MEMORANDUM
RM-3894-3
JUNE 1967

ANNOTATED BIBLIOGRAPHY OF RAND PUBLICATIONS IN COMPUTATIONAL LINGUISTICS

David G. Hays, Bozena Henisz-Dostert and Marjorie L. Rapp

The RAND Corporation
SANTA MONICA, CALIFORNIA
MEMORANDUM
RM-3894-3
JUNE 1987

ANNOTATED BIBLIOGRAPHY OF
RAND PUBLICATIONS IN COMPUTATIONAL
LINGUISTICS

David G. Hays, Bozena Henisz-Dostert and Marjorie L. Rapp

DISTRIBUTION STATEMENT
Distribution of this document is unlimited.
PREFACE

This bibliography lists and annotates 143 RAND publications on linguistic theory, linguistic research methods, computational methods for linguistics, the Russian language, the English language, information retrieval, automatic content analysis, psycholinguistics, and character readers. In order to keep the size and scope of the bibliography within convenient limits, other studies on such subjects as information theory (in the ordinary sense of probabilistic analysis of communication channels) and artificial intelligence in general (including nerve-net simulation) have been omitted here. References to each related RAND publication can, however, be found in the one-volume Index of Selected Publications of The RAND Corporation, 1946-1962, and in its successor, Selected RAND Abstracts, a quarterly series collected in annual volumes beginning in 1963.

The new or completely revised entries in this third revision of the Bibliography are 1-21, 1-22, 1-23, 1-24, 1-25, 2.2-2, 2.2-6, 2.2-7, 2.2-8, 2.2-9, 3-9, 3-10, 4.1-6, 4.2-16, 4.2-17, 5.1-1, 5.1-5, 5.2-8, 5.2-9, 5.2-10, 6-3, 7-13, 7-14, 7-15, 7-16, 7-17, 8-3, 9-11, 9-12, and 10-6.

Aside from studies that have appeared in commercial editions, the publications listed in this bibliography may be obtained by writing directly to the RAND Reports Department (1700 Main Street, Santa Monica, California 90406). Wherever a DDC document (AD) number is given, they are also available.
from the Defense Document Center (Cameron Station, Alexandria, Virginia 22314), or from the Clearinghouse for Federal Scientific and Technical Information (U.S. Department of Commerce, Springfield, Virginia 22151). From the latter two sources, requests will be filled most quickly if the AD numbers provided in the citations are used.

Virtually all of the reports have been deposited with the libraries indicated on pp. 24-27. There they are available to library patrons on the same basis as other holdings. They may also be borrowed through the Interlibrary Loan Service.

Commercially published books by RAND authors are available only through the publisher or a bookseller. Aside from such studies, the publications listed in this bibliography may be purchased by the public directly from the RAND Reports Department. The price schedule listed below was established to cover costs of reproduction and handling.

<table>
<thead>
<tr>
<th>Page Count*</th>
<th>Price**</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-49</td>
<td>$1.00</td>
</tr>
<tr>
<td>50-99</td>
<td>2.00</td>
</tr>
<tr>
<td>100-199</td>
<td>3.00</td>
</tr>
<tr>
<td>200 and over</td>
<td>4.00</td>
</tr>
</tbody>
</table>

*Page counts are included in each publication citation.

**California residents add 4 percent sales tax.
## CONTENTS

### PREFACE

- Section
  1. GENERAL
     1. New areas of application of computers
     2. Benchmarks in artificial intelligence
     3. Soviet cybernetics and computer sciences, 1960
     4. Automatic language-data processing
     5. Survey and critique
     6. The use of machines in the construction of a grammar and computer program for structural analysis
  7. Linguistic research at The RAND Corporation
  8. Linguistic analysis in machine translation research
  9. Automatic computers in machine translation research
  10. Six tasks in computational linguistics
  11. Dictionary problems in machine translation
  12. Soviet research in machine translation
  13. On cybernetics, information processing, and thinking
  14. U.S.-Japan seminar or mechanical covet costs of machine translation: Summary of U.S. contributions
  15. Cybernetics and its development in the Soviet Union
  16. Report of a summer seminar on computational linguistics
  17. Machine translation
  18. Technical writers: Educated or trained?
  19. Computational linguistics: Bibliography, 1964
  20. Computational linguistics: Bibliography, 1965
  21. Alchemy and artificial intelligence
  22. Computational linguistics: Research in progress at The RAND Corporation
  23. Computational linguistics: Bibliography, 1966
  24. Readings in automatic language processing
  25. Introduction to computational linguistics
2. LINGUISTIC THEORY ........................................... 4

2.1. DEPENDENCY THEORY ....................................... 4
1. Grouping and dependency theories ............... 4
2. Basic principles and technical variations in sentence-structure determination ... 4
3. Dependency systems and phrase structure systems .......... 4
4. Dependency theory: A formalism and some observations .......... 4
5. An annotated bibliography of publications on dependency theory .......... 5

2.2. OTHER THEORY ............................................. 5
1. An experimental study of ambiguity and context .......... 5
2. The quantification of functional load: A linguistic problem .......... 5
3. Suprasyntactics ............................................ 5
4. Computers and comprehension ................................ 5
5. Linguistic and non-linguistic "understanding" of linguistic tokens .......... 5
6. A simple proof of a theorem on self-synchronizing automata .......... 5
7. Some remarks on acceptable sets of numbers ............... 6
8. On products of finite dimensional stochastic matrices .......... 6
9. The queens grammar ....................................... 6

