Manzoni and the Microcomputer: *I Promessi Sposi*: Chapter IX*

I INTRODUCTION

1. *Premise*

This is a lexical analysis of Alessandro Manzoni's IXth chapter of the *I Promessi Sposi*. It is intended as a testing application of LDMS (Linguistic Data Management System), a research tool for scholars in the Humanities developed on IBM S/5120 and soon to be available for PC models. Taking chapter IX as a case study experimentation of the adopted referencing methodology, this paper illustrates some of the data management techniques (and, by implication only, also the diversified types of content analysis made possible) introduced by LDMS in conformity with traditional scholarly criteria. The task of evaluating the performance of LEXED — the central LDMS's module — is necessitated, despite its largely recognized relevance to independent research, by the need for pragmatic demonstration. Its sound theoretical foundations need to be better defined from a practical viewpoint to avoid relying on intuitive reasoning, vague perception and undefined procedures. The definition of concrete stops on a model scale is obtained by outlining and by discussing from a user's perspective the functional parameters and the actual values produced.

Demonstrating the system's validity on a known assigned text by measuring its performance increases users' confidence in its operational simplicity and sophisticated feature combination. The use of this system presupposes no prior knowledge in either hardware or computing languages and discredits irrelevant claims on the need to become programmers or acquire yet another specialization. Assuming previous proficiency for any aspect of this application amounts to a misunderstanding of the reasons that inspired this system.

2. *Scope of the Study*

Since one can hardly fail to notice the flurry of intellectual activity that currently surrounds the understanding of computer applications, this article addresses three interrelated areas of interest to Italianists in particular: 1. the purpose and features of LDMS as a research tool; 2. the actual process of text management including its input, stream filing techniques, and coding hierarchy (in conjunction with this, observations will be offered on the system's reliability, underlying scholarly criteria and morphic table creation); and finally, 3. the use of resulting information and a brief illustration of each compiled major reference work.

Part two stresses also the concept of programming modification as necessitated by Manzoni's prose habits; the constitution of specific sets of morphic tables versus a metatable or supertable; the degree of obtained coding discrimination with and without author intervention; and, the need for more exhaustive documentation of the "real" rather than the "formal" language actually used by an author in a certain work at a certain time period. The Word-Processing (WP) module (TXTED) and Data Base (DB) compilation are not directly discussed. Also, for space reasons Referencing and Information Retrieval are barely treated to allow greater illustration of analytical tasks and a summary interpretation of the linguistic data contained in ch. IX of *I Promessi Sposi*. This practical exploration reflects on Referencing, Computer Literacy and other traditional practices such as normalization in manuscript editing. However, it does so in general terms only as needed for a more cohesive justification of the LDMS methodology, its discriminate referencing and the advantages it offers in preparing analytically arranged information for consultation.

There are numerous reasons for conducting a computer analysis of Manzoni's "L'urtar che fece la barca contro la proda." Ch. IX was selected from among the 38 chapters of this novel because it has a fairly typical style of prose which mixes narrative with dialog. It is not excessively long, and it has an acceptable location (approximately at one fourth the total length). The text is the one edited by L. Russo and published in 1957.¹

Two basic reasons for this analysis were Manzoni's bicentennial birthday celebration,² and the general need to speak of (and perhaps clarify) the important contribution of modern technology to language and content analyses as traditionally understood by scholars. Thus, it became almost imperative for me to work — though on a limited scale and in terms of a practical demonstra-

tion/application — toward a real assessment and experimentation of LDMS, the computerized analytic system originally designed for medieval legal texts, in the event of its wider utilization with modern Italian and other languages.³ Specifically, I was curious to establish to what extent I would have to modify the already compiled morphic tables to accelerate the process of automated morpho-semantic analysis. I emphasize that I am not a "literature man" and, much less, a specialist on Manzoni. Mine is a technical interest in an admired author as an *utente di lingua*⁴ — and a master at that — whose vocabulary in the 1840 edition of the *Promessi Sposi* embodies the "concretezza, acutezza" and "umorismo" of Florentine language. This language is all the more significant because it lacks the local coloring contamination and regional features which would have turned Manzoni's writing into another provincial — and thus unfashionable — product headed for rejection. Though Manzoni's acclaimed linguistic experience, genuinely national and yet so close to French and Milanese, does show promise of yielding interesting comparative results in the end (thus adding to the critical acclaim he has received for focusing on the limitations of the Italian language), our reflections will only expand on a single chapter of his volume in terms of a concrete exemplification of a general system's lexical discrimination and coding ability. It amounts to establishing what modifications are suggested, needed or required by automated microcomputing analysis. What conflicts in other words, can modern Italian present for a system originally implemented for medieval and specialized languages? Will radical changes be required for its working elements or will integration of complementary information be sufficient? Were the critics fully right? The intention here is to illustrate the kind of data management and results that a microcomputer system such as LDMS can provide to scholars involved in the study of Manzoni or other modern and medieval writers. As a practical application, this presentation first conveys a general understanding of this research tool for humanists by outlining its general characteristics, operational modes and operational advantages as well as by listing the standard and special output forms it produces. The literary and linguistic criteria applied to ensure the system's adherence to the basic principles of traditional scholarship will also be pointed out.

The core of the presentation will introduce and subsequently discuss the concept of "morphic tables." Attention will be focused on the working principle underlying these tables and the features

which enable them to enhance the understanding of "automated" morpho-semantic and lexical analysis. Limitations of the type of analyses performed by LDMS are discussed to underline the degree of morphosemantic discrimination achieved. Concurrent content analysis features and techniques are also documented. Using ch. IX of the *Promessi Sposi* the work and modifications required in the compilation of "general," "specialized" (for a specific time period or particular literary classification), and "super or macro" morphic tables (for universal use) are illustrated to define the tasks and the purpose of "automated" literary or non-literary text analysis. Primary and accessory applications are also more realistically defined. Additionally, the overall purpose is to afford scholars greater familiarization with machines which are revolutionizing their research by improving both its quantity and quality while preserving its traditional principles and human involvement.

3. LEXED: A Tool Helping to Link Scholarship and Technology⁵

Fundamental to this presentation is a general understanding of LEXED. Briefly, it is a general lexico-morphic analytic tool designed to aid traditional scholarship by minimizing time-consuming tasks in selecting, organizing and managing linguistic data, while maximizing a researcher's concentration on the analysis of collected data.

