ELS Conference on Natural-Language Applications

September 10-12, 1985

IBM Danmark A/S, 2800 Lyngby, Auditorium

under the Patronage of

IBM Denmark's General Management

E.O. Lippmann, Chairman
"Computer-assisted Philology: Towards a Unified Edition of OSp. 'Libro de Alexandre'".

This paper shows the different problems involved in the edition of an OSp. text from two manuscripts and some fragments, written in different centuries and dialects, and how many of those problems can be solved with the aid of the computer.

The OSp. Libro de Alexandre was written c. 1230. It is a fairly long poem of more of ten thousand lines, which makes of it the longest of the OSp. poems up to the fourteenth century. The "original" and early copies are lost, whilst only two long (but non-complete) and late copies have been preserved. The first one, manuscript O, in Madrid. This is a fourteenth century copy with a strong western dialectal (Leonese) characterization. The second one (ms. P, B.N. Paris) longer than the other, but also incomplete, was written in the fifteenth century by a copyist who used eastern dialectal features (Aragonese). Four short fragments were preserved, too, which add little to our knowledge of the text.

Against the general idea of the enormous differences existing between the two manuscripts, the adequate treatment with the computer shows the possibility of arriving to an internally coherent unified edition. To reach the first stages I employed a number of editor macros and a short PL/I program; a method which could be extended to similar texts of OSp. poetry. Further treatment of the text till the complete unified edition requires a philological task which benefits from other computer advantages.

Francisco Marcos-Marin
Universidad Autonoma de Madrid
IBM Scientific Centers: Madrid and Heidelberg
Research Fellow of the Alexander von Humboldt Stiftung (1984-85)
Thursday, September 12, 1985

8:45 - 9:00  Coffee, Hall

9:00 - 9:45  **Phonemic Script: A Central Vehicle for Natural Language Processing - The Case of Hebrew**  
             Professor Uzzi Ornan, The Hebrew University, Jerusalem, Israel

               Professor Francisco Marcos Marin, University of Madrid, Madrid, Spain

10:30 - 11:00  Break

11:00 - 11:45  **The Use of Morphology in Automatic Language Processing with Examples from Swedish**  
                Professor Gunnel Källgren, University of Stockholm, Stockholm, Sweden

12:00 - 13:00  Luncheon

13:00 - 13:45  **Norwegian Language-Processing Research/Some Problems Connected with Varieties of the Latin Alphabet in Text Processing**  
                Dr. Jan Engh/Professor Even Hovdhaugen, University of Oslo, Oslo, Norway

13:45 - 14:30  **Controlled Natural Language Interface: The Best of Three Worlds**  
                Dr. Eva-Maria Mückstein, T.J. Watson Research Center, Yorktown Heights, U.S.A.

14:30 - 14:45  Break

14:45 - 15:30  **Text-Processing and Office Systems Applications in Practice**  
                Mr. Daniel Wurmsner, IBM Europe Headquarters Paris, France

15:30 - 16:00  **Highlights and Wrapup**  
                Edward Lippmann, European Language Services, IBM Denmark
PARTICIPANTS

Professor Bente Maegaard  Denmark
Dr. Kimmo Koskenniemi  Finland
Mr. Seppo Koskenniemi  Finland
Professor Maurice Gross  France
Dr. Bernard Meritaldo  France
Dr. Isabelle Warnesson  France
Dr. Herb Budd  France
Mr. Daniel Wurmser  France
Dr. Vittore Casarosa  France
Dr. Wilhelm Ott  Germany
Dr. Michael Wegner  Germany
Dr. Brigitte Barnett  Germany
Professor Uzzi Ornan  Israel
Dr. Michal Ephratt  Israel
Dr. Nicoletta Calzolari  Italy
Dr. Remo Raffaelli  Italy
Dr. Samer Attasi  Kuwait
Dr. Botrous T. Thalouth  Kuwait
Professor Gerard Kempen  Netherlands
Mr. Douwe van Essen  Netherlands
Professor Even Hovdhaugen  Norway
Dr. Jan Engh  Norway
Professor Francisco Marcos Marin  Spain
Dr. Ramon Casajuana Casanovas  Spain
Dr. Luis de Sopena Pastor  Spain
Professor Gunnell Källgren  Sweden
Ms. Karin Mårtensson  Sweden
Dr. Stephen Richardson  U.S.A.
Dr. Eva-Maria Mückstein  U.S.A.

