American Journal of Computational Linguistics Microfiche 9 : 1

NEWSLETTER OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS

VOLUME 11 - NUMBER 3 DECEMBER 1974

This issue was released for production on March 25, 1975. The editor intends to distribute American Journal of Computational Linguistics in four packs per calendar year, promptly at intervals of three months. Each pack is to contain two numbers of the Finite String. The difficulties of the first year of publication of AJCL are responsible for the delayed production of this pack, which also contains Volume 11, Number 4 of TFS. It would be a rash editor indeed who guaranteed promptness without caveat. The present editor must warn the subscriber that the end of the difficulties is not yet fixed for a date certain.

AMERICAN JQURNAL OF COMPUTATIONAL LINGUISTICS is published by the Center for Applied Linguistics for the Association for Computational Linguistics.

EDITOR: *David G. Hays, Professor of Linguistics and of Computer Science, State University of New York, Buffalo.*

EDITORIAL STAFF: *Brian Phillips, Assistant; Jacquin Brendle, Secretary.*

EDITORIAL ADDRESS: *Twin Willows, Wanakah, New York 14075.*

MANAGING EDITOR: *A. Hood Roberts, Deputy Director, Center for Applied Linguistics.*

ASSISTANT: *Nancy Jokovich.*

PRODUCTION AND SUBSCRIPTION ADDRESS: *1611 North Kent Street, Arlington, Virginia 22209.*

Table of Contents

INTERNATIONAL CONFERENCE - COLING 76 3
WORKSHOP - theoretical issues in natural language processing 4
SUMMER SCHOOL - computational linguistics at Rocquencourt . . 10
SUMMER SCHOOL - literary statistics at Cambridge, England . . 13
INTERNATIONAL CONFERENCE - computers and the humanities . . . 14
nfais CONFERENCE - information interfaces 19
IEEE CONFERENCE - computers to reach the people 21
IFIP CONFERENCE - computers in education 22
INTERNATIONAL CONFERENCE - developing countries 23
NSF DEPUTY DIRECTOR - Richard C. Atkinson nominated 24
NEH CALENDAR - grant applications 25
NATIONAL PROGRAM DRAFT - library and information services . . 26
RESEARCH PROGRESS - Index Thomisticus in press 26
COMPUTER SECURITY - AFIPS and ACM publish guides 27
OPINION: A restricted sublanguage approach to high quality translation - *Victor Raskin* 28
TECHNIQUE: Letters with variable values and the mechanical inflection of Rumanian words - *Minerva Bocsa* 38
ACL: Secretary-Treasurer's report 53
Editor's report 54

C O L I N G 7 6

INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS

UNIVERSITY OF OTTAWA
JUNE 28 - JULY 2, 1976

The conference originally planned for September 1975 has been rescheduled in order to avoid conflict with meetings on artificial intelligence (the date of that meeting was changed after the CL date was first set) and on applied linguistics.

The general plan is as described in ACJL; Card-6.

The coordinator is Dr. Guy Rondeau.

A request for further information should contain Title, name, and surname; Post held; Department; Institution; Postal address and zip code; and Field of interest.

Address: COLING 76, Linguistics, University of Ottawa, K1N 6N5.

American Journal of Computational Linguistics Microfiche 9 : 4

INTERDISCIPLINARY WORKSHOP

THEORETICAL ISSUES IN NATURAL LANGUAGE PROCESSING

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 10 - 13, 1975

Sponsored by the Association for Computational Linguistics
Supported by the Mathematical Social Science Board with funds granted by the National Science Foundation

DIRECTORS: BONNIE NASH-WEBBER AND ROGER SCHANK

The announcement and program appear on Card 25. The following account is presented as a statement of the needs and condition of the field; it was prepared by BNW and RS.

The workshop is to promote interaction among researchers in computational linguistics, psychology, linguistics, and artificial intelligence. The primary purpose of the workshop is to explore two major areas of common research, memory and knowledge, both with respect to their relation to human language behavior. Specifically, we hope to consider such questions as:

(1) What computational models and mechanisms have been proposed up to now in these areas?
(2) What aspects of human language behavior are they meant to account for?
(3) Are these models and mechanisms compatible?
(4) Is there a single global view of language understanding and use that is adequately modelled by some combination of them?

(5) Are there still significant aspects of human language use which they cannot account for?

(6) What is the best model of human language use that can be assembled out of the concepts that nave been developed in computational linguistics, linguistics, psychology, and artificial intelligence?

(7) How well does it really approximate what humans do with language?

(8) With respect to gaps in the model, is there anything currently in the wind adequate to complete them?

Thus our primary purpose is both to provide an assessment of our knowledge and goals in this area with respect to language processing and to provide a cross-disciplinary tutorial for the participants. A secondary purpose of the workshop is to discuss the important issue of valid methodology in such research.

In eight sessions, speakers will address the audience on points made by themselves or others in position papers previously distributed to all participants. After these presentations are finished, speakers and audience will take part in general discussions.

MOTIVATION AND GOALS

In recent years, researchers in several disciplines have been converging on the problem of language understanding and memory as providing a handle on the problems in their own fields. Researchers in Artificial Intelligence, concerned with building models of intelligent behavior, have started to develop and study models of conversational interaction, which naturally rely on models of language understanding. This is often done without reference to the work of researchers in other disciplines.

Linguistics, of course, has always been concerned with language, but frequently at no level higher than that of the sentence. Such theories as derived from the consideration of single

sentences out of context are difficult to apply to the task of understanding. Recently some linguists such as Chafe have shifted their emphasis towards understanding and modelling discourse, which is of direct relevance to questions of understanding.

Psychologists have also been interested in problems of language understanding and memory. While early research did look to linguistics to p-ovide plausible theories of human language behavior the then-current theories proved insufficient to the task. As a result, some psychologists have begun to create their own theories, while others have begun to pay attention to the work of AI researchers. Computer models are now appearing which are explications of these theories.

In the field of computational linguistics, the challenge of building computer systems which can carry on fluent and helpful dialogues with a user has also shifted the emphasis in the field from more efficient parsers to more capable understanders.

In spite of this convergence of many different disciplines on the same problem, there is surprisingly little communication between researchers in the different fields, apart from occasional discovery of relevant papers in one field by members of another. Without conferences of the sort that we are proposing, there is no opportunity for the individual researchers in these different fields to talk to each other directly. The value of such face-to-face confrontation and the opportunity for asking questions and exploring the applicability of techniques in areas other than the ones that the author had in mind are well appreciated. Within disciplines such opportunities exist in traditional professional conferences. Our conference will provide an opportunity for such interaction across a diversity of fields which would not otherwise be possible.

The need for such a conference is especially great in the area of natural language understanding. The many different fields which are beginning to give strong attention to this problem all have different interests and consequently different emphases on the problem. Researchers in a particular field tend to focus only on their own interests and ignore other aspects of the problem. For example, the classical transformational grammar theory has largely ignored the necessity for the theory to account for psycholinguistic and other performance aspects of language. This is a reasonable way to gain a restricted research problem, but the result of such research may suffer if the researcher does not have some general idea of the problems associated with the aspects that are being ignored. This conference will attempt to provide specialists in different fields with this type of general understanding of the problems of concern in other fields. We feel that this exposure to different aspects and emphases will have a very beneficial effect on all fields of natural language research, and that without such interchange the potential for much of that research will not be realized.

The opportunity for such a meeting as we are preparing does not currently exist elsewhere. Conferences sponsored by professional societies invariably present intradisciplinary views rather than interdisciplinary ones, and past interdisciplinary workshops have always been on a very small scale. For example, at the NSF sponsored workshop in Woods Hole, Massachusetts, in 1971, some 25 researchers interested in Computational Semantics were brought together for two weeks of worthwhile talks. While the results of such a conference are largely intangible, a number of participants in that conference have attested to the impact of that conference on their way of thinking about problems and the course of their research. We plan to adhere to the model of the Woods Hole conference, but with the following two changes.

First it is important to create the possibility for the many new researchers from different fields who have entered this area to gain an appreciation of the different emphases of other fields. Secondly, it is important that a much larger number of people whose interests are in one discipline or another be exposed to the ideas emerging from the synthesis of these disciplines. It is important that more than just a small group be able to exchange ideas.