LEXED — or XT-LEXED for the upcoming version implemented for PC models (IBM or IMB-compatible) — is a program package which is suitable for both quantitative text analysis and for data management in qualitative text analysis. My claim is that purely quantitative investigation is at best "approximate" in natural language processing. It still deserves to remain an option, but, as an innovative approach in need of refinement, it should be subordinated to qualitative analysis. All programs, for editing or text exploration, are designed as user-friendly to assist or to lead step-by-step — as the case may be — in the application process.

Recently the focus on the computer as a mere object of study has shifted to the computer as a versatile tool. This is mainly due to the fact that microcomputers are now flooding campuses, and there is a need to focus on the innovative uses of the micro's real power for graphics, communication and WP. Yet, research methodology is also an aspect of great interest to those who want to know how technological advances are affecting the evolution of higher education. LEXED is a research tool devised to accommo-

date scholars, with or without computer background, in utilizing this powerful new tool effectively in their own operations. The XT-LEXED version, supported by an IBM-Equipment Grant represents an attempt to accelerate the availability of a marketing version of this interactive Linguistic Data Management System (LDMS). This underscores the recognized academic value and potential for colleges, universities and, above all, individual scholars engaging in independent research.

Most of the consideration received by the system as one of the leading national projects in literary applications is due to the attention commanded by the innovations it introduced. Comparatively speaking, its multiple features and combinations produced a tool best suited for humanistic research with remarkable anticipation of similar projects, none of which fulfil as many functions.

Referencing, DB Compilation and Management, Analytic Tasks, WP, and Information Retrieval (including idea or concept processing) are the distinctive functions integrated for author and printer interface in various languages. This kind of multi-tasking and sophisticated capability enhances the actual market success of microcomputers. However, the majority of researchers/users have yet to enjoy their full potential and relate to them mostly as a WP or a teaching tool (mostly for tutorials at this time). Time will soon take care of this, since academic computing makes more sense now. Skepticism about the coming Information Age and about "Computer Literacy," which is admittedly a concept in a fluid state, may be warranted. However, a "doom and gloom" outlook, such as that taken in Noble's "The Underside of Computer Literacy," 1984,⁶ is going a bit too far. Noble begins his article as a reasoned and refreshing study on the perils that might result from the hasty infusion of microtechnology into homes, work places, and schools. Unfortunately, his analysis soon degenerates into a paranoid conspiracy theory — an extreme pitfall to be avoided as much as micro-hype.

II TEXTUAL INPUT AND OUTPUT (MANZONI'S *Promessi Sposi*, Ch. IX)

1. *Narrative and Input*

To grasp what the system can do for you, we must start with a basic consideration such as what do you do when you would like to use it. Knowing where to start always helps.

Chapter IX in Russo's edition consists of 650 lines of prose in 24 pages (each page is only half to two-thirds full because of the commentary) which translate into 16 LEXED files/pages of 50-51 lines including blank lines. This page reduction from 24 to 16 is attributable simply to the elimination of the notes. More revealing is the chapter's transformation into 140 pages of Stream File (SF), a vertical rendition of the text which we discuss later, but which represents the basis of most of the actual output formats (*Fig. 1*).

Manzoni's narrative is a type of prose that differs obviously from poetry but also from, say, legal prose (analyzed by the original system). The latter requires short laws that provide quite a satisfactory 'natural' referencing breakdown needed to produce a fast and easy consultation system. Long prose, however, does not have such a convenient breakdown, and the text reproduction calls for subsequent "artificial" division where reference must use either the page or line. Line references should be avoided, not so much for their being a superfluous indication, as for later operational problems they cause. For instance, discovery of a single line of omitted text can be a real adjustment headache when working with a fixed number of lines per file in a page-oriented system. Also, insertion of page indicators at text entering time is simpler (requiring one input versus one per line). And verification is not actually encumbered by this simplification except, perhaps, when extraordinarily long paragraphs are present. One example of structural change affecting the run or processing of data, may exemplify the importance of "consistency" between input and programming. A [9.164] indicates that data entered below prior to the next reference variation indicator ([9.165]) belong to ch. IX, page 164. Indicators of text are by convention set to start on file-line seven. Errors in keying the text into computer memory are a frequent occurrence, and missing a phrase or two can cause inadvertent insertion of additional lines during revision. If any insertion pushes the first text line or indicator upward, the program will either blow without executing the job or will apply a wrong cross-reference. Thus operational limitations i.e. conditions, must be kept to a minimum in programming for ease of operation.

On the other hand, a forgotten page indicator discovered at the end of the memorization process can be corrected with relative ease. If our text has 27 files and p. 183 were left out (the indicator only that is, not the actual data), proper insertion can be executed in about 10 minutes without retyping over 200 lines of text! A correct procedure to follow entails the shifting of three lines of text

for files 27 through 24 (starting on 27). Removal of the bottom three lines occurs in files 26-24 (and are added in 27). Still, sizeable omissions in text-input are one drawback in a page-oriented system.

From an operational point of view, then, precision is a basic requirement both in the traditional sense to ensure reliability of data and from a technical point of view to allow execution. In the specific case quoted earlier, failure to start our first file on line seven produced the Alpha-Index with wrong reference for the words of page 164 (the first page in this case). Their incorrect "0.000" reference appeared in the Stream File (Fig. 1). But processing was made impossible at the next stage, an error message resulted and there was no Formarium execution.

2. The Stream Files (SF)

A stream file or vertical rendition of the text accompanied by various reference fields (needed for the reconstruction of the original) is machine controlled. Its reliability is, of course, dependent upon the accuracy of the input and revisions. It may be pointed out that even when one feels there are no more misspelled words, some will show up in all likelihood. Inadvertent agglutinations are the most common error. And, while misspelled items are easy to correct, unseparated words require remaking of the entire SF. For this reason it is highly recommended that careful revision(s) take place before the SF run (Fig. 1).

