

# **Computerized editions of oral epic poetry : the evolution of the text-processor Heuro-I**

by

John Miles FOLEY

*University of Missouri - U.S.A.*

An oral epic text is not and cannot be *the* oral epic, for in oral tradition the work or song forever eludes any single or final textual representation<sup>(1)</sup>. A poem survives in oral tradition through its variants or avatars, each of them an equally legitimate version of the work and in fact a text in its own right, but none of them the definitive version<sup>(2)</sup>. To restore to the latter-day reading process a semblance of the original oral experience, it would be necessary to provide a reader not merely with the synchronic reality of the individual text-version but also with the diachronic experience of many text-versions. Ideally such an experience should include variants performed by the same and different bards, as well as variants drawn from different local traditions. In this way the reader could approximate the original oral situation by placing each version in its traditional context, by bringing to bear his knowledge of other related texts upon the text at hand. Interpretation would not be limited to perusal of one version detached from its tradition but would proceed within the traditional context which generates all text-versions.

Along with multiformity at the level of entire texts, the editor of oral epic must consider multiformity on at least two other levels -- those of phraseology and narrative design. As many scholars have shown<sup>(3)</sup>, the diction of an oral epic poem is composed of *formulas*, a unit Milman Parry first defined as "a group of words regularly used under the same metrical conditions to express a given essential idea" (*MHV*, p. 272). These substitutable verse increments are the bard's "words", his phraseological inheritance from the oral epic tradition which he and generations of singers before and after him employ to compose their songs. Formulas are of two major types. Some are fossilized into apparent tags, as in the case of so many Homeric Greek epithets ("swift-footed Achilles", *πόδας ὠκύς* 'Ἀχιλλεύς for example); others may admit more variation, as in the following system of expressions in Anglo-Saxon<sup>(4)</sup>.

System - X under wolcnum	(X under the clouds)
where X = weox (he grew up)	weox under wolcnum <i>Beowulf</i> 8
wann (dark)	wann under wolcnum <i>Andreas</i> 837
wide (widely)	wide under wolcnum <i>Genesis</i> 1950
wod (went)	wod under wolcnum <i>Beowulf</i> 714
waeter (water)	waeter under wolcnum <i>Beowulf</i> 1631
woruld (world)	woruld under wolcnum <i>Genesis</i> 916
etc.	

As the individual texts of a given work vary one to the next within limits, so the formulaic language as a whole admits a modicum of variation within certain limits set by the metrical requirements of the verse and the general semantic thrust of the essential idea underlying the formula.

To the multiformity of song and diction may be added the Protean nature of the narrative scene or *theme*<sup>(5)</sup>. Examples of themes in various oral literatures include the Homeric feast scene, which

follows a generic sequence and admits particular details in conformity with its nominal position in a tale; the Anglo-Saxon *bēot* or boast, the verbal contract to defeat the adversary that is sworn before battle and that has analogues in medieval French and Middle High German<sup>(6)</sup>; and the Serbo-Croatian typical scene of the hero arming himself (or herself) before setting out on a journey, a common theme which has counterparts in many heroic poetries. Whatever the actual content of these narrative units, however, they follow the traditional design of variation within a generalized pattern. The instances of themes are, like instances of formulas and versions of whole songs, individual reflections of oral traditional ideas which, while they are equally faithfully represented in each occurrence, cannot be reduced to any one of them. The ideas and structures which are the lifeblood of the epic tradition are multiforms : they take shape in one or another text, scene, or phrase, but their traditional meaning is the Gestalt of all their occurrences rather than the individual instance at hand.

With this brief sketch of oral traditional structure in mind, we can perhaps appreciate that to edit a poem rich in traditional meaning demands a great deal more than to edit the detached, decontextualized text; to edit the oral epic work is somehow to reactivate its traditional resonance, to provide the reader with the background necessary to a faithful extrapolation from the text to its context and on to the poem. Without such an extension from version to source, without simulating a knowledge of traditional multiforms with which to interpret otherwise apparently unique texts and structures within them, we have reduced the oral epic to a faint echo of its original self. We have, in short, edited the multi-dimensional source poem out of its natural richness and complexity into a manageable but hopelessly monolithic text; and this unfortunate process is an obviously reductive act, an exercise in oxymoron. The key to understanding the text is to recreate the poem, and for oral traditional epic recreating the poem means reinvesting the text with traditional meaning.