3. RESEARCH METHODS ......................................... 6
1. Research procedures in machine translation ............... 6
2. An introduction to computational procedures in linguistic research ... 6
3. Research methodology ..................................... 7
4. On the value of a dependency connection ............... 7
5. Manual for postediting Russian text ............... 7
6. Some Monte Carlo estimates of the Yule distribution .......... 7
7. The logic of cognate recognition in historical linguistics .......... 7
8. Language identification in the limit ............... 7
9. The termination of certain iterative processes ............... 8
10. Automatic classification in linguistics ............... 8
4. COMPUTATION .................................................. 8

4.1. CODES AND FORMATS ................................. 8
1. Manual for keypunching Russian scientific text .......... 8
2. Resume of machine codes and card formats .......... 8
3. Manual for pre-editing Russian scientific text .......... 8
4. Manual for coding Russian grammar ................. 8
5. Natural language in computer form .................. 8
6. Acquisition, archiving and interchange .................. 9

4.2. PROGRAMS AND SYSTEMS .......................... 9
1. The nature of data in language analysis .......... 9
2. MIMIC: A translator for English coding .......... 9
3. Glossary lookup made easy ......................... 9
4. Russian sentence-structure determination .......... 9
5. Connectability calculations, syntactic functions, and Russian syntax ............... 10
6. Translation of artificial languages by compiler programs ....... 10
7. Experiments with a heuristic compiler ........ 10
8. A parsing program for categorial grammars .......... 10
9. PALL: RAND's automatic address book ........ 10
10. Endocentric constructions and the Cocke parsing logic .......... 10
11. Large files in linguistic computing ........ 10
12. The catalog: A flexible data structure for magnetic tape .......... 10
13. The catalog input/output system .............. 11
14. The tabular parser: A parsing program for phrase structure .......... 11
15. Computer routines to read natural text with complex formats .......... 11
16. The catalog: A flexible structure for data storage .......... 11
17. COLLECT: A program for the retrieval of grammatical information from annotated text .......... 11

4.3. PROGRAM LANGUAGES .............................. 12
1. The evolution of concepts and languages of computing .......... 12
2. An experimental syntax-directed data structure language .......... 12
5. THE RUSSIAN LANGUAGE

5.1. TEXT AND GLOSSARIES
1. Bibliography of Russian scientific articles
2. A glossary of Russian physics on punched cards
3. A glossary of Russian physics
4. High frequency words and occurrence forms in Russian physics
5. A file of Russian text with syntactic annotations

5.2. SYNTAX
1. Order of subject and object in scientific Russian when other differentia are lacking
2. A Russian structure for comparison
3. Governors of the conjunction chto
4. Some linguistic problems of Russian graphic abbreviations
5. The suffix -aga
6. Predicative case, short form adjectives and predicatives
7. Complementation in Russian: Theory and application
8. A deep index of derivational morphology
9. Toward exploitation of a file of Russian text with syntactic annotations
10. Vowel-zero alternations in derivation

5.3. SEMANTICS
1. Pairs of Russian words with high correlation
2. Procedures for the determination of distributional classes
3. Transformational criteria for the classification of predicative genitive constructions in Russian
4. Machine translation of Russian prepositions
5. The position of prepositional phrases in Russian
6. Prepositional phrases and automatic parsing
7. Measurement of similarity between nouns
8. Studies in inter-sentence connection
9. On the predicative use of the Russian infinitive
10. Some combinatorial properties of Russian nouns ........................................ 16

6. THE ENGLISH LANGUAGE ................................................................. 16
   1. Preliminary codes and rules for the automatic parsing of English .............. 16
   2. PARSE: A system for automatic syntactic analysis of English text - Parts I & II .......................................................... 16
   3. Some aspects of the thematic organization of the English clause ................. 17

7. RETRIEVAL AND INDEXING .................................................................. 17
   1. On relevance, probabilistic indexing and information retrieval ................. 17
   2. Automatic indexing: An experimental inquiry ............................................. 17
   3. A logician's view of language data processing ........................................... 17
   4. Information retrieval: A look at the logical framework and some new concepts .......................................................... 17
   5. Probability and the library problem ......................................................... 17
   6. A proposal for the indirect retrieval of unpublished technical material ....... 18
   7. The Soviet classification scheme for literature ......................................... 18
   8. Automatic parsing and fact retrieval: A comment on grammar, paraphrase, and meaning .................................................. 18
   9. Mechanized documentation: The logic behind a probabilistic interpretation .......................................................... 18
  10. The logic of interrogating a digital computer ............................................ 18
  11. Relational data file: A tool for mechanized inference execution and data retrieval .................................................. 18
  12. The transformation of sentences for information retrieval ......................... 19
  13. Relational data file I: Design philosophy .............................................. 19
  14. Relational data file II: Implementation .................................................. 19
  15. A computer system for inference execution and data retrieval .................. 19
  16. A formal system for the logical analysis of temporal relationships between intervals of time ........................................ 20
  17. A simple scheme for formalizing data retrieval requests .......................... 20
8. CONTENT ANALYSIS ........................................... 20
   1. Automatic language-data processing ........................ 20
      in sociology ............................................ 20
   2. Automatic content analysis: Some entries for a transformation catalog ... 20
   3. Processing natural language text .......................... 20

9. PSYCHOLINGUISTICS .......................................... 20
   1. An information processing theory of verbal learning ................ 20
   2. The simulation of verbal learning behavior .................... 21
   3. Performance of a reading task by an elementary perceiving and memorizing program ........ 21
   4. A theory of the serial position effect ........................ 21
   5. A net to simulate Morse-code learning ....................... 21
   6. Similarity and familiarity in verbal learning .................. 21
   7. Construction of a simulation process for initial psychiatric interviewing ................ 21
   8. On the construction of a simulation of the initial psychiatric interview ... 21
   9. Computer simulation of human behavior ....................... 21
  10. Generalization of an elementary perceiving and memorizing machine ............. 22
  11. Linguistic relativity and the language learning process ...................... 22