Actually, two such SFs are compiled — CNC or concordance SF and IDX or index SF. The latter is a modified version of the CNC stripped of all punctuation. Diacritics often cause problems given the tremendous variability of writers' prose habits. Two such instances in Manzoni's writing are the use of double hyphens and suspension dots. Manzoni interrupts a character's speech and, between his/her remarks, he interjects his own comments using hyphens as delimiters of his character's continued sentence. Suspended utterances are also new in the sense that they did not occur in medieval legal texts. Without program changes, this causes the appearance of incorrect entries and offset counts.

The SF stages are three. The first is an uncoded-SF which does not have to be printed except for error-free verification. The second is the automatically coded SF, in which each word presents a letter tag (N,O,V) or combination thereof. The last is the author revised-SF or final version to be used for output. In this phase, any N,O,V followed by a pound (#) sign should be carefully re-

viewed since together they represent a "default" value. The latter is decided by the author; in this case it is always an N value. These nominal forms are tagged only after all other priority tagging has been satisfied.

3. *SFs, Morphic Table Coding and Lemmatization*

Priority tagging refers to the running order of the programs which identify the parts of speech. Previously compiled tables (grammatical(Os), verbal(Vs), nominal(Ns)) are integrated in sequenced runs. Any sequential ordering reflects the implied hierarchy imposed by a language system with univocal instances enjoying top priority. Frequency of occurrence is another principle used in establishing discriminating criteria. More rarely, ingenuity may provide a perfect solution (O.F. *a* [prep] versus *a* [verb] using the Italian letter *à*⁷ which is not available in the three so-called French alphabet variations on the IBM S/5120). And, of course, the last applied criterion is the "default" principle. All conflicts still extant must then be solved by author intervention, which explains the interfacing characteristic so fundamental to the system in ensuring both maximum flexibility and accuracy.

Author alteration (intervention) is complementary in nature and purpose, since it must be understood that morphic table compilation and application are not meant to be exhaustive. The aim is not to get a 100% correct performance. Though possible — and prohibitively expensive — it would result in a text-dependent set of tables that would be too "text-specific." The program usefulness would be limited and prospects for implementing it on varied types of writing would be poor. Rather than shooting for a 98% score on a single text, it would seem wiser to settle for an 85% to 90%, but have a table set that is applicable to largely different prose types.

The constitution of the morphic tables thus entails an underlying operational philosophy. To accommodate personal taste, it is now possible to provide the system with a choice typology. A composite set of three tables, or four at the most, with the last being a "meta-table" or "super-table," would insure the possibility of trial runs before the final choice. (Depending on textual length, I would first make sure I had a competent assistant; it's a job I wouldn't wish on anybody!).

At this point, there is no space to discuss further the preceding observations, but some reflections on lemmatization and SF-editing (revision) are useful. Traditional lemmatization is too

"democratic." Linguistically, it annihilates verbal individuality for the sake of a formal system analysis and ends up saying too little. It may inform me that Manzoni uses *accennare* five times in ch. IX when in fact he does not. Not once does such a word-form occur in ch. IX. Even granting that the presence of *accennando*, *accennarono*, *accennati*, *accennava* & *accennò* is no revelation in itself, we can be confident that these words together are a stylistic indication of an author's narrative form(s). Moreover, we are not interested in documenting directly a formal language system. Rather, we want to identify the "real" one as used by an author in a specific time and genre.

A scientifically complete description of a formal language system is hardly based on a few major authors' output. It rather is the end result of a thorough search of language use by centuries (or cross-century patterns) yet to be achieved in Italian or — for that matter — other languages. Condensing or lemmatizing under the infinitive is another example of arbitrary elimination in research which increases subjectivity rather than objectivity and collects what I call "mutilated" Data Bases (you still do not know whether that DB actually reflects its author or to what extent it does/does not). The text must be respected and preserved in much the same way a manuscript is. "Normalization, integration, recensio, emendatio etc." should not be used as a crutch. Rather, it should be the instrument for consistently eliminating ambiguity where it really exists. Oscillations are not ambiguous and are now a measurable entity. In terms of our morphic coding, the "real" language system analysis respects this fairly rigid, self-imposed criterion, making separate word form entries for any existing clitic combination. Thus for ch. IX, *accordarlo*, *accorgersi*, *accostarsi*, *accompany*, *aiutarci* and *alletterla* are among the first 100 formal entries elevated to individuality.

Coding achieves automatic morpho-semantic discrimination not only for *condotta(V)/condotta(N)*, *faccia/faccia*, or *fra(N)[frate]/fra(O)*, but also for *colti* where the phonetic opposition *o/ɔ* of closed and open vowels is taken care without subscripts since the nominal N-code resolves the conflict. *Batter* is another case for N/V use resolution. It is also interesting because of the pound sign(#) juxtaposed after the N-code to indicate its "defaulted" nature. As such its lexical value reflects the nominalizing use of infinitives. Considering Manzoni's frequent use of apocopated forms, this "default" instance is not acceptable in principle and was modified since many more are the actual occurrences of apoco-

pated infinitives requiring a V-code. The use of a pound sign as a global value of defaulted instances can become a measure of not "overdone" tables.

Gran is also presently a default value, but it is scheduled to go into the nominal table or the grammatical one. This is to avoid its mistaken identification as a verbal entity. The frequency of apocopated verb forms ending in *-an* and *-on* makes it subject to ambiguity once the new rule is implemented. It will be quite interesting to see whether this rule change will have a positive impact throughout the entire novel in as much as ch. IX seemed to justify its adoption. (I would hate to see the degree of its validity applied to Venetian documents).

4. Auto-coding Revision and Morphic Table Expansion

As for the author's manual revision of the SF, any correction introduced in the coding fields is automatically applied to the inactive SF(CNC or IDX). Using a slow S/5120 and assuming entry-by-entry revision (brought to the screen in batches of 11), I reviewed 700 entries the first hours and about 900 the second hour for the first 1,600 records. The actual average at the end of the job was right around 1,000 words/hr. It is not unlikely that the full chapter revision (7,266 words), which at the indicated speed could take from eight to ten hours, could be reduced to one to two hours — if not less — by using the proper tables. (In the experiment the medieval tables were adopted).