The question for the prospective editor thus becomes how to accomplish that reinvestiture, how to remake the poem from the text. Early collectors and editors of Serbo-Croatian oral epic, concerned as they were with other dimensions of what Murko called *la poésie populaire épique*<sup>(7)</sup> characteristically put together what amounted to anthologies constructed along generic lines. Either a focal hero, like Kraljević Marko in the Serbian tradition or Djerdjelez Alija in Moslem epic song, or an area, such as the borderland Krajina, often served as the unifying feature of a published collection. With the appearance of the first two volumes of the Parry-Lord material in 1953-54, the series which has continued under the title *Serbo-Croatian Heroic Songs*<sup>(8)</sup> a first step was taken toward placing the individual text in its traditional context. Here were multiple versions of a poem by the same and different singers within the local tradition of Novi Pazar; by reading through the volume, one can gain some understanding of the poetic context out of which each text emerges. Later volumes in this series carry on the task of documenting the oral epic tradition in the various regions of Yugoslavia in the 1930s and afterward. As the edition progresses, the effect of increased coverage will be cumulative : as more and more material becomes available, the individual texts modulate more and more toward the poems and the tradition which they instance.

As fine and necessary an edition as *Serbo-Croatian Heroic Songs* is, finally it cannot fully restore

traditional meaning to a decontextualized text, in large part because of the endemic limitations of the typographical recension. No printed book can present all available versions of a song concurrently, relating the variant texts by their traditional structures of formula and theme. As post-oral media, printed books must rely on appendices or other conventions of scholarly apparatus to accomplish the work of comparison and reference. In addition, the reader himself must provide the algorithm of reference, as it were : he must find the variant, search its separate inventory of themes and formulas, refer back to the variant text, make the comparison, draw the necessary conclusions, and then return to the original text to continue his reading. The responsibility of establishing a traditional context for a work is thus left entirely to the reader, who must be intellectually nimble enough to juggle object text, variant texts, and appendices while still keeping the progress of the narrative in mind. The printed book, it quickly becomes obvious, simply is not equal to the task of providing an editorial medium for oral epic.

On the other hand, automatic text-processors such as my system HEURO-I offer a far more satisfactory means of extending text to oral poem, of restoring traditional context to a work. As the name implies, HEURO-I is a heuristic series of computer programs which "reads" an object text in terms of its tradition, thus recreating the lost context and revivifying the work. The text-processor recognizes formulaic and thematic elements in an object text and automatically searches its data file of reference texts for other occurrences of the same structures. If correlative patterns are found in the data file, HEURO-I reports them and, depending on their size and importance, either prints them out alongside the object text or offers the reader the opportunity to choose whether he wishes to view them or not. At every point, the reader is able to compare traditional collocations and to interpret the text before him not as an *oeuvre complète* but as one version of a multiform. The object text is no longer unique, no longer cut off from its tradition by the tyranny of the printed book; rather it becomes again a text contextualized, a poem reinvested with the lifeblood of its tradition.

HEURO-I performs these routines hierarchically, proceeding from the macrostructure of themes to the microstructure of formulas. At each level the system concords and lists all occurrences in the data file and offers the analyst the choice of viewing any or all of them alongside the relevant line or passage from the object text. For example, consider the case of the theme of "A Captive Shouts in Prison". As shown below, one instance begins at line 122 :

122 Pocomilijo Alagić Alija THEME 12

The designation THEME 12 refers to a dictionary of typical scenes, abbreviated in the machine-readable texts themselves in order to save space; when the system enters on a thematic search, it prints out the full descriptive name of the theme under the initial line :

122 Pocomilijo Alagić Alija THEME 12

THEME 12 - A CAPTIVE SHOUTS IN PRISON

It then searches the data file for all instances of this narrative unit and reports the first lines and

passage citations, with beginning and ending line numbers<sup>(9)</sup> :

2.42	Pocmilijo tri bijela dana	THEME 12	2.49
3.55	A cmilijo Ograšćić Alija	THEME 12	3.85
3.695	Pocmilijo Alagić Alija	THEME 12	3.729
11.91	Pocmilijo sužanj nevoljnice	THEME 12	11.127

The reader can now choose on the basis of such a report to view any or all of the passages cited, in whole or in part, moving freely along lines of traditional equivalence from one text to another. In the present example, he may want to focus on the passage beginning at line 695 of text 3, since it seems from that notation that a second THEME 12 occurs in that song. This second instance within the same song betokens a double cycle, a multiple of the simplex pattern described above<sup>(10)</sup>, and provides an illustration of how song-types evolve by combination and permutation within an oral tradition.