10. CHARACTER READERS ......................................... 22
    1. A digital simulation of an aided adaptive character reading machine .... 22
    2. An aided adaptive character reader for machine translation of languages ................ 22
    3. The RAND Tablet: A man-machine graphical communication device .......... 22
    4. Abstraction and pattern classification .................................. 22
    5. Computer recognition of on-line, handwritten characters .................. 23
    6. On the development of equitable graphic I/O ............................ 23

11. TRANSLATIONS ................................................ 23
    1. Exact methods in linguistic research .................................. 23
    2. Two operators for determining agreement for automatic syntactic analysis .......... 23
3. Problems of algorithmic composition of subject indexes

DEPOSIT LIBRARIES IN THE UNITED STATES

DEPOSIT LIBRARIES ABROAD

AUTHOR INDEX

PUBLICATION NUMBER INDEX

23

24

27

28

33
Annotated Bibliography of RAND Publications in Computational Linguistics

1. General

1. NEW AREAS OF APPLICATION OF COMPUTERS
A. Newell
P-3114, November 1960, 4 pp.
In 1960, computers were in use for engineering calculation, business-data processing, preparation of Bible concordances, prediction of election returns, discovery of new chemical names, composition of elementary music, solution of double-crostics, and playing solitaire. Future applications suggested were the handling of large hydrodynamic problems and of raw language, and the development of elementary language-understanding programs.

2. BENCHMARKS IN ARTIFICIAL INTELLIGENCE
F. L. Grubbins
A list of tasks, presently performable by most human beings, that may eventually be performed successfully by computers. These tasks can serve as milestones in the field of artificial intelligence.

3. SOVIET CYBERNETICS AND COMPUTER SCIENCES, 1960
E. A. Feigenbaum
AD 266 129
Feigenbaum describes his experiences as a delegate to the International Congress on Automatic Control held at Moscow, June 27 - July 7, 1960. He discusses (i) certain aims of the conference; (ii) some Soviet research projects in artificial intelligence and biocybernetics; and (iii) general Soviet attitudes, techniques, and directions in the cybernetic and computer-related sciences. He concludes that Soviet research in the computer sciences lags behind Western development, but that the gap is neither large nor based on a lack of understanding of fundamental principles. The Soviets will progress rapidly if and when priority, in terms of access to computing machines, is given to their research.

4. AUTOMATIC LANGUAGE-DATA PROCESSING
D. G. Hays
(Unsponsored)
ALDP includes content analysis (i.e., of novels, editorials, discussion-group transcripts, interviews), MT, etc. Translation is taken as illustrative: Input and output problems, dictionary lookup, sentence-structure determination, and output construction are discussed. Phrase-structure and dependency theories are briefly compared. The first routine capable of finding all structures for a given sentence under a given grammar with a feasible number of operations (due to John Cocke) is reported, and compared with the earlier RAND program that could find at most one structure per sentence. Semantics, and its relation to information-retrieval and psychological ALDP systems, is treated briefly.

5. STUDIES IN MACHINE TRANSLATION--I:
SURVEY AND CRITIQUE
H. P. Edmundson, K. E. Harper, and D. G. Hays
AD 150 672
This very early survey noted that both words and sentence structures had to be translated. Glossary compilation and lookup, idioglossaries, thesauruses, methods, idiom recognition, translation of prepositions, construction of grammars by semiautomatic methods, early sentence-structure determination routines, and hardware requirements for input, storage, and output were reviewed. Great stress was laid on empirical linguistics.

6. THE USE OF MACHINES IN THE CONSTRUCTION OF A GRAMMAR AND COMPUTER PROGRAM FOR STRUCTURAL ANALYSIS
K. E. Harper and D. G. Hays
P-1588, January 1959, 15 pp.
The cyclic research process is briefly described, including mention of the use of structural concordances in identification of new syntactic classes. The RAND sentence-structure determination routine, based on precedence and dependency theory, is outlined.

7. LINGUISTIC RESEARCH AT THE RAND CORPORATION
D. G. Hays
P-1900, February 1960, 22 pp.
In 1960, the RAND project was interested in postediting as a basis for linguistic research, the use of semi-automatic research methods to follow postediting, and distributional analysis of transformations and semantics. The notion of idiom-in-structure is mentioned here: it is a fixed combination of words having distinctive meaning, but free to occur in different orders.
8. LINGUISTIC ANALYSIS IN MACHINE TRANSLATION RESEARCH
H. P. Edmundson
P-1526, April 1958, 10 pp.
In: Martha Boaz, ed., Modern Trends
in Documentation, Pergamon Press, New
York, 1959, pp. 31-37.
A popularized exposition of some funda-
mental ideas in linguistics and mathe-

matics.

9. AUTOMATIC COMPUTERS IN MACHINE-TRANSLATION RESEARCH
G. Hays
P-1321, April 1958, 10 pp.
In: Martha Boaz, ed., Modern Trends
in Documentation, Pergamon Press, New
York, 1959, pp. 38-44.
A popularized exposition of the cyclic
research method, of phrase-structure
parsing, and of requirements for com-
puters with special features.