To illustrate more precisely how Manzoni's word forms differ from XV century Italian, I am listing below, either by table or in a narrative form, the additional forms of the required expansion beginning with the grammatical forms (Table I):

TABLE I. Grammatical Forms Expansion

<i>a'[ai], c', n', m', t', v'</i>
<i>ai, io, ah, oh, ih, eh</i>
<i>ne', de', col, cui, due, ivi, fin, già, tre, sur, lui, lei,</i>
<i>sin, noi, voi, chi, sul</i>
<i>anzi, com', male, ott', otto, ogn', gran, pari, nove, poco</i>
<i>que', solo [taken out: fine, vero]</i>
<i>anch', ancor, tant', cert', dalla, dalle, dallo, dagli,</i>
<i>dall', dagl', fuori, entro, negl', mezzo, [di] nuovo, nulla</i>
<i>ognun, punto [affatto], quasi, senz', tutt', sette, sugli,</i>
<i>sugl', sulla, sulle, sullo</i>
<i>appena, almeno, allora, adesso, cinque, costui, costei,</i>

*grand', dietro, dentro, laggiù, mentre, nessun, spesso, ovvero, quand', quando, quest', quegli, quindi, perciò, talora, troppo, [a] stento, vicino
altrove, appunto, addosso, costoro, giacché, davvero, giammai, lontano, indosso, intanto, nemmeno, insieme, neanche, neppure, quaggiù, quattro
ciascuna, ciascuno, ciascuno, ciascuna, talvolta, affinché, soltanto, indietro, tuttavia, qualcosa, allorché, malgrado, qualunque, alquanto, quantunque, piuttosto. [in] succinto, [in] disparte, qualcheduna, qualcheduno*

Among these forms, Manzoni's love for apocopation is clear. For the rest, most noticeable is the appearance of modern subject pronouns and forms obviously not used in medieval times (XIV-XV century).

In comparison, the variations in the N-tables and the V-tables are almost proportionately inverse. By increasing Ns, one can introduce a good many generalizing V-rules using endings as matching criteria. Generalization carries, however, a shortcoming in the sense that adding nominal forms to the tables tends to make these quite large. It must be pointed out also that the verbal, nominal, default value, and compound-words tables are run by a single program. So the priority sequence must be exact and devoid of contradictions. On the other hand, rule repetition — again, when not contradictory — does not prevent correct executions from being carried out as desired. For instance, if it has been established that all accented endings such as *-à, -í, and -ò* are to be labeled as Vs, all abstract words such as *età, città, pietà, vanità, verità, ansietà, volontà, facilità, necessità, formalità, familiarità, familiarità, immobilità, probabilità* and *superiorità*, actually occurring in ch. IX, must be part of the nominal table or face incorrect tagging. Noticeable is the lack of accented words ending in *-í* and *-ò* from ch. IX.

As can be inferred from the examples used, the programs' structure is such that it allows for separate or joint full word analysis (distinguished in letter verbs or letter nouns) or for analysis of parts of words such as endings (distinguished in letter endings-verbs, letter endings-nouns and letter endings-compound words VO). All labels are arbitrary and chosen for logical convenience. On the basis of what has been said, it can also be surmized that a majority of the N-table's elements are terms presenting some kind of conflict with the sweeping generalizations of the V-table. A

typical case is represented by words ending in *-no* (3ppl/s). Both regular (*cappuccino, destino, inchino*) and diminutive forms (*principino, cervellino*) are affected. Of course, nouns and adjectives ending in *-are, -ere, -ire* (*genere, affare, camere, altare, libere, povere, tenere, [il] piacere, maniere, [l'] avvenire, carattere, chiacchiere, faccendiere, esemplare, alveare*) must be tabled. Similarly, if *-ssa, -sse, -sso* and *-ssi* endings are generally selected as verb forms, irregular indicatives (*scrisse*) and regular imperfect subjunctive forms, can still be handled correctly if all nouns and adjectives like *rosso, [le] mosse, spasso, fisso, eccesso, fattoressa, badessa, interesse* are properly tabled. Gerunds require that particular attention be given to nominal forms like *iracondo, mondo, benda, comando, vicenda, and reverenda* to avoid a V-coding execution. Present participles, on the other hand, are identified by their letter endings (*-ante, -anti, -ente, -enti*) as nouns. A default value in their case would not work because the participial V-endings would pick them up as such; the default value is the last value applied by the hierarchy. The intricacy of rule overlap and overtake is demanding but easily mastered through experience (a misnomer at times for frustration).

III OUTPUT CONSULTATION AND DATA INTERPRETATION

Having established the direct relationship between input and output, its delicate nature can be seen at the moment of general information retrieval and "observation" analysis (an appropriate nickname I chose since by this time you cannot change results!). Also, on-line retrieval of information can no longer be altered now, short of going back to the SF stage or the text itself.

A normal question which often arises at this point, and also at the time interest is first generated in a system, is "what is the benefit for the time and work involved?" I answer that this way: besides the 500ca. printed pages of actual output⁸ from 24 original pages of text amounting to 640 lines of prose, and besides the accuracy and objectivity achieved, there is the truly new "dynamic nature" of the data base destined for future use, refinement and eventual network diffusion.

1. Indices

For users, the thrill comes from seeing the processed data come to life. Its interrelated nature makes it quite easy to proceed with verification and cross-referencing for either comparative or investigative purposes. Even the much maligned Alpha Index can be an immediate and valuable source of information. It is, in fact, a

good tool of establishing the most distinctive words used by Manzoni in ch. IX.⁹ In 24 short pages, 2,300 word forms are morphically and semantically discriminated.¹⁰ A major complaint is the necessity of keying in the text. Yet, in many cases, forms like *gastigo* or *soffogata* do strike the author's curiosity and prompt him or her to assess how many times such initial and internal sonorizations occur. It can be quickly determined that we have also *distrigarsi* and *nudrita*, all with a count of one except for *gastigo* with three. These cases are clearly in the minority, but they still deserve attention if we are to grasp the meaning of the survival of these and other phenomena that cannot be treated here.