Since a long workshop would be very difficult on the scale we are proposing, we are relying on an early circulation of position papers to familiarize all participants, speakers and audience alike, with the current ideas on natural language understanding in each of the fields. These preprints will also serve the valuable job of informing those unable to participate in the workshop of these current ideas, and we therefore intend to make copies of the preprints widely available.

ORGANIZATION

The two sessions of each day will be held in the morning and in the late afternoon. The long break for lunch in between will facilitate discussions of the morning's topic, without the need to get back to another session immediately.

All sessions will be open to the public, and we expect about 150 people to participate. MIT was selected as a site so that the widest group of interested people might be able to come at a reasonable cost. MIT has made a large air-conditioned lecture room available, and will also provide low-cost dormitory housing for the participants.

The sessions will not simply be introductions to working systems or well-known theories. Position papers will have been distributed to all participants at least a month in advance of

the workshop to familiarize them with the ideas of each speaker. The sessions will consist of short presentations (10-15 minutes) by the speakers outlining their already circulated ideas. Following a break, each speaker will have the opportunity to respond to earlier remarks, after which discussion will be opened to the audience. A session chairman will be responsible for maintaining the level and direction of the session.

SUMMER SCHOOL

COMPUTATIONAL LINGUISTICS

DOMAINE DE VOLUCEAU - ROCQUENCOURT
78150 LE CHESNAY - FRANCE

MAY 26--30, 1975

Sponsored by the Institut de Recherche d'Informatique et d'Automatique

DIRECTOR: M. ANDREEWSKY

FEE: 750 FRANCS

INTRODUCTION

Generalities about the methods, problems, and applications of computational linguistics. (Level of analysis of content, documentation, indexing, aids to diagnosis,programmed instruction in natural language, etc.) Is automatic analysis of language possible? Limits and possibilities of results. Possible applications, justification.

MATHEMATICAL METHODS IN LINGUISTICS

The essential components of the structure of language (French as the example. Usable strategies in the automatic analysis and production of text. Difficulties encountered. From language to algorithms to programming. Does language have a mathematical structure? How does it come out in French? Difficulties in the choice of methods of analysis and production. Adequacy relations between natural and programming languages.

LINGUISTIC ANALYSIS *(Les methodes d'apprentissage)*

Presentation of an operational discovery procedure which, beginning with a corpus analyzed grammatically, makes it possible to obtain automatically a syntax allowing disambiguation. It is impossible to foresee all the peculiarities of language. One must therefore arrange to integrate new linguistic data as they arrive; for that a discovery procedure is necessary.

AUTOMATIC DOCUMENTATION

A mathematical model of content analysis, used in automatic indexing and in interrogating documentation systems. Presentation of realizations. Automatic documentation is a privileged field of application for the most advanced methods. It is equally a particular viewpoint from which to see linguistics.

The first part of the course will be treated jointly by M. Andreewsky and M. Fluhr.

USE AND PROOF OF THEOREMS IN LINGUISTIC AND INFORMATIC APPLICATIONS

M. Pitrat, C.N.R.S.

AUTOMATIC CONTENT ANALYSIS OF SCIENTIFIC TEXT WRITTEN IN NATURAL LANGUAGE

M. Daniel Herault, Universite Pierre et Marie Curie

Discursive double articulation of scientific discourse: the hypersyntactic and hypersemantic components. Definition of semantic content. Role of the underlying derivational system; at the level of a text, in the realization of the double articulation: semantic units (predicates), principal modifiers, and associated syntactic structures. Informatic realization for the Slavic languages. Remarks on German, Romance, and Japanese. Elaboration of an advanced documentation system: integration of this research in an MT system.

AUTOMATIC TRANSLATION

M. Vauquois, GETA Grenoble

Automatic translation by syntactic analysis: 1960-1970. The process of translation as the step following source analysis, followed by transfer to the level of surface syntax. Notions of structural descriptors and their diverse representations in linguistic schools. Types of grammars and algorithms which permit automatic analysis or production. Evolution of different strategies aimed at new approaches to MT. Levels of transfer in MT: Surface syntax (Japanese experiments); transformational grammar (American experiments); pivot languages (Grenoble expériments; Mel'chuk-Zholkhowski theory); critique. Current research in MT: Practical work (machine aided translation--pre-editing, revising short and middle-term possibilities); long-term research (aspects of semantic calculus; experiments in man-machine communication in question-answering systems; influence of semantic research on automatic translation).

MAN-MACHINE DIALOGUES AND SPEECH

M. Gueguen, E.N.S.T.

Objective analysis of language: physical structure of the signal of language, classic methods of analysis, analysis by modeling (linear prediction, analysis by synthesis). Automatic recognition: system organization, acoustic preprocessing, levels of recognition, use of linguistic data; realizations and open problems in automatic comprehension of speech. Speech synthesis: devices and their commands (vocoders with channels, formants, simulation of the vocal tract); levels of synthesis; synthesis by rules. Perspectives and conclusions: the help of computational linguistics in the area.

ROUND TABLE: COMPUTATIONAL LINGUISTICS AND LINGUISTICS

SUMMER SCHOOL

L I T E R A R Y S T A T I S T I C S

CAMBRIDGE UNIVERSITY, ENGLAND

JULY 13 - 19, 1975

Sponsored by the Association for Literary and Linguistic Computing

ACADEMIC SECRETARY	M. H. T. Alford, Esq. 2, Sidgwick Avenue Cambridge, England
PRINCIPAL LECTURER	Norman Thomson IBM and Southampton University
GUEST LECTURERS	H. Sykes-Davies, Cambridge A. Q. Morton, Edinburgh Y. T. Radday, Haifa R. W. Bailey, Michigan K. W. Kemp, Cardiff
LECTURE TOPICS	Estimation and confidence intervals Design of experiments Analysis of variance
TUTORIAL TOPICS	Exercises based on the lectures Statistical validity of the work carried out Practical demonstrations
ONLINE FACILITIES	Available to students
INFORMATION	*Those who inform Mr. Alford of their hope of attending will receive further information. A tentative reservation of living space is suggested.*
FEES	Tuition L20 for nonmembers, L17 members Accommodation about L30; room and all meals, from dinner 7/13 through breakfast 7/19.

SECOND INTERNATIONAL CONFERENCE ON
COMPUTERS AND THE HUMANITIES
UNIVERSITY OF SOUTHERN CALIFORNIA
APRIL 3 - 6, 1975

ORGANIZING COMMITTEE

Robert Dilligan Rudolf Hirschmann	407 Founders Hall, USC, Los Angeles 90007
Joseph Raben	Queens College, CUNY
Donald Ross	University of Minnesota
Todd K. Bender	University of Wisconsin
Grace C. Hertlein	California State University, Chico

ADVISORY COMMITTEE

John R. Allen, Richard Bailey, Emmett Bedford, Roy Boggs, T. H. Howard-Hill, Winfried Lenders, Willy Martin, Joan Smith, John B. Smith, F. de Tollenaere, Stephen V. F. Waite, Stanley N. Werbow, Roy A. Wisbey, Antonio Zampolli

EDITORIAL NOTE

A tentative list of contributions appears on the following frames. It is too late to stimulate attendance, but it shows the scope of current activities and may suggest.further exchange of information.