10. SIX TASKS IN COMPUTATIONAL LINGUISTICS
K. E. Harper, D. G. Hays, D. S. Worth,
and T. W. Ziehe
AD 264 769
(Air Force Office of Scientific Research: AFOSR-TN-1362)
(i) Distributional semantics. 42 distribu-
tional classes are listed. For example,
some verbs always take an animate
subject (103 such verbs were found in a
block of about 100,000 occurrences of
physics verbs).
(ii) Discovery procedures for transfor-
mations. If two structures are trans-
formationally linked in the sense of
Harris, they should tend to occur in
text with about the same words filling
them.
(iii) Development of a Russian word-
family glossary. The forms in the RAND
physics glossary were segmented by syn-
chronic criteria into constituent morphs.
About 2,500 families, each having a com-
mon root, were identified.
(iv) A text processor. This system is
used to transcribe printed pages of
text on tape, list from tape, and make
corrections. Codes and formats are
given.
(v) A glossary editor. Preliminary dis-
cussion of a system for maintenance of a
deeply organized glossary file.
(vi) Text-improvement program. Discus-
sion of the nature of changes in the RAND text file, in-
cluding both raw text and editors' notes,
and of procedures applied to eliminate
them.

11. DICTIONARY PROBLEMS IN MACHINE TRANSLATION
K. E. Harper
In: Paul L. Garvin, ed., Natural
Language and the Computer, McGraw-
215-222.
Punched-card and magnetic disk or drum
systems are inherently slow. Tape sys-
tems have been used in several ways:
(i) Alphabetized text is compared with
alphabetized dictionary. (ii) Text is
compressed into core storage and the
dictionary read from tape. (iii) The
dictionary is compressed into storage by
use of letter-table addressing. The
second and third systems are fast enough
to make the need for special-purpose
equipment doubtful. Treatment of
"idioms" (input units containing internal
blanks) and of word segmentation are
mentioned, and methods of collection of
entries for MT dictionaries are con-
sidered. Harper also discusses the form
and content of grammar and semantic
codes and target-language equivalents.

12. SOVIET RESEARCH IN MACHINE TRANSLATION
K. E. Harper
AD 244 725
In: H. P. Edmundson, ed., Proceedings of the National Symposium on
2-12.
In 1960, a survey of the available lit-
erature indicated that Soviet research-
ers were working hard on MT, developing
glossaries and routines for input analy-
sis, transfer, and output synthesis, but
also developing sophisticated linguistic
theories applicable to the problem.
Computing facilities seemed to be un-
available except for short demonstra-
tions.

13. ON CYBERNETICS, INFORMATION PROCESSING, AND THINKING
M. E. Maron
P-2879, March 1964, 41 pp.
AD 435 484
In: N. Wiener and J. P. Schade, eds.,
Cybernetics of the Nervous System (Pro-
gress in Brain Research, vol. 17),
It is the purpose of this paper to examine
the origins, development, and present status
of those key cybernetic notions that pro-
vide an information-flow framework within
which to attack one aspect of the question
of how a person thinks; i.e., the question
of the information mechanisms and processes
which underlie and are correlated with
thinking.

After an introductory survey of the
scope and ramifications of the informa-
tion sciences, the cybernetic way of
looking at information processing in the
nervous system is examined so as to
see what sense it provides new and
sharp tools of analysis for the neuro-
physiologist. With this as background,
the problem of artificial intelligence
is considered and with that the logical
and linguistic difficulties in talking
about the relationship between thinking
and brain activity. An information-flow
model of an artificial brain mechanism is
described whose activity, it is argued,
is the correlate to activity such as
perceiving, learning, thinking, knowing, etc.

This leads finally to a consideration of
the impact of these notions on theoretical
neurophysiology and its attempt to
frame suitable hypotheses, and on episte-
modology which is concerned with the logical
analysis of measures, methods, and
techniques which can justify the activi-
ty of knowing.

14. U.S. - JAPAN SEMINAR ON MECHANICAL
TRANSLATION: SUMMARY OF U.S. CONTRIBU-
TIONS
D. G. Hays
AD 433 850

The summary examines MT as a branch of
computational linguistics, reviews diver-
tent theories in linguistics, notes cur-
rent problems of descriptive linguistics,
sketches computational problems, and
notes administrative problems of inter-
national cooperation.

15. CYBERNETICS AND ITS DEVELOPMENT IN
THE SOVIET UNION
R. Leven and M. E. Maron
AD 602 705

Following a brief introduction, the
Monumental traces the birth of cyber-
netics and sketches its early develop-
ment and its growth and emergence in
the West. The second half of the
Monumental examines Soviet cybernetics.

16. REPORT OF A SUMMER SEMINAR ON COMPU-
TATIONAL LINGUISTICS
D. G. Hays
AD 431 868

(=National Science Foundation)
The Seminar was held from July 8 through
August 30, 1963, in Santa Monica. The
staff taught elementary programming,
linguistic theory, mathematical linguis-
tics, and existing programs for linguis-
tic operations to twenty professors of
linguistics. The report includes a de-
tailed schedule, an analysis and evalua-
tion of the Seminar by the participants,
and a list of all persons involved.

The evaluation indicates that most par-
ticipants, while attempting to obtain
further education in this area, will
meanwhile begin using computers in their
own research and that many of them will
also teach this subject - some in special
courses, some in general linguistics
courses, and others only to research
students.

17. MACHINE TRANSLATION
K. E. Harper
In T. A. Sebeok, ed., Soviet and
European Linguistics, Current
Trends in Linguistics, vol. 1,
Mouton & Co., The Hague, 1963,
pp. 133-142.

After writing about 20 algorithms for
different pairs of languages, Russian
linguists turned their attention more to
basic studies of language and development
of theory. Harper scans the literature
and summarizes a variety of recent work.