Other noticeable preferences over more modern forms are some variants such as *divozione*, *enimma*, *rispingerne*, *scandololi*, *tristalmente*; the use of prosthetic *i* in *iscorgendo* and in *isgarbatezze*; his attention to technical terms in feminine clothes in *soggolo* and *scollo*; and *coonestare*, which I would choose as the distinguishing literary term featured in this chapter.¹¹

His use of apocopated forms which enhance his literary tone stimulates the use of the Alpha Index. The frequencies of use with a good many verbs and a few nouns produce an off-setting result of ch. IX as shown (Table II):

TABLE II: *Apocope**Equivalence**Full Form*

<i>furon/furono</i>	3/1		<i>cuor/cuore</i>	1/2
<i>far/fare</i>	16/9(+ 1N)		<i>dir/dire</i>	1/13
<i>dar/dare</i>	3/1		<i>eran/erano</i>	6/9
<i>chieder/chiedere</i>	2/0			
<i>esser/essere</i>	9/7			
<i>aver/avere</i>	7/1			
<i>voler/volere</i>	1/0			
<i>poter/potere</i>	1/0			
		<i>saper/sapere</i>	1/1	
<i>ben/bene</i>	13/7			
Prepositional preference shows:				
<i>col/con</i> 5/82 and <i>sur/sul</i> 2/4				

What the Alpha Index also indicates is the consolidation level which actually reveals the number of word forms used compared to the total number of words: 2,300 versus 7,266 (31.6% reduction). Numbers and percentages will have to be analyzed chapter by chapter and overall to arrive at any meaningful determination.

Finally, for ch. IX we have only compiled three Word Indices (WI): WI-Alpha Index; WI-Decreasing Frequency or Numeric Index; and, WI-TopoOnomastic Index (Fig. 2-4). When fully implemented the LEXED system can produce eleven additional indices: WI-R (Reverse or Rhyme I.); WI-O (Functional Words); WI-N (Nominal); WI-V (Verbal); WI-(O+N); WI-(O+V); WI-(N+V); WI-(L+P); WI-L (Toponomastic); WI-P (Onomastic); WI-(N+L+P) [other possibilities should be regarded as purely mathematical]. Also, using the Glossary SF-field and the selection SF-field, another set of indices could be produced based on the desired numbers of specialized dictionaries and coded-selections.¹²

2. *Morphoformarium*

What this output adds to the basic type of indices is the cross-reference of each work form. It is also viewed as a cheaper replacement of expensive concordances since it carries location and frequencies telling exactly what the author has/has not used. It is built on the same morphosemantic discriminating criteria of the concordance (Fig. 5).

3. *Concordance*

It adds contexts to each occurrence and often makes a direct verification of the text unnecessary except for cases where the quoted contexts are insufficient (Fig. 6).¹³ Each line is 180 characters long with 48 before a word. Each word has 27 spaces reserved with seven for its reference and morphocode. Maximum length of word and trailing context is 125.

IV CONCLUSION

In the 1st analysis it can be said that LEXED helps scholars produce and interpret a complete set of reference works while reserving the memorized data for multiple manipulations as well as for future incrementation. The microcomputer can serve as the basis for more objective content analysis and, where Manzoni is concerned, may reveal interesting elements which help clarify some of the most debated aspects of this writer's character, which is now receiving renewed attention as his bicentennial approaches. Apparent contradictions in his works may prove he was a man more sensitive to life than to mere principles.

FIG. 1

Stream File of Chapter IX

Phase 3: Edited. Records 51-100

CODED LEX, STRM, CNC FILE								
REC £	MOR	VAR	FLG	CON	DCT	SEL	WORD	TXT.PAR X
51	O						tre	0.000
52	V						resero	0.000
53	N						tristamente	0.000
54	N						grazie	0.000
55	O						al	0.000
56	N						barcaiolo,	0.000
57	O						"Di	0.000
58	O						che	0.000
59	N						cosa?"	0.000
60	V						rispose	0.000
61	O						quello:	0.000
62	V						"siam	0.000
63	O						quaggiú	0.000
64	O						per.	0.000
65	V	0					aiutarci	0.000
66	O						l'	0.000
67	O						uno	0.000
68	O						con	0.000
69	O						l'	0.000
70	O						altro,"	0.000
71	O						e	0.000
72	V						ritirò	0.000
73	O						la	0.000
74	N						mano,	0.000
75	O						quasi	0.000
76	O						con	0.000
77	N	≠					ribrezzo,	0.000
78	O						come	0.000
79	O						se	0.000
80	O						gli	0.000
81	V						fosse	0.000
82	V						proposto	0.000
83	O						di	0.000
84	V						rubare,	0.000
85	O						allorché	0.000

86	N	⤴	Renzo	0.000
87	V		cercò	0.000
88	O		di	0.000
89	V	0	farvi	0.000
90	V		sdrucchiolare	0.000
91	O		una	0.000
92	N		parte	0.000
93	O		de'	0.000
94	N	≠	quattrinelli	0.000
95	O		che	0.000
96	O		si	0.000
97	V		trovava	0.000
98	O		indosso,	0.000
99	O		e	0.000
100	O		che	0.000

FIG. 2

Alpha Index

Morpho-Codes & Frequencies

WORD FREQUENCY INDEX — ALPHABETICAL SEQUENCE

*** (N ≠)	1	accattar (V)	1
-curiosa (N ≠)	1	accennando (V)	1
-e (O)	1	accennarono (V)	1
-gran (N ≠)	1	accennati. . . (V)	1
-o (O)	1	accennava (V)	1
a (O)	130	accennò (V)	2
a' (O)	1	accettaron (V)	1
abbandono (N)	2	accettata (V)	1
abbassò (V)	1	accetti (V)	1
abbattimento (N)	1	accoglieva (V)	1
abbia (V)	3	accommiatò (V)	1
abbiam (V)	3	accomodato (V)	1
abbiamo (V)	2	accompagnata (N ≠)	1
abbondante (N)	1	accompagnati (N ≠)	1
abbondio (N)	1	accompagnava (V)	1
abisso (N)	1	accompagnò (V)	2
abitare (V)	1	acconsentirò (V)	1
abituale (N ≠)	2	accordarlo (V)	1
abituamente (N)	1	accordate (V)	1
accade (V)	1	accordo (N ≠)	1

accorgerà (V)	1	agnese (N)	10
accorgersi (VO)	1	ai (O)	2
accorgevan (V)	1	aiutarci (VO)	1
accostarsi (VO)	1	aiuto (N ≠)	3
accostatevi (VO)	1	al (O)	38
accozzando (V)	1	alcune (O)	6
accresciuta (V)	1	alcuni (O)	2
acqua (N ≠)	1	alla (O)	13
acquietò (V)	1	alle (O)	25
ad (O)	7	alle (C)	13
adamo (N)	1	allettarla (VO)	1
addosso (O)	1	allevato (V)	1
addussero (V)	1	allogate (N ≠)	1
adesca (V)	1	allora (O)	12
adesso (O)	1	allorché (O)	1
adorna (V)	1	almeno (O)	4
affacciassero (V)	1	alquanto (O)	3
affacciavano (V)	1	alquanto (O ≠)	1
affare (N)	2	alta (N)	1
affermativo (N ≠)	1	altare (N)	1
affetto (N ≠)	3	alterato (N ≠)	1
affettuoso (N ≠)	1	altero (N ≠)	1
affinché (O)	1	alti (N)	1
afflitta (N ≠)	1	alto (N)	2
afflizione (N ≠)	1	altra (O)	9
aggiunse (V)	1	altre (O)	14
agio (N ≠)	1	altrettanto (O)	1
agli (O)	1	altri (O)	13