TENTATIVE PROGRAM

MUSIC

Raymond Erickson Queens College, CUNY	*The DARMS project*
Bo Alphonce Yale University	*INTRIX: A scanner for pitch-class patterns in multipart music*
Gary Nelson Oberlin Conservatory	*A formalization of musical syntax*
Fred T. Hofstetter University of Delaware	*National differences and similarities in the use of melodic intervals during the mid-19th to early 20th centuries*
Jerome R. Wenker Sperry-UNIVAC	*On the analysis of musical analysis*
Lynn Trowbridge University of Illinois	*A computer processing system for Renaissance music*
Norbert Boker-Heil Staat. Inst. Musikforsch.	*A computerized indexing of Renaissance music*
Fred T. Hofstetter University of Delaware	*Development of a center for computational musicology*
Michael J. Ramey UCLA	*Computer application to the comparative study of musical instruments*

VISUAL ARTS

Luraine Tansey San Jose Comm. College	*Pre-columbian art: sites and chronology compu erized*
Eleanor Guralnick	*The proportions of Archaic Greek sculptured figures: a computer study*
Hiroshi Kawano	*Markov process theory of pictures*
Elizabeth M. Lewis West Point	*Computer coding for a micrographic index in art*
James E Gips & George N. Stiny UCLA	*Computer models for aesthetics*
Leonard Meyers Calif St.	*Computer animated film as visionary art*

COMPUTER-ASSISTED INSTRUCTION

Tej Bhatia University of Illinois	*New directions and issues in CAI*
Robert L. Oakman University of S. Carolina	*A videotape course for computer education in the humanities*
Peter Zoller Wichita State University	*A CAI approach to Black English*
Francine Ouellette University of Montreal	*JEUDEMO: A practical workshop*
George O'Brien University of Minnesota	*CAI: Siren songs and a skeptic*

DICTIONARIES AND CONCORDANCES

Sarah K. Burton Hunter University of Alabama	*Evolution of languages, Part I: Romance etymology*
Sidney Berger UC Davis	*Compiling a concordance*
Johannes B. Casser University of Montreal	*The Index Thomisticus: A test-case*
Robert Benson UCLA	*A proposed computer concordance of Medieval Latin*
Andrew T. Crosland U. S. Carolina, Spartanburg	*The concordance and the study of the novel*
Donald M. Lance University of Missouri	*The use of the computer in determining the geographical distribution of items*
Edward A. Kline University of Notre Dame	*Computer applications in Middle English dialectology*
Michael M. T. Henderson U. Wisconsin, Madison	*Use of an interactive program in analyzing data for a dialect dictionary*
Paul Bratley & Serge Lusignan University of Montreal	*Some problems and solutions in the edition of a dictionary*
Richard W. Bailey University of Michigan	*Inter-active lexicography: Some uses of* Michigan Early Modern English Materials

DATA BASES

Vincent J. Ryan
UCLA
Computerized concurrent indexing

Charles Dollar
Nat. Archives & Records
Scholars, computers, and the National Archives

LINGUISTICS

Gerard Salton
Cornell University
On the role of words and phrases in the automatic content analysis of texts

Annette Paquot-Maniet
University of Laval
Le vocabulaire caracteristique de l'avare chez Plaute et chez Moliere

Dirk Geens
AvTL
Automatic syntagmatic analysis of English

Patricia Lang
SWRL
L.A.P.: A system for processing text

Jean-Guy Meunier
U. Quebec, Montreal
A system for interactive text processing and content analysis

Edward R. Gammon
Calif. St. U., Fresno
Numerical taxonomy in linguistics

Burghard B. Rieger
Tech. Univ., Aachen
On a tolerance-topology model of natural language semantics

Robert A. Ariew
Pennsylvania State U.
Andre Breton's Poisson soluble: A computer-aided study

Jay Leavitt & John Lawrence Mitchell
University of Minnesota
Gap recurrence: A lexico-statistical measure

David Sankoff
University of Montreal
Correlates of speakers' word frequency parameters in a corpus of spoken French

Barron Brainerd
University of Toronto
On the distributions of articles and pronouns

TEXTUAL ANALYSIS

Robert Cannon
University of S. Carolina
An optional text collation algorithm

Todd K. Bender
U. of Wisconsin, Madison
A literary work conceived in positional notation

Giorgio Buccellati
UCLA
Computer aided analysis of Cuneiform texts

Eric Poole
University of Kent
The computer in textual collation and stemmatic analysis

STYLISTICS

Richard Williams Wichita State University	*Diction and social class in seventeenth century Spanish drama*
Colin E. Martindale University of Maine	*The Night Journey: Patterns of regressive imagery in journeys to Hell*
Tommy Joe Ray University of Mississippi	*Theme as style*
James Joyce UC Berkeley	*Computational model of stanzaic patterns in English*
Donald Ross University of Minnesota	*Keats' odes and sonnet--style and genre*
John Odmark Universität Regensburg	*Computers and stylistic analysis*
David H. Chisholm University of Arizona	*Phonological patterning in German verse*
Geoffrey J. D. E. Archbold University of Victoria	*Repetition, a characteristic of Ammianus Marcellinus' style*
Stephen Waite Dartmouth College	*Effects of genre and some stylometric features: evidence from Cicero's works*
Pierre Laurette Carleton University	*La petite liseuse do poem automate de lecture/réécriture*
Daniel L. Greenblatt University of Missouri	*Variable rules and literary style*

SEVENTEENTH ANNUAL CONFERENCE

NATIONAL FEDERATION OF ABSTRACTING AND INDEXING SERVICES

I N F O R M A T I O N I N T E R F A C E S

ARLINGTON, MARCH 4 - 5, 1975

CONFERENCE COMMITTEE

Ben H. Weil	Exxon Research and Engineering Co.
Joseph Coyne	National Technical Information Service
Ann Farren	BioSciences Information Service
A. Hood Roberts	Center for Applied Linguistics

P R O G R A M

BIBLIOGRAPHIC CONTROL

Ellis Mount Columbia University	*Bibliographic standards work--nationally*
Eric Clyde Canada Inst. S-T Information	*Bibliographic standards work--internationally*
Lawrence Livingston Council on Library Resources	*CONSER project*
Maureen LeFever BIOSIS	*BIOSIS/CAS/Ei bibliographic guide for authors and editors*

USER ASPECTS

Judy Wanger System Development Corp.	*Impact of on-line bibliographic services--a preliminary report*
Colin K. Mick Applied Communication Res.	*Impact of on-line search services on public library operations*
James L. Carmon University of Georgia	*Roles of intermediary and users in bibliographic retrieval systems*

REPACKAGING OF ABSTRACTS

Irving Zarember Amer. Petroleum Institute	*API/DERWENT "Patent Alerts"*
James Cape Energy R&D Administration	*ERDA Bibliographic Data Base*

Douglas A. Fisher
BIOSIS

HEEP

MILES CONRAD MEMORIAL LECTURE

Melvin S. Day
National Library of Medicine

Sharing--the hope of the seventies

DOCUMENT ACCESS

James L. Wood
Chemical Abstracts Service

nfais member services study report

Margaret H. Graham
Exxon Res. & Eng'g Co.

API-CAIS experimental metropolitan library service

Paul Zurkowski
Information Industry Assoc.

Business implications

Roger Summit
Lockheed Retrieval Service

On-line ordering of documents

COMPUTER TECHNOLOGY TO REACH THE PEOPLE

PROGRAM EXTRACTS

Martin L. Rubin HumRRD	*Lister Hill: A national CAI network*
Susan Wittig University of Texas Austin	*CAI in the composition classroom: some practical answers and some philosophical problems*
Kerry Mark Joëls Ames Research Center	*The megauniversity of Athens: A scenario for the future*
O. Firschein & R. K. Summat Lockheed Information Systems	*Computerized retrieval in a public library setting*

I F I P SECOND WORLD CONFERENCE
C O M P U T E R S I N E D U C A T I O N
MARSEILLES
SEPTEMBER 1 - 5, 1975

TO BE HELD UNDER THE HIGH PRESIDENCE OF THE

FRENCH MINISTRY OF EDUCATION

Under the patronage of UNESCO, OECD, and the Commission of European Communities; with the assistance of the Direction Générale de l'Industrie and the Intergovernmental Bureau for Informatics; in cooperation with the International Commission on Mathematical Instruction and the International Commission on Physics Education.

Organized by the Association Française pour la Cybernétique Economique et Technique

REGISTRATION

Registration is 500 F. until May 1; 600 F. thereafter. Write to AFCET - B.P. 571 - 75826 Paris CEDEX 17 for forms and details.

PROGRAM

Twenty papers have been invited; 740 contributions have been submitted to referees.

Information about the scope of the conference was published on AJCL Card 8.