During this second phase of analysis, HEURO-I will locate and concord all themes as it encounters them. Having accomplished that concordance and given the reader the opportunity to contextualize whatever typical scene he meets in the object text, the system then proceeds to a line-by-line formulaic scan, a second level of analysis which continues until the next thematic label turns up. During the formulaic phase, the reader is automatically provided with correlative phraseology in other texts in the data file : the actual corresponding line in a 2- to 4- line context is printed out alongside the line from the object text. An example follows :

OBJECT TEXT	DATA FILE
"E cu li mene, nerodjena majko,	A govori Bojičić Alija :
	"Ču li mene, nerodjena majko, 1.21
	Ja sam sret 'o tvoja Bojičića,

In practice, the formulaic routine operates through a key-word-in-context processor; that is, the program locates related phraseology by matching key words which are chosen for their relative importance. As the system evolves further, I plan to make the selection mechanism more sophisticated, and particularly to teach the computer enough about the morphology of Serbo-Croatian to enable it to overcome the vagaries of inflection and dialect. As it stands now, however, the basic heuristic algorithm is not language-dependent; thus HEURO-I can be easily adapted to any oral literature whose transliteration is machine-readable, and this flexibility is a great advantage at the beginning stages of establishing computerized editions.

At each of these two levels, then, HEURO-I seeks out analogs for traditional structures, first recognizing the pattern in the object text and then locating other occurrences in the data file. The file of reference texts is constructed by the reader before analysis starts : it may consist of any combination of texts, from the object text alone through the corpus of texts sung by a single *guslar* to a more

extensive sample drawn from the *guslar's* local tradition or even more widely. The textual referent for the heuristic routines can, in other words, be modified to suit the reader's or analyst's specific purposes. Since the reader has the option of viewing any or all of the thematic and story pattern correspondences, as well as of attenuating the sensitivity of the formulaic search, he can in effect teach himself the poetic tradition at the pace and in the depth necessary to the task at hand.

At present HEURO-I operates exclusively on approximately 10,000 lines of Serbo-Croatian oral epic, a group of eight texts chosen from among those I am preparing for a volume of *Serbo-Croatian Heroic Songs*. The system makes available to the uninitiated reader an edition of any one of these texts in a uniquely active context; according to the selection made for the data file, any of these works is automatically contextualized, presented against the background of its tradition. As the reader proceeds through each work, summoning analogs to the object text to reveal its multiform structure, he begins to revivify the performance record, to transmute text back into poem. For the analyst, HEURO-I provides a way to study multiformity at every level on a scale and with an accuracy otherwise impossible. The text-processor glosses any line or any passage by charting its traditional identity over the 10,000-line sample; as one learns more and more about a given traditional structure, he is more and more able to describe the basic processes of oral tradition itself. For the reader, HEURO-I constructs a uniquely faithful edition, recalling a traditional ambience forever lost in the printed text. For the analyst, it assembles a structural profile of the object text and the tradition of which that text is one momentary example.

A part from enlarging the data base for Serbo-Croatian oral epic and improving the sensitivity of automatic text-processing routines, I plan in the future to extend the capabilities of HEURO-I in interdisciplinary directions. Highest in order of priority is developing editions of selections from certain ancient Greek and Anglo-Saxon texts, works which are thought to derive from oral traditions. To be able to read the Homeric poems as the oral palimpsest they constitute, especially in the light of encouraging theoretical progress on oral traditional structure made by classicists in recent years, is, it seems to me, of the first importance. And the Old English epic *Beowulf*, whose formulaic and thematic patterns are well documented, should profit from the kind of edition that HEURO-I makes possible. Of course, the size of the textual sample and our limited knowledge about the actual performance and transmission of the Greek and Old English epics preclude assembling a text and tradition model as faithful as in the case of the Serbo-Croatian material, both these kinds of limitations will always be with us in dealing with dead-language traditions. At the very least, the construction of computerized editions promises to recover more of the poems involved than could the usual typographical recensions. In addition, I plan a series of such editions in modern English translation, so that students in courses like my Oral Literature seminar at the University of Missouri can have the opportunity to teach themselves poetic traditions whose languages they do not control. As I believe it should, HEURO-I thus commands a wide spectrum of possible uses, from original-language editions and analytical tools for scholars to pedagogically oriented editions in translation for students. At all points the fundamental premise remains the same: to recontextualize the individual text, to reinvest the poem with the richness of its oral tradition<sup>(11)</sup>.