FIG. 3

Numeric (Decreasing) Index

WORD FREQUENCY INDEX — NUMERIC SEQUENCE

e (O)	266	non (O)	104
di (O)	218	una (O)	96
che (O)	214	con (O)	82
la (O)	156	per (O)	81
a (O)	130	d' (O)	74
il (O)	123	le (O)	73
in (O)	108	si (O)	68
un (O)	108	era (V)	60

l' (O)	54	un' (O)	17
più (O)	54	far (V)	16
ma (O)	51	gli (O)	16
della (O)	49	avrebbe (V)	15
da (O)	46	disse (V)	15
del (O)	42	questo (O)	15
sua (O)	42	altre (O)	14
come (O)	41	lucia (N)	14
al (O)	38	né (O)	14
quel (O)	38	all' (O)	13
i (O)	33	alle (O)	13
se (O)	30	altri (O)	13
gertrude (N)	27	ben (O)	13
è (V)	26	cosa (N)	13
alla (O)	25	dell' (O)	13
quella (O)	25	dire (V)	13
signora (N ≠)	25	due (O)	13
aveva (V)	24	fu (V)	13
nel (O)	23	madre (N ≠)	13
s' (O)	23	monaca (N ≠)	13
tanto (O)	23	nella (O)	13
guardiano (N)	22	poi (O)	13
quando (O)	22	allora (O)	12
questa (O)	22	chi (O)	12
lei (O)	21	così (O)	12
loro (O)	21	delle (O)	12
monastero (N)	21	già (O)	12
suo (O)	21	giovine (N ≠)	12
de' (O)	20	ha (V)	12
fosse (V)	20	senza (O)	12
lo (O)	20	sue (O)	12
qualche (O)	20	tra (O)	12
quale (O)	20	altro (O)	11
anche (O)	18	ch' (O)	11
o (O)	18	ci (O)	11
padre (N ≠)	18	dopo (O)	11
quelle (O)	18	essa (O)	11
ciò (O)	17	lui (O)	11
donne (N ≠)	17	poco (O)	11
ogni (O)	17	poteva (V)	11

FIG. 4

Topo-Onomastic Indices

Non-Condensed Type (ch. IX)

ONOMASTIC INDEX — ALPHABETICAL SEQUENCE

Adamo; (9.167)	Gertrude (9.183)
Agnese, (9.164)	Gertrude, (9.184)
Agnese, (9.166)	Gertrude, (9.185)
Agnese (9.167)	Gertrude, (9.185)
Agnese, (9.167)	Gertrude (9.185)
Agnese (9.168)	Gertrude (9.186)
Agnese; (9.170)	Gertrude (9.186)
Agnese, (9.171)	Gertrude (9.186)
Agnese (9.172)	Gertrudina (9.176)
Agnese, (9.173)	Gertrudina, (9.178)
Agnese (9.173)	Lucia, (9.164)
barocciaio (9.167)	Lucia (9.164)
Cristoforo (9.167)	Lucia (9.166)
Cristoforo, (9.173)	Lucia, (9.167)
Cristoforo (9.173)	Lucia (9.167)
Cristoforo, (9.174)	Lucia, (9.168)
Don Abbondio, (9.164)	Lucia, (9.170)
fra Cristoforo (9.164)	Lucia (9.170)
fra Cristoforo!" (9.167)	Lucia, (9.171)
Gertrude, (9.175)	Lucia, (9.171)
Gertrude (9.176)	Lucia (9.171)
Gertrude, (9.177)	Lucia (9.172)
Gertrude (9.178)	Lucia, (9.172)
Gertrude (9.178)	Lucia, (9.173)
Gertrude (9.179)	padre guardiano; (9.167)
Gertrude, (9.180)	padre guardiano (9.167)
Gertrude (9.180)	padre guardiano, (9.167)
Gertrude (9.181)	padre Cristoforo, (9.168)
Gertrude (9.181)	padre guardiano (9.170)
Gertrude (9.181)	padre Cristoforo, (9.172)
Gertrude, (9.181)	Renzo (9.164)
Gertrude (9.181)	Renzo (9.164)
Gertrude (9.182)	Renzo (9.165)
Gertrude, (9.182)	Renzo (9.166)
Gertrude, (9.183)	Renzo (9.166)
Gertrude (9.183)	vicario delle monache, (9.180)
Gertrude (9.183)	
Gertrude, (9.183)	

TOPONOMASTIC INDEX — ALPHABETICAL SEQUENCE

convento de' cappuccini, (9.166)	Monza, (9.166)
Lambro; (9.165)	Monza (9.168)
Milano, (9.168)	Monza: (9.177)
monastero della signora, (9.167)	Monza, (9.185)
Monza (9.165)	Spagna, (9.167)
Monza, (9.165)	