CONFERENCE ON NATIONAL PLANNING FOR INFORMATICS IN

D E V E L O P I N G C O U N T P I E S

BAGHDAD

NOVEMBER 2 - 6, 1975

For full information:

Dr. H. A. Al-Bayati
Director General
National Computers Centre
P.O. Box 3261 - Saadoon
Baghdad, Iraq

Mr. A. A. M. Veenhuis
Intergovernmental Bureau
for Informatics
P.O. Box 10253
00144 Rome, Italy

NEW DEPUTY DIRECTOR FOR
NATIONAL SCIENCE FOUNDATION

Richard C. Atkinson, Stanford psychologist, is President Ford's nominee, according to a March 4, 1975, release.

Atkinson, a creative designer and user of mathematical models for memory, learning, and behavior, is assistant dean of the school of Humanities and Sciences at Stanford and chairman of its Psychology department.

A member of the National Academy of Sciences, National Academy of Education, and the American Academy of Arts and Sciences, Dr. Atkinson has written or edited ten books and more than a hundred professional papers. He was educated at Chicago and Indiana, and has taught at UCLA and Michigan.

NATIONAL ENDOWMENT FOR THE HUMANITIES

CALENDAR OF APPLICATION DEADLINES

SPRING AND EARLY SUMMER

1975

May 9 Research Grants Beginning after January 1, 1976

Simone Reagor, Division Director - 202-382-1072

Six programs: Research tools (dictionaries, bibliographies, guides, and catalogs). Research centers (major research collections with topical focus). International conferences for the Bicentennial (the deadline for this program is past). State and local history (scholarship and archives). Editing (of historical and literary papers of scholarly value). General research.

May 12 Fellowships For 1976-1977

James Blessing, Division Director - 202-382-1491

Independent study and research for scholars, teachers, writers, and other interpreters of the humanities who have produced or demonstrated promise of producing significant contributions to knowledge. Six months (to $10,000) or twelve months (to $20,000).

June 26 Public Programs Beginning after December 15, 1975

John Barcroft, Division Director - 202-382-1111

Museum personnel program. University or internship programs, seminars, or workshops to train interpreters.

July 1 Education Programs Beginning after January 1, 1976

Roger Rosenblatt, Division Director - 202-382-5891

Program grants for critical re-examination of the content, organization, and method of presentation of a group of related courses or an ordered program of study in the humanities. The central topic can be a region, culture, era, etc.; or a program can be defined by a curricular level. Limit, $180,000 in three years.

REVISED DRAFT

A National Program for Library and Information Services

National Commission on Libraries and Information Science
Suite 601, 1717 K Street NW Washington 20036

The commission expects to introduce draft Federal legislation during 1976. The draft reflects comments received in letters and obtained through regional hearings and the professional press.

INDEX THOMISTICUS

Roberto Busa, S.J.
Fondamente Nove 4885
30121 Venezia
Italy

Progress during 1974: 32 volumes of the Index Thomisticus, 23 volumes of the Concordantia Prima, and 9 volumes of the Indices Distributionis, making 36,000 pages in all, photocomposed. "Photocomposition time was 60 second per page: slow but perfect!" according to Father Busa. Ten volumes have been printed, bound, and published. Some 20 to 25 volumes are still to be prepared.

SECURITY

AFIPS SYSTEM REVIEW MANUAL

The first of a series on system improvement

Edited by Robert L. Patrick; based in part on a review of literature conducted by Mary Elizabeth Stevens.

The object is to specify methods of gathering data so that computing center managers, auditors, and system designers can assess their security needs. *$10.00 from AFIPS Press, 210 Summit Avenue, Montvale, New Jersey 07645.*

ACM EXECUTIVE GUIDE

A booklet for executives and managers--those to whom EDP managers report, has been prepared by the Institute for Computer Sciences and Technology of the National Bureau of Standards and the Association for Computing Machinery, with financial assistance from the National Science Foundation.

Why? *A man substituted deposit slips, magnetically coded with his account number, for the blank ones available on a bank's customer counter.....*

Who? *programmer, janitor, or even manager....*

Can data in a computer system be completely protected? No.

Terminals are the least secure points.

The booklet lists technical and managerial solutions to partially protect against these and other problems.

Dennis K. Branstad and Susan K. Reed
Systems and Software Division
Institute for Computer Sciences and Technology
National Bureau of Standards
Washington, D. C. 20234

American Journal of Computational Linguistics Microfiche 9: 28

O P I N I O N

A RESTRICTED SUBLANGUAGE APPROACH TO HIGH QUALITY TRANSLATION

Victor Raskin
Institute of Philosophy
Hebrew University
Israel

This paper deals with an approach to the problems of automatic high quality translation and, more generally, of automatic language data processing, based on the restriction of the input of MT and other systems to a certain type of sublanguage. The approach was proposed by the present author in the framework of a general theory of sublanguages (see Raskin, 1971) and subsequently tested and used by his own and other groups of researchers in the USSR on the material of diverse restricted sublanguages (see ibid; Gorodeckij and Raskin, 1971, Pt.1). The paper consists of two parts. Part 1 contains a very brief exposition of the basic principles of the approach. In Part 2 some advantages of the approach over other (unspecified) approaches unrestricted in this way are mentioned in the context of a few important problems of high quality translation. Since these problems were also discussed by the contributors to *Feasibility Study on Fully Automatic High Quality Translation*, at certain points of Part 2 the paper enters a dialogue of a sort with some of them (all the quotations and references followed by a name only are taken from the contribution by the corresponding author to the said report).

1. SOME BASIC PRINCIPLES OF THE RESTRICTED SUBLANGUAGE APPROACH

It might be observed that in most cases when the practical need of constructing an MT system arises, its input, i.e. the linguistic material which is to be subjected to such treatment, is highly restricted by certain conditions: it is usually a relatively *narrow field of science or technology* with texts which are *relatively homogeneous*, with a *limited vocabulary*, a *restricted set of syntactic constructions*, a highly *structured substance of the content plane*, and a relatively *constant system of values for all the relevant pragmatic parameters* which are determined in this case not by the individual properties of any particular situation of communication, as is usually the case in casual communication, but rather by the position of the field itself among the contiguous fields as well as in non-linguistic reality, in general. For such restricted sublanguages a simple algorithm of automatic processing was constructed and proved to be highly efficient in its practical applications.

The algorithm is based on an over-important property which follows, logically and practically, from the features which characterize the class of restricted sublanguages in the theory of sublanguages, including those which were emphasized above and which result in the irrelevance of all surface structure distinctions among sentences with identical deep structure or the exact synonymity of all paraphrases (and, in fact, even near-paraphrases). This property implies that each stem in the vocabulary of a restricted sublanguage tends to play certain permanent role in

all the situations described by those sentences where the stem occurs, no matter whether it takes the form of a verb, noun, or any other part of speech. A minimal sufficient inventory of these roles, which are given the status of semantic characteristics of stems, is compiled (usually it does not exceed 15 items) and each dictionary entry is assigned a certain characteristic. Then a scheme of the maximally extended sentence of the restricted sublanguage, a maximal deep structure of a sort, is postulated in such a way that each sentence (or rather, each clause) can be represented as a (partial) realization of this structure. Such structures can embed, nest in, etc., each other. The dictionary of the restricted sublanguage with all its entries being assigned semantic characteristics and the schema of the maximally extended sentence of the restricted sublanguage are the two instruments on which the universal-algorithm is founded. Texts of the restricted sublanguage constitute its input, the output being a sequence of (partially) filled, ordered and subordinated schemes/deep structures. By making the semantic characteristic assigned to each stem of the restricted sublanguage, more or less detailed, one may control the depth of semantic analysis. With its subalgorithms of "ellipsis analysis", "homogeneous parts analysis", "boundary analysis", the algorithm operates as a universal Turing machine in the sense, that having been fed the universally standardized information on a particular restricted sublanguage, it proceeds to analyze it in the universal way and is equally applicable to each and every restricted language.

Is Restricted Sublanguage Approach, RSA, applicable to all, or at least most relevant cases or can it be applied only in a few exceptional situations? It has been argued elsewhere (see Raskin, 1971; Ch.4.1) that the first alternative holds true while in the cases in which a polythematic informational system is needed it seems worthwhile to treat the processed texts as belonging to several distinct restricted sublanguages, and after distinguishing them with the help of a not too complicated device, to make use of the technique developed for restricted sublanguages.