18. TECHNICAL WRITERS: EDUCATED OR
TRAINED?
V. Peterson
P-3069, March 1965, 8 pp.
AD 612 380

Two possible innovations in the computer
field which may influence the future ofeditors and technical writers are dis-
cussed; a typewriter-like console and a
tablet. These changes should not affect
either the responsibilities of the editor
or his training and qualifications.

19. COMPUTATIONAL LINGUISTICS:
BIBLIOGRAPHY, 1964
D. G. Hays and R. Ma
AD 613 311

The bibliography cites 809 U.S. and
foreign articles, reports, and books par-
ticularly relevant to the fields of
computational linguistics and documen-
tation, with selective coverage in the
fields of classification theory, compu-
tation and programming, computers and
hardware, nonnumerical application of
computers, and psycholinguistics. In the
area of linguistics, a fairly broad view
of structural theory and semantics is
taken without being exhaustive.

20. COMPUTATIONAL LINGUISTICS:
BIBLIOGRAPHY, 1965
D. G. Hays, B. Henisz-Dostert,
M. L. Rapp
P-4986-PR, April 1966, 74 pp.
The bibliography cites 700 U.S. and
foreign articles, reports, and books rele-
vant to the fields of computational
linguistics and documentation which
appeared during late 1964 and through-
out 1965. For a more detailed account of
the coverage see 1-19.

21. ALCHEMY AND ARTIFICIAL INTELLIGENCE
H. L. Dreyfus
AD 625 719

An examination of the difficulties of sim-
ulating cognitive processes on com-
puters in the areas of game playing,
problem solving, language translation,
and pattern recognition reveals that the
attempt to analyze intelligent behavior
in digital computer language systemat-
ically excludes three fundamental human
forms of information processing: fringe
consciousness, essence/accident discrimi-
ation, and ambiguity tolerance.

There are four distinct types of in-
telligent activity: 1. Associationistic
(Irrelevance of meaning and context),
2. Non-formal (Dependent on meaning and
context, which are not explicit), 3.
Simple formal (Meanings completely ex-
plcit and context-independent), and
4. Complex formal (in principle, same
as 3.; in practice, internally context-
dependent, independent of external
context). Only areas 1. and 3. do not presuppose those human forms of information processing and can be programmed.

Significant developments in artificial intelligence in the remaining two areas must await computers which would combine their capacity for fast and accurate calculation with the short-cut processing made possible by the fringes of consciousness and ambiguity tolerance.

22. **COMPUTATIONAL LINGUISTICS:**
**RESEARCH IN PROGRESS AT THE RAND CORPORATION**
D. G. Hays P-3456, August 1966, 9 pp.
AD 657 929
RAND's linguistic research program is reviewed from its establishment in 1957. The current and projected work of the linguistics group is discussed in such areas as linguistic theory (in particular dependency and transformational), descriptive studies of Russian and to a lesser degree English, and the development of basic computing tools, especially the text and catalog system.

23. **COMPUTATIONAL LINGUISTICS:**
**BIBLIOGRAPHY, 1966**
The bibliography cites 914 U.S. and foreign articles, reports, and books relevant to the fields of computational linguistics and documentation which appeared during late 1965 and throughout 1966. For a more detailed account of the coverage see 1-19.

24. **READINGS IN AUTOMATIC LANGUAGE PROCESSING**
D. G. Hays, ed. (Unsponsored)
The introductory (Hays) mentions the computer as reducer of linguistic data and deliverer of conclusions from theoretical premises. It also mentions applications in social science, library work, translation, etc. It lists some techniques, said to unify the fields: file management, textual input and output, dictionary lookup, and parsing. Among the papers in the volume are 3-3, 4-1-5, 4-2-5, and a brief expository paper on parsing by the editor.

25. **INTRODUCTION TO COMPUTATIONAL LINGUISTICS**
D. G. Hays (Unsponsored)
A textbook with chapters introducing computers and algorithms; storage structures; external storage; acquisition, storage, and presentation of textual data; dictionary lookup; parsing strategies; techniques for storing and using grammars; context-sensitive and transformational parsing; stratal conversion; concordances; techniques for linguistic research; documentation; and automatic translation.

2. **Linguistic Theory**

2.1. **Dependency Theory**

1. **GROUPING AND DEPENDENCY THEORIES**
AD 250 237
Immediate-constituent theory (phrase-structure theory) is based on a topology of grouping, whereas dependency theory uses a topology of trees, each minimal syntactic unit occupying a node in a tree. A concept of correspondence between the two kinds of structures is defined, and the two topologies are compared.

2. **BASIC PRINCIPLES AND TECHNICAL VARIATIONS IN SENTENCE-STRUCTURE DETERMINATION**
Basic principles of the RAND method of sentence-structure determination include the isolation of grammatical detail from the structure of the computer program and postulation of a certain word-order rule, the rule of projectivity, that is realized in the program. Technical variations control the order of establishment of connections, the format of the grammar used in testing agreement, and other matters.

3. **DEPENDENCY SYSTEMS AND PHRASE STRUCTURE SYSTEMS**
Information and Control, vol. 8, no. 3 (June 1965), pp. 304-337.
The formalism of dependency theory was first published here, together with proofs (i) that the RAND SSD routine is adequate for recognition of sentences produced by the formalism; (ii) that dependency and phrase-structure are weakly equivalent, since every language that has a finite grammar of one kind has a finite grammar of the other kind as well; and (iii) that the two theories are almost strongly equivalent, in the sense that all but a specified subclass of phrase-structure grammars correspond to dependency grammars, characterizing the same language and assigning corresponding structures to their sentences.