FIG. 5

Morpho-Formarium

Codes, Absolute Frequency

Cross-Reference, Relative Frequency

* * * FORMARIO GENERALE * * *

*** (N ≠) [1] 9.174

-curiosa (N ≠) [1] 9.173

-e (O) [1] 9.182

-gran (N ≠) [1] 9.,173

-o (O) [1] 9.182

a (O) [130] 9.164, 9.165(4), 9.166(10), 9.167(3), 9.168(8), 9.169(3),
 9.170(4), 9.171(8), 9.172(6), 9.173(8), 9.174(9),
 9.175, 9.176(4), 9.177(3), 9.178(4), 9.179(4),
 9.180(5), 9.181(4), 9.182(3), 9.183(8), 9.184(6),
 9.185(15), 9.186(6), 9.187(3)

a' (O) [1] 9.170

abbandono (N) [2] 9.170, 9.182

abbassò (V) [1] 9.170

abbattimento (N) [1] 9.187

abbia (V) [3] 9.167, 9.173(2)

abbiam (V) [3] 9.170, 9.174, 9.180

abbiamo (V) [2] 9.166, 9.176

abbondante (N) [1] 9.166

abbondio (N) [1] 9.164

abisso (N) [1] 9.186

abitare (V) [1] 9.178

abituale (N ≠) [2] 9.177, 9.186

abitudinalmente (N) [1] 9.176

accade (V) [1] 9.182

accattar (V) [1] 9.181

accennando (V) [1] 9.168

accennarono (V) [1] 9.167

Concordanza — Capitolo IX de *I Promessi Sposi*
 Discriminazione: Morfosemantica
 Rimandi: Capitolo e pagina

CONCORDANZA GENERALE

nne dopo. Fra essa l' ultima figlia del principe	***, (N #)	gran gentiluomo milanese, che poteva contarsi tra i più doviziosi	9.174
questa signoral- pensava tra sé, per la strada;	-curiosa (N #)	davvero! Ma chi la sa prendere per il suo verso, le fa far ci	9.173
il suo piano. -O mi vorranno forzare,- pensava,	-e (O)	io starò dura; sarò umile, rispettosa, ma non acconsentirò: non si t	9.182
la lettera di ragguaglio all'amico Cristoforo,	-Gran (N #)	cervellino che è questa signoral- pensava tra sé, per la strada:	9.173
ure, e fatto, com' ora si direbbe, il suo piano,	-O (O)	mi vorranno forzare,- pensava, -e io starò dura; sarò umile, rispett	9.182
primogenito, destinato a conservar la famiglia,	a (O)		9.174
ità. Non che tutte le monache fossero congiurate	a (O)		9.177
stento le sue, e, stringendo forte forte la mano	a (O)	Agnese, disse con voce soffogata: "A rivederci", e partì. Le donne si	9.166
o: quando sarai madre badessa, allora comanderai	a (O)	bacchetta, farai alto e basso". Qualche altra volta il principe, ripr	9.176
e si fermavano a tormentarlo più distintamente e	a (O)	bell' agio. Che poteva mai esser quella punizione minacciata in enimm	9.185
ato, senza tante faccende, ha condotto l' affare	a (O)	buon porto, in un batter d' occhio. Sarà contento quel buon	9.173
er un di que' dispetti della sua guardiana, andò	a (O)	cacciarsi in un angolo della camera, e lì, con la faccia nascosta tra	9.186
re e signora illustrissima," disse il guardiano,	a (O)	capo basso, e con la mano al petto: "questa è quella povera giovine,	9.170
farle sentire la sua suggestione. Di rado, e solo	a (O)	certe ore stabilite, era ammessa alla compagnia de' parenti e del pri	9.182
natamente pronta a tutto ciò che potesse piacere	a (O)	chi doveva doveva	9.187
desse nelle camere della fattorezza; e andò solo	a (O)	chieder la grazia. Dopo qualche tempo, ricomparve giulivo, a dir loro	9.168
caro delle monache, o da qualche altro deputato	a (O)	ciò, affinché fosse certo che ci andava di sua libera scelta: e quest	9.180
tal supplica. E a fine d' indurla più facilmente	a (O)	ciò, non mancaron di dirle e di ripeterle, che finalmente era una mer	9.180
più buona di loro; piangerò, pregherò, li moverò	a (O)	compassione: finalmente non pretendo altro che di non esser sacrific	9.182
re intatta la sostanza al primogenito, destinato	a (O)	conservar la famiglia, a	9.174
ovine contadina inesperta, non pensava più tanto	a (O)	contenersi; e i suoi discorsi divennero a poco a poco così strani, ch	9.174
un soggolo, che si stendeva alquanto sul petto,	a (O)	coprire lo scollo d' un nero saio. Ma quella fronte si raggrinzava sp	9.169
di questo dice ch' era un borgo antico e nobile,	a (O)	cui di città non mancava altro che il nome; dice altrove, che ci pass	9.165
ttore tanto impreveduto, tanto diverso da quello	a (O)	cui eran destinate; si figurava che avesser potuto cader sotto gli oc	9.185