2. RESTRICTED SUBLANGUAGE SOLUTIONS TO SOME PROBLEMS OF HIGH QUALITY TRANSLATION

Semantics and pragmatics and the quality of translation.

Recent developments in semantic and syntactic theory have demonstrated the practically indefinite potential depth of a complete linguistic description which seems to require much scarcely accessible (at present, if not in principle) information on "speech act conditions, conversation rules, and semantic interpretation which must be associated in an idiosyncratic way with the lexical item in question", on "a theory of illocutionary acts", on "a theory of discourse which relates the use of sentences in social and conversational situations", and on "a theory of natural logic" (Fillmore), while the pragmatic dimension of the text is said to include answers to such heterogeneous questions as "by whom the text was produced, for which kind of audience it was meant, which kind of background knowledge the producer of the text assumed to be

available to the audience, the time, the place, and other parameters of the situation in which the text was produced, etc." (Bar-Hillel).

Now, it is obvious that for an adequate translation, no matter whether it is human or automatic, all this highly complex information must be obtained and taken into consideration, otherwise the quality of translation falls down sharply. It is equally obvious that all this is far beyond the linguist's reach at the present stage of linguistic knowledge.

In order to arrive at a practical solution of this problem one should impose some restrictions on the process of MT. In other words, certain criteria of the quality of translation should be formulated, and if necessary and possible, lowered. One might try to restrict the output of an MT system in the sense that it should certainly not produce what the user does not actually need It is evident that the user of a translation of a scientific or technical text will not require as much finesse and subtlety as the user of a translation of a literary text. Some (e.g. Garvin) are prepared to go even further and construct systems which would produce clearly inadequate though still tolerable translations (in a sense nobody has even succeeded in defining) in order to gain in speed. Now, when "machine-aided translation" or similar approaches are suggested, a restriction is imposed on what the computer is supposed (and thought of as capable) to do.

The restriction on the input in RSA determines, of course, some restrictions on the output (but, certainly, not to the extent

of tolerating barely acceptable translations). On the other hand, rather on the contrary, the simplicity and easier observability of the material of a restricted sublanguage make automatic translation feasible, allowing at the same time and for the same reason for the total accountability of the sublanguage which makes it possible to account for and use for the practical purposes of translation all the complex semantic and pragmatic information which might be relevant for translation. Of course, what makes it possible is that the degree of complexity of such information in the restructed sublanguage is very much inferior to what might be observed in language as a whole. What follows, however, is that restricting the input of an MT system to a sublanguage of a certain type RSA ensures high quality translation within the sublanguage and no further restrictions or lowering of the quality of translation is necessary.

It should be mentioned at this point that RSA shares with "machine-aided translation" the property of requiring a limited amount of predetermined and routine human participation prior to automatic processing.

Syntax and semantics, lexicon and grammar. One of the major claims of RSA is that, at least in applications to restricted sublanguages, intricate and labor-consuming syntactic algorithms (cf. Mel cuk, 1964) are redundant. The universal algorithm is based on semantics and is designed to use linguistic information of "lower" linguistic levels (viz. morphology and syntax) in a few exceptional

cases of semantic ambiguity. This emphasis on semantics rather than on syntax in automatic language data processing systems takes on a new value when compared to current discussions of the relations of syntax and semantics in linguistic theory and the existence of a clear-cut boundary between them. Probably influenced by the tendency, at present prevailing statistically in theoretical linguistics, to claim the priority of semantics over syntax, and, moreover, to negate the existence of the boundary, even those researchers in MT who do not seem to be influenced by RSA also speak in favor of such a "semantically-based" position (e.g. Mey). The latter position is indirectly reinforced by the fact that purely syntactic contributions to the *Study* (e.g. Petrick) fail to prove their pertinence to the problem of actual realization of MT bearing instead on the relation of recent theoretical innovations to the feasibility of MT (see below).

Thanks to its basic principles and internal organization RSA came independently to a justification of the claim made by Garvin that it is operationally more effective to delegate most of the grammatical information used in an MT system to the lexicon rather than to the parsing algorithm.

Linguistic theory and feasibility of MT. RSA seems to contribute to the solution of the major dilemma concealed in this phrase by providing, in a way similar to the one discussed above in connection with semantic and pragmatic problems, an interesting halfway position, a middle ground of a sort which in a sense combines

some relevant properties of the two extremes.

In the light of quite a number of promising developments and achievements in linguistic theory, the pertinent question is whether these have, do, or will, contribute anything to MT, or the latter, as Lyons thinks, "will neither contribute very directly to, nor depend very directly upon, advances in linguistic theory." This basically defeatist position has at least two aspects, the one being that language is claimed to be too complicated to be successfully subjected to automatic processing, an opinion many theorists would subscribe to, and the other, proclaimed by MT operationalists" (e.g. Garvin) that much of what has been recently proposed in grammatical and semantical theory is far too strong for MT, and much weaker models, as a possible theoretical basis for practically feasible MT are required. The latter consideration is interestingly illustrated by the fact of the recent emergence of working automatic systems of language data processing, quite close in their restrictiveness to RSA though, rather contrary to it, not aiming at theoretical generalization, which use "analysis-based grammars" (cf. Winograd, 1972).

-However, it is natural for the linguist to be suspicious of any attempts to base an MT system on a theory or a model, which has been demonstrated to be inferior to some other theory or model. Any serious attempt to make use of any linguistic knowledge for any purpose must, he might feel, be based on an adequate theoretical framework, otherwise the ever present danger of ad hoc decisions could hardly be avoided. What might be missed in this

argumentation is the fact that, when dealing with computerized applications of linguistics, we impose on the linguistic material a fundamentally different phenomenon, with laws and logic of its own, which may be very foreign to the nature of human language and the mental mechanism which underlies it, and this might force us to give up purely linguistic theories, even if they seem based on the properties inherent in man's nature, and to adopt, in man-machine partnership, a compromising approach which would account for both human and machine nature. It is not unimaginable, though rather distressing if true, that, due to the essential difference between the two, no linguistic theory claiming or exhibiting the property of adequacy to the nature of human language can be directly "computed", i.e. taken in by the computer;

It seems, and this is substantiated by the material of some papers contributed to the *Study* (e.g. Karttunen), that the more dependent on some recent development in "pure linguistics" a paper is, the less pertinent to MT it becomes. The contradiction between linguistic theory aiming at adequacy and practical needs of MT and, for that matter, other problems of computational linguistics, is self-evident. In this situation RSA seems to be doing a unique job of reconciling the two extremes, since on the material of a restricted sublanguage it might turn out that the application of a grammar based on adequate linguistic theory would be quite practical and there would not be any need to seek more feasible ad hoc solutions. Besides that, RSA may contribute a great deal to what is essentially

an issue between "theory" and "practice" by:

(1) providing a suitable "testing ground" for various conflicting theories or models, both for those which claim linguistic adequacy and analysis-based ones;

(2) allowing one to select the most preferable alternative on the basis of complete and easily accessible evidence which might be relevant for the choice;

(3) enabling one to limit the strength of a too powerful but valid theory or model by making suitable modifications on the basis of easily observable linguistic material of the restricted sublanguage.

The basic principles of RSA make one think of its language independence.

REFERENCES

Gorodeckij, B. Ju., and V. V. Raskin, 1971. *Methods of Semantic Investigation of a Restricted Sublanguage*, Moscow: Moscow University Press (in Russian)

Mel'cuk, I. A., 1964. *Automatic Syntactic Analysis*, Novosibirsk USSR Academy of Sciences Press (in Pussian)

Raskin, V. V., 1971. *Towards a Theory of Linguistic Subsystems*, Moscow: Moscow University Press (in Russian - an English translation is in preparation by Mouton)

Lehmann, W. P., and R. Stachowitz, 1971. *Feasibility Study on Fully Automatic High Quality Translation*, Griffiss Air Force Base, Rome Air Development Center, RADC-TR-71-295

Winograd, T., 1972. *Understanding Natural Language*, New York: Academic Press

T E C H N I Q U E

LETTERS WITH vARIABLE VALUES AND THE MECHANICAL INFLECTION OF RUMANIAN WORDS

Minerva Bocşa
University of Timişoara
Romania

The generation by computer of written Rumanian words faces two difficult problems: to produce automatically the numerous alternations which modify the stem and to add the inflectional endings, building a rich set of classes and subclasses. The mechanical morphological analysis is also complicated because of the stem's phonetic alternations.

