derivational phonology provides a more plausible account of those speech events than OT does (Bromberger, Halle 1997). The counter-argument would be that the objection depends on construing phonology as a performance model, which Bromberger and Halle are careful not to do. Speech production and perception cannot really adjudicate between theories of phonology without an explicit account of how these theories are involved in the use of language.

After the completion of SPE it was a natural step to put the new theory to a test in other domains. This led Halle to investigate, in collaboration with
S. J. Keyser, the role of stress in metrical verse. They laid out a conception of meter as a matching between a simple abstract template (such as an alternating sequence of Strong and Weak positions) related to linguistic representations by correspondence conditions – in the case of English, the STRESS MAXIMUM PRINCIPLE (Halle, Keyser 1968, 1971). Their purely constraint-based approach to the correspondence between metrical pattern and text can be seen as a precursor of developments such as autosegmental phonology, and even of the correspondence theory that later grew out of Optimality Theory. In his subsequent work on meter Halle switched to a mixed theory that conceived meter as a set of parsing operations, but still governed by constraints.

The basis of Halle’s new theory of poetic meter was the bracketed grid approach to stress that he had developed with Vergnaud and Idsardi (Halle, Vergnaud 1987, Halle, Idsardi 1995). As representations of metrical structure, bracketed grids are equivalent to labeled trees. But by reifying the brackets and grid columns as elements that can be inserted, deleted, and moved they opened up a formal Pandora’s toolbox that can do much more than the classical metrical theory of stress. This was the toolbox that was exploited in his work with Fabb (2008). They recast meter as a bottom-up directional parsing procedure that constructs bracketed grids in satisfaction of certain well-formedness conditions. Critics noted that in spite of its richness, this theory of meter shares with the phonological stress theory that it is based on an inability to account for top-down effects. These are in principle unproblematic for constraint-based approaches, which can deal with them by ranking constraints on higher domains ahead of constraints on lower levels to construct an optimal parse.

The book’s boldest theoretical claim is that rhythm is not a constitutive property of meter, but an epiphenomenon that arises in performance as a side-effect of counting syllables or other prosodic units. This view seems at odds with the formalism of bracketed grids itself, which is designed precisely to represent rhythm as periodic alternation of prominence at a hierarchy of levels. It is also at odds with the simple observation that meter always constrains prominence-defining categories such as stress and weight, and with the intimate connections between poetic meter and musical rhythm. The lasting value of the book is to bring a large variety of metrical systems from world literature into the theoretical discourse by working out precise and detailed formal analyses of them.

A second domain in which the theory of stress begged to be tested is historical change. Earlier stages of English yielded some intriguing results (Halle, Keyser 1971), but the real challenge was the movable stress of languages like Russian and the similarly behaving pitch accent of Greek and Sanskrit. Simply listing the paradigms seemed a rather unenlightening type of analysis, but the
tools of SPE did not offer a better one. During one of Halle’s sabbatical years he and I met regularly in his study in Widener Library to tackle this problem. We arrived at a COMPOSITIONAL approach which derives the accent of words from the lexically specified accentual features of their constituent morphemes (Halle, Kiparsky 1979, 1981, Kiparsky, Halle 1977, Halle 1997d, 2001). A default rule (the BASIC ACCENTUATION PRINCIPLE) erases all accents but the leftmost one, and assigns an accent to the left edge of an unaccented domain. Our proposal has only recently begun to gain some traction among Indo-Europeanists.

Morphology is an area that Halle launched twice, each time to great effect. In an early essay (Halle 1973) he raised the problem of gaps in the lexicon – semantically, syntactically, and phonologically possible words which are not actual words, such as *derival and *arrivation. He advocated an overgenerating morphology in which the missing items are generated by word formation rules, but marked as not being subject to lexical insertion, and therefore unavailable to actual use. By highlighting issues of productivity, compositionality, and the treatment of lexical exceptions, the study inspired much of the early work on morphology in generative grammar.

Halle returned to morphology in joint work with Marantz to set forth a new theory, Distributed Morphology (DM), which is realizational but morpheme-based. It holds that all words are built from roots and affixes. Roots are unspecified for lexical category, and become nouns, verbs, or adjectives in virtue of being complements of a functional head v, n, or a, which may be either null, or realized by a derivational affix such as -ize and -ion. DM posits a rich apparatus of morphological rules, such as fission, fusion, impoverishment, and metathesis. There is no “lexicon”; the phonological and morphological content of morphemes is stored in different modules. DM has been enormously successful and is now by far the dominant approach to morphology among syntacticians, even among adherents of minimalism. But the debate continues, for important properties of morphology that come free out of the box in lexicalist approaches, notably locality, inward dependency, cyclicity, and the Mirror Principle, must be stipulated by extrinsic constraints on derivations in DM.

Morris Halle was the most Pāñinian of linguists in his ability to detect patterns in language, in his methodological commitment to an exhaustive accounting of the data, and in taking seriously the theoretical formalism and the principle of simplicity.

In recognition of his achievements Halle was awarded an endowed chair at MIT in 1976, and a lifetime Institute Professorship in 1981. He was elected president of the Linguistic Society of America in 1974, and received honorary doctorates from Brandeis University and the University of Chicago, as well
as several prestigious scientific prizes. His students expressed their admiration, affection, and gratitude by editing and contributing to three festschriften dedicated to him (1973, 1984, 2013).

On April 2, 2018, Morris Halle died of heart failure, in peace and surrounded by his sons David, John and Tim, and their children.

References


Halle, Morris 2011. 50 years of Linguistics at MIT, Lecture 10. https://www.youtube.com/watch?v=g8g0A8QR61s


Harris, Florence; Harris, James 1974. The development of the Linguistics program at the Massachusetts Institute of Technology. http://ling50.mit.edu/harris-development


Jakobson, Roman; Fant; Gunnar; Halle, Morris 1952. Preliminaries to speech analysis: the distinctive features and their correlates. Cambridge, MA: Acoustics Laboratory, Massachusetts Institute of Technology.