## FOOTNOTES

- (1) For perspective on the rapidly evolving field of oral literature research, see my "Oral Literature : Premises and Problems," *Choice* (December 1980), 9-16. The discipline is founded in the pioneering work of Milman Parry, whose writings are reprinted or published from manuscript in Adam Parry, ed., *The Making of Homeric Verse : The Collected Papers of Milman Parry* (Oxford : Clarendon Press, 1971), hereafter cited as *MHV*; and of Albert Lord, whose writings through 1980 are summarized in my "Introduction : The Oral Theory in Context", in *Oral Traditional Literature : A Festschrift for Albert Bates Lord*, ed. John Miles Foley (Columbus : Slavica Press, 1980), pp. 32-51. Comparative extensions of oral theory, which now entail dozens of literatures, date from Lord's classic *The Singer of Tales* (1960; rpt. New York : Atheneum, 1968 et seq.), hereafter cited as *Singer*. For a partial bibliography of the field through 1972, see Edward R. Haymes, *A Bibliography of Studies Relating to Parry's and Lord's Oral Theory* (Cambridge, Mass. : Harvard University Printing Office, 1973); see further my *Oral Theory : An Annotated Bibliography and Introduction* (New York : Garland Publishers, forthcoming).
- (2) See Lord, "Songs and the Song", in *Singer*, pp. 99-123.
- (3) For the example of Old English formulaic diction, see the summary in Foley, *Oral Traditional Literature*, pp. 52-79.
- (4) The citations are taken from the computer-generated concordance of Old English verse : Jess B. Bessinger, Jr. and Philip H. Smith, ed. and progr., *A Concordance to the Anglo-Saxon Poetic Records* (Ithaca : Cornell University Press, 1978).
- (5) On the stereotypical scene, which occurs throughout an oral tradition with some adjustments to conform to the particular narrative environment, see Lord, *Singer*, 68-98.
- (6) See, e.g., Alain Renoir, "Oral-Formulaic Theme Survival : A Possible Instance in the *Nibelungenlied*", 65(1964), 70-74.
- (7) Matija Murko, *La Poésie populaire épique en Yougoslavie au début du XXe siècle* (Paris : Honoré Champion, 1929).
- (8) Milman Parry, Albert Lord, and David Bynum, colls., eds., trans., *Serbo-Croatian Heroic Songs (Srpskohrvatske junačke pjesme)*, vols. 1-4, and 14 (Cambridge, Mass. : Harvard University Press, 1953-80), with additional volumes in preparation, including a collection of the epic songs from stolac which constitute the data base for HEURO-I.
- (9) The following lines are quoted from Stolac texts in the Milman Parry Collection of Oral Literature at Harvard University (see note 8) with the permission of the curator. The editions of the texts are my own.

- (10) On story pattern morphology in oral epic, see my "The Traditional Structure of Ibro Basic's 'Alagić Alija and Velagić Selim', " *Slavic and East European Journal*, 22 (1978), 1-14; "The Oral Singer in Context : Halil Bajgorić, *Guslar*, " *Canadian-American Slavic Studies*, 12 (1978), 230-46; and " *Beowulf* and Traditional Narrative Song : The Potential and Limits of Comparison", in *Old English Literature in Context : Ten Essays*, ed. John D. Niles (London and Totowa : Boydell-Brewer and Rowman & Littlefield, 1980), pp. 117-36, 173-78.
- (11) For a fuller explanation of HEURO-I than is possible here, see my "Editing Oral Epic Texts : Theory and Practice, " *TEXT (Yearbook of the Society for Textual Scholarship)*, 1 (1981), forthcoming.