For example, the Rumanian words

UNIVERSITATE / UNIVERSITĂŢI	(university)
SERIOS / SERIOŞI / SERIOASA	(serious)
PUTEA / POT/POŢI / POATE	(may)
VEDEA / VĂD/VEZI / VĂZUI / VADA	(to see)

present the alternations

(1) A/Ă, T/Ţ, S/S, O/OA, U/O/OA, E/Ă/A, D/Z

Phonetic rules describing the occurrence of these stem modifications have several exceptions and must include the presence or absence of stress, which is not marked in ordinary

experiments in mechanical translation from English into Rumanian [16] and so on. Phonetic alternation in Rumanian has been investigated by Lombard [11], Felix [7], Juilland and Edwards [10], Augerot [1], and others.

The preparatory work for our automatic linguistic task has several stages:

Examine the inflection of each word.

Establish the set of phonetic alternations.

Attach a specific variable letter to each alternation. In our conception [4] these are different from those of [9, 14, 15].

Design a binary code for the variable letters, taking into account the possibilities of the IRIS 50.

Detach morphological parameters.

Code each word.

Punch a deck of cards.

The card file is the Morphological Dictionary. It is exploited by the programs in various ways. Here the working principles of a program to produce the paradigm (set of inflected forms) of each word in the Morphological Dictionary are presented.

In this process the computer writes the inflected forms in the p positions of the paradigm P. The stem allomorphs constitute a set A with n elements. The different distributions of the allomorphs of A in P are described by a set G of *grouping functions*

spelling. Nevertheless, the words with nonconstant stem are too numerous to be considered irregular. The method of storing the several allomorphs of the stem for automatic inflection misses the natural unity of the word.

We have constructed a mechanical *Morphological Dictionary*, containing 2058 written Rumanian words with a *synthetic representation* of all these phonetic alternations. An algorithm based on this representation generates the inflectional noncompound forms of these words. They are Rumanian nouns, adjectives, and verbs, the main part belonging to the *basic word stock* [8, 17]. About 45 percent of them present stem alternations.[1]

The algorithm whose logic was given in [3] is the background of a set of programs written in the programming language ASSIRIS for the French computer IRIS 50 and its Rumanian counterpart FELIX C-256. The programs were recently run at the Territorial Electronic Calculus Center of Timisoara, verifying the algorithm.

The synthetic representation uses G. C. Moisil's notion of *letters with variable values* [14, 15], which V. Gutu Romalo developed [9]. The setting of our research is Marcus's theory of mathematical linguistics [12, 13], Diaconescu's study of word segmentation and the degree of regularity [5, 6], Domonkos's

---

[1]It seems that in Rumanian only 20 percent or even less of the total number of words have these phonetic alternations, but in our dictionary reference is made generally to the most frequently used words, with relative frequency above 0.22% [17].

identified by numerals. Thus grouping function 00 associates allomorph *a* in *A* with positions 1, 2, 5, 6, ... in *P*, allomorph *b* in *A* with positions 3, 8, ... in *P* , etc. The different *partitions* of *A* are called *allomorph configurations* and symbolized by a/b (with $n = 2$), ab/c, a/bc, a/b/c, ... (with $n = 3$), etc. A *variable letter* maps the elements of the partition into the Rumanian alphabet A, Ă, Â, B, ..., Z, Ø (here Ø represents the empty letter). Thus the variable letter T/C with the configuration ac/bd has the *realization* T in allomorphs *a* and *c*, and another realization C in allomorphs *b* and *d*. Not all of the theoretically possible variable letters exist in Rumanian; we found 85.

The set of fixed, variable, and empty letters is called the generalized Rumanian alphabet. A version of it is given in [2]. Words can be represented in this alphabet in either external or internal code.

The program operates in several steps which are described and then illustrated.

Input. In the Morphological Dictionary, the fixed letters are punched in accordance with the standard card code. Each variable letter is punched as a numerical prefix of one or two decimal digits followed by a letter. Part of speech, number of allomorphs, word length, stem length, etc. appear as parameters.

1. Recoding. The computer reads the word on the punched card and recodes it into an internal code; each letter is one byte. A fixed letter has zone E or F (leading four bits 1110 or 1111); variable letters have other zones. The recoding instruction in IRIS 50 is TRTR (translate and test).

2. Realization. The program reads the word byte by byte. If the zone is E or F, it writes the byte into the allomorph registers. If the zone is less than E, the program constructs a realization for each allomorph and stores it in the allomorph register.

The principles that govern the decoding of a variable letter into realizations are given in [3]. As an example, take the rule for regular variable letters (zone 0, 1 ... 7). Each regular variable letter has two realizations, and in the internal code the zone of each realization is F. The numeric of one realization is identical with the numeric of the regular variable letter, and the numeric of the other realization is greater by 1. The method of encoding partitions for regular variable letters is explained on the next frame.

The next program stage is on frame 43.

ALLOMORPH CONFIGURATIONS FOR REGULAR VARIABLE LETTERS

Eight zones (0, 1, ..., 7) encode regular variable letters. Each stem has two, three, or four allomorphs. Each partition of the paradigm has two members for a regular variable letter; the numeric of the variable letter is copied into the allomorphs of the first member of the partition, and incremented by 1 into those of the second member.

	Number of Allomorphs		
Zone	2	3	4
0	a/b		ac/bd
1	a/b		a/bcd
2			ab/cd
3	a/b		ac/bd
4	a/b	a/bc	ad/bc
5	a/b	a/bc	a/bcd
6		ac/b	acd/b
7		ab/c	ab/cd

3. Recoding. The program recodes the allomorphs into EBCDIC by another TRTR instruction.

4. Distribution. The program distributes the allomorphs to their locations in another region. The word's grouping function controls the process.

5. Inflection. The program adds the inflectional endings to the right of the stem allomorph in conformity with the class and subclass noted on the punched card.

6. Printing. The program condenses the empty letter and prints the inflected forms.

We illustrate concisely these phases for two words from our Morphological Dictionary, the verbs A PUTEA (may), and A VEDEA (to see). They have, respectively, four and five different allomorphs of the stem.

Input. The content of the card is

PUTEA	P8U19A8TEA	V4	100403
VEDEA	V9E9DEA	V5	070300

8U, 19A, 8T, 9E, and 9D are variable letters in the external code.

Some morphological parameters are

V	verb; part of speech
$\begin{matrix}4\\5\end{matrix}$	number of allomorphs
$\begin{matrix}10\\07\end{matrix}$	word length
$\begin{matrix}04\\03\end{matrix}$	stem length
$\begin{matrix}03\\00\end{matrix}$	grouping function

1. After translation into the internal code the words are represented in storage as

EA 84 A9 86 F2 FO

E6 92 93 F2 FO

EA, F2, FO, and E6 represent the fixed letters P, E, A, and V. 84, A9, 86, 92, and 93 represent the variable letters U/O, Ø/A, T/T, E/Ă/A, and D/Z. The symbol Ø will be replaced by blank.

2. The four or three stem letters, specified by 04 or 03 on the punched card, give the following four or five allomorphs.

a	EA F5 FF FA		*a*	E6 F2 FC
b	EA F6 FF FA		*b*	E6 F1 FC
c	EA F6 FF FB		*c*	E6 F2 FD
d	EA F6 FO FA		*d*	E6 F1 FD
			e	E6 FO FC

The program decodes the irregular variable letter 84 and produces the realizations U and O (bytes F5, F6) in the allomorphs *a* (U) and *b*, *c*, *d* (O), in accordance with a translation table. (3) The allomorphs are translated into EBCDIC.

4. The allomorphs are placed in new registers as specified by the grouping functions 03 and 00.

PU T, PU T, PO T, PO T, POAT, PU T, PU T, PO T, ...

VED , VED , VAD , VEZ , VED , VED , VED , VĂD , ...