4. **DEPENDENCY THEORY: A FORMALISM AND SOME OBSERVATIONS**
AD 602 648
Dependency grammars characterize the class of context-free languages, assigning to each sentence of a characterized language a tree structure with minimal syntactic units at the nodes. Both production and recognition procedures are given. Either transformational or stratified linguistic systems can be constructed on the basis of dependency theory; more attention is given to the latter possibility. Semantic and psychological considerations are cited as motivating specific features of the theory, but they are no more necessary as justifications of this theory than for others.

5. AN ANNOTATED BIBLIOGRAPHY OF PUBLICATIONS ON DEPENDENCY THEORY
D. G. Hays
AD 613 469
Seventy-one annotated citations in the field cover early works, formal analyses, parsing procedures, synthesis procedures, natural languages (English, Czech, French, and Russian) and discourse analysis.

2.2. Other Theory

1. AN EXPERIMENTAL STUDY OF AMBIGUITY AND CONTEXT
A. Kaplan
Mechanical Translation, vol. 2, no. 2

Using English text, native speakers of English, and the criterion of distinguishing among dictionary definitions, Kaplan finds that relatively little context before or after a word suffices to eliminate most ambiguity.

2. THE QUANTIFICATION OF FUNCTIONAL LOAD: A LINGUISTIC PROBLEM
C. F. Hockett
(Revised from P-2338, June 1961)

The function of a phonemic system is to keep the utterances of a language apart; distinctions that do more work are said to carry a heavier functional load. Hockett proposes criteria for indices of functional load which are satisfied by three measures of entropy (all used by Shannon in information theory); the three measures are in binitis, in Shannons, and as relative entropy. The units analyzed can be phonemes, allophones, or components. The load carried by a contrast is non-negative; it is zero if the contrasted units are identical or in complementary distribution, Hockett's measure is the change in the entropy of the system if the contrasted phonemes are coalesced. He discusses some problems peculiar to the allophonic case.

3. SUPRASYNTACTICS
D. S. Worth
AD 276 152
Suprasyntactic structures are those of negation, interrogation, and emphasis.

The last of these is most fully treated in this report. Noting that stress on the governor in a combination (i.e., the head in a construction) is ambiguous, since it can mean emphasis either of the item itself or of the whole construction, Worth goes on to give numerous other examples of devices for emphasis in various languages and grammatical situations, fitting them all into a transformational model.

4. COMPUTERS AND COMPREHENSION
M. Kochen, D. M. Mackay, M. E. Maron, M. Scriven, and L. Uhr
RM-4085-PR, April 1964, 72 pp.
AD 437 589
Understanding entails construction of an internal model. The input to be understood must be handled not merely as symptomatic of a state of affairs, but as descriptive of a state. Comprehension is distinguished from learning. The highest level of comprehension involves ability to learn, to model the self and the interlocutor as well as the domain of discourse, and thus to engage in a fully reciprocal dialogue.

5. LINGUISTIC AND NON-LINGUISTIC "UNDERSTANDING" OF LINGUISTIC TOKENS
D. M. Mackay
AD 432 308
This paper explores the distinction between recognizing an utterance, on the one hand, as a "symptom" of the state of affairs in or confronting its originator and, on the other hand, as a linguistic tool. It is suggested that although only the first kind of understanding is required to enable a computer to acept data and answer questions in verbal form, such ability is no guarantee of its comprehension of all aspects of language. To attain full linguistic comprehension, the program must also embody at least a "skeleton representation" of the linguistic context in which an utterance originates and from which it derives its linguistic significance as a goal-directed operator. By use of this richer model of its situation, a computer should be able to meet more sensitive tests of linguistic comprehension.

6. A SIMPLE PROOF OF A THEOREM ON SELF-SYNCHRONIZING AUTOMATA
D. M. Landi
AD 636 097
A more direct, short proof is given for verifying Winograd's theorem that a finite state, completely specified automaton will resynchronize itself with probability 1. if and only if there exists a finite sequence of input letters called a universal synchronizer of the automaton.
7. SOME REMARKS ON ACCEPTABLE SETS OF NUMBERS
M. P. Schützenberger
P-3413, August 1966, 10 pp.
AD 637 123
Consideration is given to the application of negative results on the so-called acceptable sets of numbers to the case of push-down automata with the help of conventional analytic techniques. Two examples are given, and some related unsolved problems are discussed.

8. ON PRODUCTS OF FINITE DIMENSIONAL STOCHASTIC MATRICES
M. P. Schützenberger
P-3459, August 1966, 10 pp.
AD 638 847
This study presents and verifies a partial extension of a theorem of Wolfowitz on products of indecomposable, aperiodic stochastic matrices.

9. THE QUEENS GRAMMAR
S. L. Lieman
AD 645 451
The famous eight-queens problem is presented and then generalized into a language over strings of integers. Various formalisms are shown to be able to describe this language, including arithmetical relations defined on each pair of integers in the string, programs for actually generating and parsing strings in the language, and a "relational grammar". It is shown that context-sensitive grammar and transformational grammar are likely to be capable of defining the language. The simplicity and elegance of the arithmetical formulation and the computer programs that are easily derived from it are contrasted to the counterintuitive properties of relational grammars, context-sensitive grammars and transformational grammars that define the same language. A natural language counterpart to the Queens language is given, and new measures of linguistic adequacy are defined. The implications of these new measures of adequacy to efforts for finding appropriate linguistic formalisms is discussed.