Ass. Gen. Mgr. Jørn Meldgaard  IBM Denmark
Mr. Bent Rosenkrands  IBM Denmark

Mr. Edward Lippmann  ELS
Mr. Walter P. Molden  ELS
Mr. Peter Moldt  ELS
Mr. Søren Brandt  ELS
Mr. Svend Gjerløv  ELS
Ms. Jytte Mouritzen  ELS
ELS Conference on Natural-Language Applications
September 10-12, 1985
IBM Danmark A/S, 2800 Lyngby, Auditorium
under the Patronage of
IBM Denmark's General Management
Proceedings
E.O. Lippmann, Chairman

Professor
Francisco
Marcos Marin

European Language Services
IBM

Francisco Marcos-Marín
Universidad Autónoma de Madrid
Centro Científico de IBM en Madrid
IBM Wissenschaftliches Zentrum Heidelberg
August 9, 1985

The *Libro de Alexandre* is a thirteenth-century Spanish learned epic poem. Its subject is the life and exploits of Alexander the Great, who was king of Macedonia, at the north-eastern part of Greece, in the IV c. b. C. He defeated Darius, king of Persia, and Porus, king of India, and founded an enormous empire, from Egypt, where Alexandria was named after him, to the Indian subcontinent. His kingdom vanished at his death, but the cultural bridges between Eastern Asian cultures and Greece were in so short a time, deeply rooted.

The Spanish poem is composed of slightly over 2600 stanzas of approximately fourteen-syllable monorhymed quatrains called, in Spanish, *tetraestrosos monorrimos* or *cuadernas via*. This type of poetry is also called *mester de clerecia* as it was "clerk work", a learned sort of poetry employed by educated people, who usually belonged to the clergy, i.e. the authors were actually monks.

The aim of our study is to provide the public with a unified edition of the two extant manuscripts and the four little fragments preserved up to date.

The "original" manuscript, or whichever copy that truly deserved that name, is lost. There are, however, two more or less complete copies, not of that "original" but of other intermediate copies. It means that the two long extant manuscripts, albeit survivors of a long tradition, have a poor quality, plus a large number of differences.

The *Osuna* manuscript (ms. O) is held at the Biblioteca Nacional de Madrid. It is a fourteenth-century manuscript on parchment, written in a language which shows a strong western (Leonese) influence. It consists of 2510 stanzas.

The *Paris* manuscript (ms. P) belongs to the Bibliotheque Nationale de Paris. It has been dated to the early fifteenth century and is written on paper. It is comprised of 2639 quatrains. Its most noteworthy feature is its strong eastern (Aragonese) dialectal character.

There are, therefore, some *lacunae*, which can be filled up with text of the other manuscript. Those excerpts which are only extant in one manuscript are left as they are in our edition.

In addition to the two more or less complete manuscripts, there are four fragments. No one contains text which lacks in the longer manuscripts, so they don’t fill any gap. Nevertheless, they are useful to solve some questions.

The manual comparison of more than 10,700 lines is a tremendous task, where the scholar gets easily lost. A computer assisted edition offers also both a real possibility and a new perspective.

Francisco Marcos Marin
Universidad Autonoma de Madrid (Spain)
IBM Wissenschaftliches Zentrum Heidelberg
Centro Científico de IBM en Madrid

ALEXANDRE MS. PARIS

3 3 Segu\textless n\textgreater t que yo entiendo quij lo qujsier saber aura de my solas en cabo muy' grañt plazer aprendera buenas gestas q<ue> sepa Retraher auerlohan por ello mucho's a coñoçer

ALEXANDRE MS. MADRID (OSUNA)

3 3 Quij oyr lo q<ui>sier / a todo mjo creer aura de mj solaø / en cabo g<ra>nt plazer aprendera bonas gestas / q<ue> sepa retraer auer. loan por ello / muchos a co<n>nosçer
ALEXANDRE P WITH UNIFIED WRITING (ALEJO-P)

3 A SEGUNT QUE IO ENTIENDO QUI LO QUISIER SAUER
   B AURA DE MI SOLAS EN CAUO MUI GRANT PLASER
   C APRENDRA UUENAS GESTAS QUE SEPA RETRAER
   D AUERLOAN POR ELO MUCHOS A CONOSER

ALEXANDRE O WITH UNIFIED WRITING (ALEJO-O)

3 A QUI OIR LO QUISIER A TODO MIO CRER
   B AURA DE MI SOLAS EN CAUO GRANT PLASER
   C APRENDRA UONAS GESTAS QUE SEPA RETRAER
   D AUER LOAN POR ELO MUCHOS A CONOSER
ALEXANDRE COMÚN (VERSION 2)

(from the Paris manuscript)

3 a qui lo quisier
b Aura de mi solas en cauo grant plaser
c Aprendra gestas que sepa retraer
d por elo muchos a

ALEXANDRE COMÚN (VERSION 3)

(from the Paris manuscript)

(After splitting of clitics)

3 a qui lo quisier
b Aura de mi solas en cauo grant plaser
c Aprendra gestas que sepa retraer
d Auer lo an por elo muchos a conoser
ALEXANDRE COMÚN (HAND INTERVENTION)

3  a  qui lo quisier

b  aura de mi solaz en cabo  grant plazer

c  aprendera  b  nas gestas que sepa retraer

d  auer-lo-an por ello muchos a conocer

PARTIAL PROVISIONAL RECONSTRUCTION (ESSAY)

3  a  Qui  oir lo quisier  a  todo mio  creer
[O is followed, P in note]

b  Aura de mi solaz en cabo grant plazer
[P: muy grant]

c  Aprendra  buenas  gestas que sepa retraer
[SCANFILE, Med: buenas]

d  Aüer-lo-an por ello muchos a conocer

Alexandre