5. The inflectional endings are added.

PU TEA, PU TERE, PO T, POŢI, POATE, PU TEM, PU TEŢI, PO T, ...

VEDEA, VEDERE, VĂD, VEZI, VEDE, VEDEM, VEDETI, VĂD, ...

6. The computer condenses the empty letter in A PUTEA and prints theinflected forms.

The variable-letter method has the advantage of keeping the unity of the word in the Morphological Dictionary and producing the inflected forms correctly. At the same time it regularizes the greatest part of the irregular words. The only irregular verbs that still remain are A AVEA (to have), A DA (to give), A FI (to be), A LUA (to take), A STA (to stand). The other so-called irregular verbs A BEA (to drink), A MINCA (to eat), A RELUA (to retake), A USCA (to dry), A VREA (to want), and all the other semiregular verbs belonging to the third conjugation [5, 14] are regular for our algorithm, and so are the irregular nouns SORĂ-SURORI (sister), NORA-NURORI (daughter-in-law), OM-OAMENI (man), etc.

The program contains 1455 ASSIRIS statements and generates the inflected forms for all the 2058 words included in the Morphological Dictionary in 1 minute 39 seconds. It represents an experimental verification of our algorithm and may be extended without essential modifications to all other Rumanian words, coded in the same way.

Another program meant for users receives a word from the punched card without its special external code or grammatical parameters, looks for it in the Morphological Dictionary file now stored on the magnetic disk, and, if it is found, produces the paradigm of the word. Examples of its output appear on the next two frames.

Subsequent frames exhibit the complete internal and external codes.

The variable-letter method enables us to form an easy algorithm for morphological analysis, as indicated in [2].

TRANSCRIBED OUTPUT

Cuvîntul cerut : PUTEA Forma flexionară : Paradigma

Răspunsul ordinatorului :

1. PARADIGMA VERBULUI A PUTEA

Nr.prs.	Prezent indicativ	Imperfect	Perfect simplu	Mai mult ca perfect	Prezent conjunctiv	Imperativ
Sg. I	POT	PUTEAM	PUTUT	PUTUSEM	POT	
II	POȚI	PUTEAI	PUTUȘI	PUTUSEȘI	POȚI	POȚI
III	POATE	PUTEA	PUTU	PUTUSE	POATE	
Pl. I	PUTEM	PUTEAM	PUTURĂM	PUTUSERĂM	PUTEM	
II	PUTEȚI	PUTEAȚI	PUTURĂȚI	PUTUSERAȚI	PUTEȚI	
III	POT	PUTEAU	PUTURĂ	PUTUSERĂ	POATE	

Modurile nepersonale : Infinitiv PUTEA PUTERE

Participiu PUTUT

Gerunziu PUTÎND

TRANSCRIBED OUTPUT

Cuvîntul cerut : VEDEA Forma flexionară : Paradigma

Răspunsul ordinatorulus :

1. PARADIGMA VERBULUI A VEDEA

Nr. pers.		Prezent indicativ	Imperfect	Perfect simplu	Mai mult ca perfect	Prezent conjunctiv	Imperativ
Sg.	I	VĂD	VEDEAM	VĂZUI	VĂZUSEM	VĂD	
	II	VEZI	VEDEAI	VĂZUȘI	VĂZUSEȘI	VEZI	VEZI
	III	VEDE	VEDEA	VĂZU	VĂZUSE	VADĂ	
Pl.	I	VEDEM	VEDEAM	VĂZURĂM	VĂZUSERĂM	VEDEM	
	II	VEDEȚI	VEDEATI	VĂZURĂTI	VĂZUSERĂȚI	VEDEȚI	VEDEȚI
	III	VĂD	VEDEAU	VĂZURĂ	VĂZUSERĂ	VĂDĂ	

Modurile nepersonale : Infinitiv: VEDEA VEDERE

Participiu: VAZUT Gerunziu : VĂZÎND

GENERALIZED RUMANIAN ALPHABET (EXTERNAL AND INTERNAL CODE)

Zone Num.	0 0000	1 0001	2 0010	3 0011	4 0100	5 0101	6 0110	7 0111	E	F
0 0000	ØA Ø/A	1A A/Ø			4A A/Ă	5A Ă/A		7A A/Ă	Â	A
1 0001	ØĂ Ø/A				4Ă Ă/E	5Ă E/Ă	6Ă Ă/E	7Ă Ă/E	F	Ă
2 0010		1E E/Ø							K	E
3 0011	ØÎ Ø/Î		2Î Î/Â a/b	3Î I/Î	4Î Î/I				M	Î
4 0100	ØI Ø/İ	1I I/Ø							Q	I
5 0101	ØU Ø/U	1U U/Ø			4U U/O	5U O/U			R	U
6 0110		1O O/Ø							V	O
7 0111				3S S/S	4S S/Ş	5S Ş/S	6S S/Ş	7S Ş/S	W	Ş
8 1000	ØS Ø/S								X	Ş
9 1001	ØC Ø/C	1C C/Ø	2C C/Ø	3C T/C	4C C/T	5C T/C	6C C/T	7C C/T	Y	C
A 1010				3T Ţ/T	4T T/Ţ	5T Ţ/T	6T T/Ţ	7T T/Ţ	P	T
B 1011		1G G/Ø	2G G/Ø						G	Ţ
C 1100	ØN Ø/N	1N N/Ø	2N N/Ø		4D D/Z	5D Z/D	6D D/Z	7D D/Z	N	D
D 1101	ØH Ø/H				4Z Z/J		6Z Z/J		H	Z
E 1110			2B B/Ø						B	J
F 1111		1L L/Ø							L	B1

GENERALIZED RUMANIAN ALPHABET (EXTERNAL AND INTERNAL CODE)

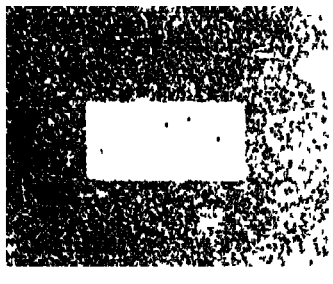
Zone Num.	8 1000	9 1001	A 1010	B 1011	C 1100	D 1101	E 1110	F 1111
0 0000	8A A/Ø ad/bc	9A A/Ø ae/bcd	10A Ă/E/A a/bd/c				Â	A
1 0001	8G G/P/Ø a/b/cd	9Ă A/Ă abe/cdf	11A E/A ac/bd				F	A
2 0010	8E E/I a/bc	9E E/Ă/A ac/bd/e	12A A/E/Ø a/b/cd				K	E
3 0011	8O O/U abc/d	9D D/Z abe/cd	13A Î/A/Ø a/b/cd				M	Î
4 0100	8U U/O a/bcd	9Z D/Z/Ø ab/c/de	14A E/A ab/cd				Q	I
5 0101	8Ş S/Ş abc/d	9J D/Z/Ø ac/bf/de	15A Î/Ă/A ab/c/d				R	U
6 0110	8T T/Ţ abd/c	9T T/Ţ/Ø a/b/cd	16Ă Ă/A ad/bc				V	O
7 0111	8D D/Z/Ø a/b/cd	9Z T/Ţ/Ø ab/c/de	17Ă Ă/A ab/cd				W	S
8 1000	8L L/Ø ac/bd	9Ş S/Ş abd/c	18A ØA abd/c				X	Ş
9 1001	8C C/P ab/cd	9G G/Ø ac/bd	19A Ø/A abc/d				Y	C
A 1010							P	T
B 1011							G	Ţ
C 1100							N	D
D 1101							H	Z
E 1110							B	J
F 1111							L	Bl