- a parlatorio, guardò in giro dove fosse la signora
 a lle sue compagne, la più franca, e pronta sempre
 a bbe più dolorosa, ch' egli potrebbe venir presto
 a partir di nascosto dal suo paese, per sottrarsi
 a grazia. Dopo qualche tempo, ricomparve giulivo,
 a informato di voi, in quest' affare. Tocca a voi
 a alle maniere della carceriera, la quale (spesso,
 a n facevano specie alle due donne, non esercitate
 a è la figlia e la madre non sapevan più come fare
 a o in quando, i pensieri della religione venivano
 a ivere sconosciuta, e dove nessuno ardisca venire
 a narcì più di rado, a rispingerne la rimembranza,
 a a nascosta tra le mani, stette per qualche tempo
 a ito e del misterioso che abbiām veduto in lei, e
 a ib. VI, Cap. III, pag. 358 et seq.) che ha avuto
 a rità di metterci al sicuro, giacché s'iam ridotte
 a trascrivere e sottoscrivere una tal supplica. E
 a forza, di modo che la povertà è stata ridotta
 a son molto rari, la consigliera fece pagar questo
 a sempre a dar consigli risoluti. Questa suggerì
 a ti i grandi del mondo si servono dei doni di Dio
 a tirò le mani, anche lui, e, come fuggendo, corse
 a duto in piedi, il guardiano si fermò, e si voltò
 a ntimenti di ben diverso genere contribuivan pure
 a "Accostatevi, quella giovine," disse la signora
 a me de' loro figliuoli!" Agnese mortificata diede
 a padre guardiano," aggiunse subito, rivolgendosi
 a lui, con una compietezza studiata. "Anzi," continuò, "ci ho già pensat
- (O) cui fare il suo inchino, e, non iscorgendo persona, stava come incant
 (O) dar consigli risoluti. Questa suggerì a Gertrude d' informar con una
 (O) dar nuove e sentime; tanto che si risolvette di partire. Si concerta
 (O) de' gravi pericoli; e ha bisogno, per qualche tempo, d' un asilo nel
 (O) dir loro che venissero avanti con lui; ed era ora, perchè la figlie e
 (O) dirci se questo cavaliere era un persecutore odioso". In quanto all'
 (O) dire il vero, provocata da lei) si vendicava ora facendole paura di q
 (O) distinguer monaca da monaca: e il padre guardiano che non vedeva la s
 (O) distrigarsi dall' interrogazioni pressanti della fattressa. Attraver
 (O) disturbare quelle feste brillanti e faticose. Ma la religione, come l
 (O) disturbarla, quand' anche. . . "Quali pericoli?" Interruppe la signor
 (O) divezzarsene. Né più a lungo, o più volentieri, si fermava in quelle
 (O) divorar la sua rabbia. Sentì allora un bisogno prepotente di vedere a
 (O) far comprendere i motivi della sua condotta, in quello che avvenne do
 (O) far menzione di quella persona medesima, non nomina, è vero, né lei,
 (O) far questa
 (O) fine d' indurla più facilmente a ciò, non mancaron di dirle e di ripe
 (O) fuggir da casa sua". "Accostatevi, quella giovine", disse la signora
 (O) Gertrude, con tante beffe sulla sua dappocaggine. La lettera fu conce
 (O) Gertrude d' informar con una lettera il padre della sua nuova risoluz
 (O) gloria sua, e in vantaggio del prossimo, come vossignoria illustrissi
 (O) governare la sua bestia. Dopo una sera quale l' abbiamo descritta, e
 (O) guardar se gli altri venivano; quindi entrò e s' avviò al monastero;
 (O) intervallì a scemare quella sua antica avversione: talvolta il rimor
 (O) Lucia, facendole cenno col dito. "So che il padre guardiano è la boc
 (O) Lucia un' occhiata che voleva dire: vedi quel che mi tocca, per esser
 (O) lui, con una compietezza studiata. "Anzi," continuò, "ci ho già pensat

9.168

9.181

9.166

9.171

9.168

9.171

9.168

9.186

9.170

9.168

9.180

9.171

9.185

9.186

9.174

9.165

9.172

9.180

9.171

9.181

9.181

9.171

9.166

9.168

9.186

9.171

9.172

9.173

NOTES

- * The original paper was prepared for a public lecture held on St. Michael's campus of the University of Toronto, co-sponsored by St. Michael's College and the Department of Italian Studies of the University of Toronto.
- 1 A. Manzoni, *I Promessi Sposi*, ed. L. Russo (Firenze: La Nuova Italia, 1957). The edition happened to be the one most conveniently available. The "testing" purpose of this essay did not call for critical evaluation of later editions.
 - 2 I am thinking primarily of the Spring 1985 Congress in Lecco, Italy.
 - 3 Latin and Old French (Poems) are the languages for which LDMS was implemented on IBM S/5120, besides Medieval and Modern Italian.
 - 4 G. Devoto, *Il linguaggio d'Italia* (Milano: Rizzoli, 1974), p. 304.
 - 5 This is a summary description. For a fuller documentation of LEXED see L.F. Farina, "A Linguistic Data Management System," in *Computers and the Humanities*, 17, 3 (1983), published in Spring 1984.
 - 6 Douglas Noble, "The Underside of Computer Literacy," in *Raritan*, 3 (Spring 1984), 37-64. Also on computer literacy concept, see L.F. Farina, "Lessico e analisi lessicomorfica," *Acts*, 17th International Congress of Romance Philology, University of Aix-en-Provence, September 1983.
 - 7 Common homographic instance encountered in Old French manuscripts such as "La vie sainte Merguerite." In general LEXED handles quite well polysemic instances, though not exhaustively.
 - 8 Actual output amounts to four Indices, the Formarium and the Concordance. Working output like the SFs are not counted or the total would double. Also, the 450 pp. for the three described stages can be reduced by at least one-third.
 - 9 "Uncommon, unexpected or strange forms" if preferred.
 - 10 The Index (2,300 forms in this case) represents a substantial reduction compared to the 7,266 corresponding occurrences. A whole new discourse can be opened up here in reference to a verification process of all actual changes introduced by Manzoni in his 1840 edition. The ample bibliography on this subject still needs technical verification.
 - 11 This literary term implies a justification by means of false argumentation presented either as honest, legitimate reasoning or with the pretext of one's authority and prestige.
 - 12 Three separate fields are presently available for the handling of 1. variants; 2. selected items for inclusion in a Glossary/Dictionary; and, 3. syntactic or conceptual string extrapolation for syntactical problem analysis as well as for searches and studies of "figurative" language comprising the traditional figures of speech known as synecdoche, metonymy, hyperbole, personification, irony, metaphor, simile etc. (which must all be identified and selected before arguing on their taxonomic structure). For discussions on taxonomy see J.G. Carbonelle, "Metaphor — A Key to extensible semantic analysis," *Proceedings of the XVIIIth Annual Meeting of the Association for Computational Linguistic and Parasession on Topics in Interactive Discourse* (Philadelphia: University of Pennsylvania, 1980); D. Gentner, "Are Scientific Analyses Metaphors?" In D.S. Miall ed., *Metaphor: Problems and Perspectives* (Brighton, Sussex: Harvester Press, 1981); A. Ortoni, "Beyond Literal Similarity," *Psychological Review*, 86,3 (1979), 161-79; id., "The Role of Similarity in Similes and Metaphors," in A. Ortoni ed., *Metaphor and Thought* (Cambridge-New York: Cambridge University Press, 1979); A. Ortoni, R.E. Reynolds, J.A. Arter, "Metaphor: Theoretical and Empirical Research," *Psychological Bulletin* (Sept. 1978), 919-43; A. Ortoni, D.L. Schallert, R.E. Reynolds, S.J. Anthos, "Interpreting Metaphors and Idioms: Some Effects of Context on Comprehension," *The Journal of Verbal Learning and Verbal Behaviour*, 17 (1978), 465-77.
 - 13 On-Line Information Retrieval possibilities are documented in L.F. Farina (Aix-en-Provence, 1983).