3. Research Methods

1. RESEARCH PROCEDURES IN MACHINE TRANSLATION
D. G. Hays
AD 268 941

Processing natural language on a computer calls for precise, accurate, voluminous knowledge of the linguistic behavior of the speakers or authors whose utterances or writings are to be processed. That knowledge can be acquired by analysis of naturally produced or elicited texts, or from comments appended to texts by editors. Syntax is described in terms of the dependency model, and indicators of syntactic connections (inflection, function words, occurrence order, and punctuation or intonation) are discussed. Under the heading of semantics, the problem of the unit of analysis is raised and the evidence that ambiguity exists is quoted. Calculation of the meaning of a sentence from the semantic properties of its parts and of their syntactic interrelations is discussed; the translation unit, or semantic unit, is the smallest linguistic unit that permits such calculation by simple rules. A scheme for automatic sentence-structure determination, the Cocke routine with an adaptation (suggested by Lecerf) to produce dependency structures, separates occurrence-order operations from all others, and obtains all possible structures for each sentence. Several plans for semantic recognition are mentioned, but none has been tested thoroughly, and all of them together still seem inadequate.

The text-based research procedures described are cyclic: some information about the language is available, and text is processed in accordance with that information, the results analyzed, and the new information used in processing more text. Standardization of equivalents, determination of contextual criteria for selection of equivalents, and improvement of syntactic analysis are considered as typical problems in linguistic research.

2. AN INTRODUCTION TO COMPUTATIONAL PROCEDURES IN LINGUISTIC RESEARCH
D. G. Hays
(Unsponsored)

The mathematician begins with explicitly formulated rules or conventions and works out their consequences, which are the sentences of his formal language. The native speaker of a natural language acquires a set of conventions by overt and covert learning; the linguist's task is to discover the covert conventions. Three methods of study are the distributional, the form-meaning, and the psychological; it is widely believed that the three methods should lead to the same results, but there are not yet proofs for this conjecture.

A method with posteditors is of the form-meaning type. Technical problems (worksheet formats, etc.) are discussed, especially for the case of multiple structures per sentence.
Analysis can use computers merely for making concordances, or for deeper parts of the work. Cross-tabulations, e.g., of inflectional suffix of dependent by inflectional suffix of governor, are suggested, with techniques for analysis of these frequency matrices.

Methods of fully automatic linguistic analysis have been proposed. Some suggestions of Harris, Garvin, and Lamb are reviewed, and a needed procedure for determining dependencies is offered.

In conclusion, the notion of extrapolation from text to language is examined.

3. STUDIES IN MACHINE TRANSLATION--2: RESEARCH METHODOLOGY
H. P. Edmundson and D. G. Hays
RM-2060, December 1957, 24 pp (out of print)
AD 150 666
(Also published as P-1251, same date)

The procedure is cyclic. In each cycle, the new text is compared with an existing glossary, and items in the text but not in the glossary are added to it; sentence-structure determination is attempted with an existing grammar, posteditors correct errors, and the errors are analyzed for the sake of improving the grammar; translation is attempted, postedited, and analyzed for the sake of improving the translation system.

4. ON THE VALUE OF A DEPENDENCY CONNECTION
D. G. Hays
AD 257 280
(Air Force Office of Scientific Research; AFOSR-TN-150)

In sentence-structure determination, values are numbers assigned to types of syntactic relations in such a way that connections of higher value are established in preference to connections of lower value. A given text in which sentence structures are known (as by post-editing), the values of some syntactic relations can be estimated by the following plan: assign value 1 to a relation provided no relation is known to have lower value; assign value 2 to a relation provided all relations known to have lower value are also known to have value 1; etc. The same procedure can be used for assigning adjectives to order classes, and for other similar purposes.

5. STUDIES IN MACHINE TRANSLATION--8: MANUAL FOR POSTEDITING RUSSIAN TEXT
H. P. Edmundson, K. E. Harper, D. G. Hays, and B. J. Scott
(Also published as P-1624, November 1959)
AD 249 679
Mechanical Translation, vol. 6 (November 1961), pp. 55-71.

Instructions for (i) choice of English equivalents, (ii) marks of English inflections, insertions, etc., and (iii) indication of dependency structure, all relative to a specific worksheet format and coding scheme. Includes some semantic and morphological suggestions about dependency analysis.

6. SOME MONTE CARLO ESTIMATES OF THE YULE DISTRIBUTION
H. A. Simon and T. A. Van Hornen
AD 604 799

Some Monte Carlo simulations of the Yule process, viewed as a stochastic process for generating the frequency distributions of words in text. Part I presents data for the case that cannot be solved analytically in closed form, where the rate at which new words enter the text is variable. Part II compares Monte Carlo simulations with two sets of data in which both the empirical frequency distributions and the actual rate at which words enter the text are known.

7. THE LOGIC OF COGNATE RECOGNITION IN HISTORICAL LINGUISTICS
M. Kay
AD 605 823

This paper presents a formalization of one step in the comparative method, that in which modern derivatives of prehistoric phonemes are recognized. The basic assumption is that the words of a hypothetic prehistoric language should be constructed in such a way that the total number of phonemes in the language and of the statements that need to be made to account for the forms of the modern words. The theory is sufficiently specific to provide an algorithm for a computer program and provides a basis on which more efficient heuristic procedures might be built.

8. LANGUAGE IDENTIFICATION IN THE LIMIT
E. N. Gold
RM-4136-PR, July 1964, 40 pp.
AD 602 071

Information concerning a language is presented by a text, an unending succession of examples of strings in the language, or by informant, a source that tells whether or not successive strings are in the language. Decision procedures
generate successive guesses as to the identity of the language, which is said to be identified in the limit if these guesses are the same and correct after some finite time. Six variations of each of the two basic identification situations are considered. It is found that identification from informant is powerful enough to identify in the limit primitive recursive languages, which include context-free languages. Use of textual information, however, is so weak that, except for one special case, which appears to have no practical implications, not even regular languages are limiting identifiable.