REFERENCES

1. Augerot, James E. *A study of Rumanian morphophonology.* Dissertation, University of Washington, 1968.
2. Bocşa, Minerva. Codage de l'alphabet généralisé du Roumain pour l'ordinateur IRIS 50 (FELIX C-256). *Cahiers de linguistique théorique et appliquée*, No. 10, Fasc. 2, 1973.
3. Bocşa, Minerva. Algoritm pentru generarea cuvintelor limbii române în ASSIRIS. *Revista de analiză numerică şi teoria approximaţiei*, vol. 3, fasc. 1, Cluj, 1974.
4. Bocşa, Minerva. *Dictionar morfologic automat al limbii române pentru ordinatorul IRIS 50 (FELIX C-256).* Dissertation, Universitatea din Timişoara, 1974.
5. Diaconescu, Paula. Contribuţii la definirea şi clasificarea verbelor regulate în limba română. *Studii si cercetări lingvistice, no. 2, an XI,* Bucureşti, 1960.
6. Diaconescu, Paula. *Structură şi evolutie în morfologia substantivului românesc.* Editura Academiei Republicii Socialiste România, Bucureşti, 1970.
7. Felix, Jiři. Asupra alternanţelor fonologice din flexiunea verbala romanească. *Studii şi cercetări lingvistice, no. 6, an XVI.* Bucureşti, 1965.
8. Graur, Alexandru. *Incercare asupra fondului principal de cuvinte.* Bucureşti, 1954.
9. Guţu Romalo, Valeria. *Morfologie structurală a limbii române.* Editura Academiei Republicii Socialiste România, Bucureşti, 1968.
10. Juilland, Alphonse, and P. M. H. Edwards. *The Rumanian verb system.* Mouton, The Hague, 1971.
11. Lombard, Alf. *Le verbe roumain.* Lund, 1954.
12. Marcus, Solomon. *Lingvistica matematică.* Editura didactică şi pedagogică, Bucureşti, 1966.

13. Marcus, Solomon, *Algebraic linguistics, Analytical models.* Academic Press, New York and London, 1967.
14. Moisil, Grigore C. Probleme puse de traducerea automata, Conjugarea verbelor în limba romana scrisa. *Studii şi cercetari lingvistice,* no. 1, Bucureşti, 1960.
15. Moisil, Grigore C. Problèmes posés par la traduction automatique. La dèclinaison en roumain écrit. *Cahier de linguistique theorique et appliquée,* no. 1, Bucureşti, 1962.
16. Nistor Domonkos, Erica. *Algoritm de traducere automată din limba engleza in limba romana.* Editura Didactică şi Pedagogica, Bucuresti, 1966.
17. Săteu, Valeriu. Observaţii asupra frecvenţei cuvintelor în operele unor scriitori romani. *Studii si Cercetari Lingvistice* no. 3, Bucureşti, 1959.

American Journal of Computational Linguistics Microfiche 9 : 53



ASSOCIATION
FOR
COMPUTATIONAL
LINGUISTICS

SECRETARY-TREASURER

A. Hood Roberts

REPORT FOR 1974

MEMBERSHIP

As of March 22, 1975, ACL had received dues for 1974 from

543 individuals :
132 libraries : 675

For 1973, the comparable figure is about 400.

FINANCIAL ACCOUNTS

Cash on hand August 1, 1973		$3,677.82
Receipts		8,119.00
Membership dues	$3,886.00	
IBM contribution for meeting	300.00	
AFIPS distribution of surplus	3,228.00	
Receipts at 1973 annual meeting	705.00	
		11,796.82
Disbursements		9,532.62
The Finite String.	$1,483.15	
Office supplies & expenses	891.62	
Expenses at 1973 annual meeting	347.85	
Expenses for 1974 annual meeting	300.00	
AFIPS dues	500.00	
Bank fee	10.00	
Certificates of deposit	6,000.00	
Cash on hand July 25, 1974		2,264.20
Savings accounts		6,130.39
Net worth July 25, 1974		$8,394.59

EDITOR S REPORT

The 1974 series of this new journal has been produced from the first 44 manuscripts received. Of this number,

10 have been published as microfiches
4 have been published as notes in The Finite String
10 are with the author for minor revision
7 are with the author for major revision
9 have been rejected
4 were returned to the author without substantive comment because they fall outside our scope

More than 2 out of 3 manuscripts go back to the author for some reason. Perhaps somewhat more than 1 out of 3 will eventually appear in this Journal. If our rejection rate remains low, it will be thanks to the microfiche format, which permits us to publish more pages per year than most journals.

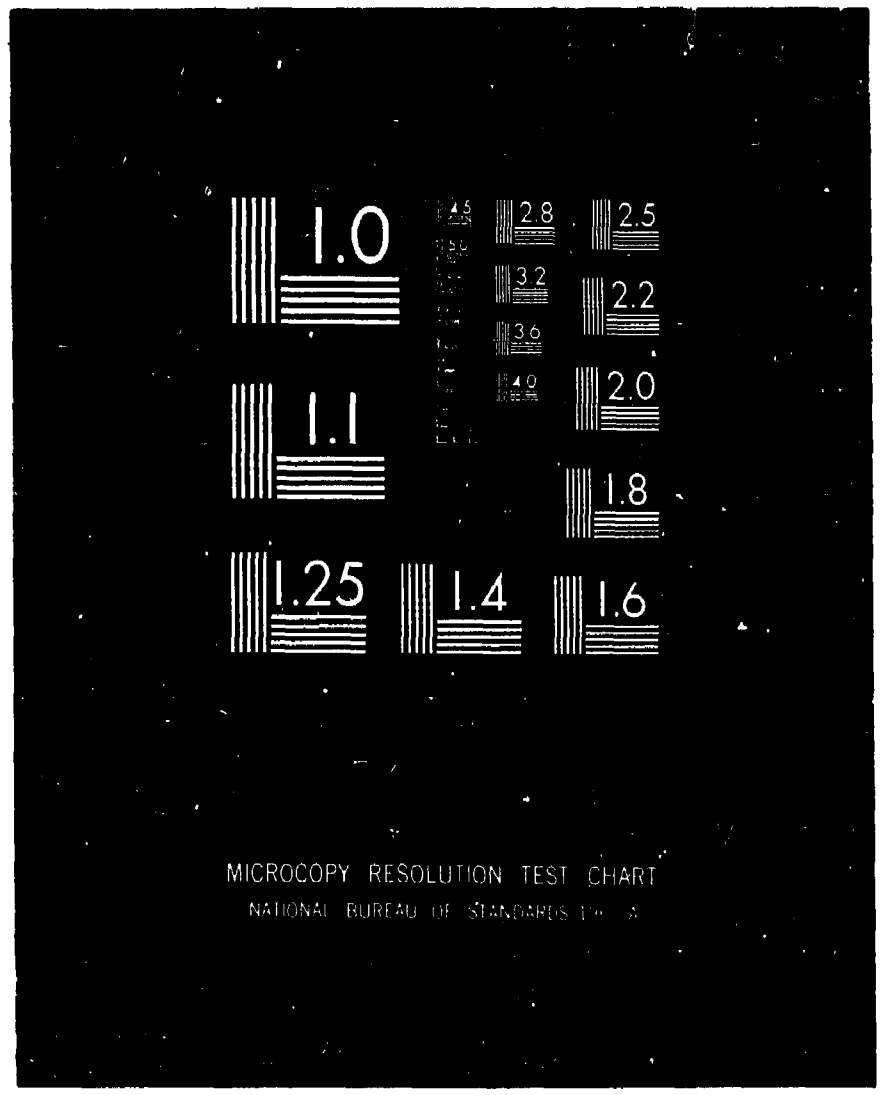
The 1974 series contains about 900 frames. As more contributors take advantage of the space available to them, individual fiches will be filled more often. As more workers in the field become aware of the quality of the audience, the rapidity of publication, and the value of space in which to make a full presentation of methods and results, the number of contributions per year will increase. We anticipate a larger series for 1975.

A new policy, effective immediately will probably make it possible in almost all instances to give the contributor the answer in about a month that has been the intent of the editor from the beginning.

The editor acknowledges with gratitude the support of the National Science Foundation, the general help of the Center for Applied Linguistics, the support provided by the Executive Committee and membership of the Association for Computational Linguistics, and the imagination and effort rendered by several persons--not including the editor--which led to the creation of AJCL. The work of the Editorial Board has been, and must continue to be, one of the foundations of the journal. Those whose names appear with mine on the first frame of this fiche work hard and well to make the journal useful to its readers.

The comments of all concerned are always welcome.

END


1.0
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.1
1.8
1.25
1.4
1.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A