9. THE TERMINATION OF CERTAIN ITERATIVE PROCESSES
R. M. Needham
AD 641 667
This study shows that a certain class of iterative processes must terminate. The processes in question arise mostly in automatic classification, and are used for finding clusters of points or objects in a wide variety of spaces.

10. AUTOMATIC CLASSIFICATION IN LINGUISTICS
R. M. Needham
AD 644 011
Prepared for publication in The Statistician, this paper emphasizes the difference between methods appropriate for the classification of words and the usual forms of numerical taxonomy. Because words often have more than one meaning, linguistic classifications are overlapping. Techniques that yield a classificatory hierarchy are rankly inappropriate. The basic problem is not whether the right data have been chosen but how to handle the very large samples necessary. Randomization and construction of similarity matrices are likely to be unproductive. Independent searches for separate classes seem to work best. A nonmatrix technique employing a cohesion function and search for class profiles is outlined, which usually requires less than ten iterations for termination.

4. Computation
4.1. Codes and Formats

1. STUDIES IN MACHINE TRANSLATION--5:
MANUAL FOR KEYPUNCHING RUSSIAN SCIENTIFIC TEXT
RM-2061, December 1957, 10 pp.
AD 150 668
Card layout for punching one form occurrence per card; keyboard layout for Cyrillic; punctuation code; and code for non-Cyrillic.

2. STUDIES IN MACHINE TRANSLATION--3:
RESUME OF MACHINE CODES AND CARD FORMATS
H. P. Edmundson, D. G. Hays, and R. I. Sutton
(Also published as P-1352, same date)
AD 156 048
Codes for the Cyrillic alphabet, for text location, for punctuation, for non-Cyrillic matter in text, for glossary location, for English equivalents, etc. Card formats for text and glossary, assuming the use of punched-card equipment for many steps in data processing.

3. STUDIES IN MACHINE TRANSLATION--4:
MANUAL FOR PRE-EDITING RUSSIAN SCIENTIFIC TEXT
H. P. Edmundson, D. G. Hays, E. K. Renner, and D. V. Mohr
AD 259 231
Instructions for sequence numbering and classification of non-Cyrillic material. An elaborated code for non-Cyrillic material is included, and full-line-per-card punching is assumed.

4. STUDIES IN MACHINE TRANSLATION--6:
MANUAL FOR CODING RUSSIAN GRAMMAR
K. E. Harper, D. G. Hays, and D. V. Mohr
AD 253 635
Part-of-speech, case, number, gender, tense, case requirements of cardinal numbers, nouns, adjectives, adverbs, verbs, and participles, functions of short-form adjectives, aspect, degree, reflexivity, mood, person, identification of particles, prepositions, and conjunctions, classification of non-Cyrillic forms, and special classification with respect to prepositional relationships.

5. NATURAL LANGUAGE IN COMPUTER FORM
M. Kay and T. W. Ziehe
AD 610 527

This Memorandum describes a scheme for recording text in computer-readable form in such a way that all meaningful typographical distinctions are represented in a standard way, provision is made for texts in different languages and different alphabets and for subsidiary material such as parallel translations and comments of interest to users and librarians. The basic set of encoding conventions is indefinitely extensible to accommodate new kinds of material.

Very large bodies of data require special facilities, and these have been provided.
by embedding the text-encoding scheme in
a general file-maintenance system. This
provides for a comprehensive set of
labels for different-sized units of text
and makes it easy to retrieve any given
unit. It also provides means of correct-
ing and revising material in the file.

It is expected that files of computer-
usable text will be built up in a variety
of ways and, in particular, that key-
punching of text for this express purpose
will become steadily less important. Com-
puter programs are described which simpli-
fy conversion of text from these various
sources into the standard format.

The final section discusses the problem
of printing text which has been recorded
in the standard format and describes a
flexible program for doing this.

6. ACQUISITION, ARCHIVING AND INTER-
CHANGE
D. G. Hays
AD 644 559

"Acquisition" is used to mean the con-
version of information from a form de-
signed for human consumption to one de-
signed for machine input. This can be
done using a standard keyboard device
such as a typewriter (perhaps slightly
modified, perhaps with a variable char-
acter set). Many different kinds of
printed material can be encoded in ways
that retain the important typographical
distinctions, if the encoding conventions
used for each item are stored with the
item to enable it to be printed out for
reading later. The RAND Text and Catalog
System for the machinable archiving of
printed text is described, and a format
of still greater generality is outlined.
A common encoding and formatting scheme
is necessary to permit interchange of
materials between archives, and it must
be such as to allow the publication of
high-quality extracts.
(See also RM-4590-PR, RM-4540-PR)

4.2 Programs & Systems

1. THE NATURE OF DATA IN LANGUAGE ANAL-
YSIS
T. W. Ziehe and S. L. Marks

Data used for and resulting from the
analysis of natural language exhibit many
different relationships; in the files
that must be maintained for linguistic
research, both the data and the relation-
ships are subject to frequent modifica-
tions. Ziehe and Marks describe a gener-
ralized notation for the relationships,
making it possible to treat the relation-
ships themselves as data on which comput-
er routines can operate without adapta-
tion as the relationships change.

2. MIMIC: A TRANSLATOR FOR ENGLISH
CODING
H. S. Kelly
P-1936, March 1960
Int: H. P. Edmundson, ed., Proceedings
of the National Symposium on Machine
Translation, Prentice-Hall, Englewood

The translator and interpreter described
here accept instructions, of limited
variety but English-like form, simpli-
ifying the task of the linguist who must
write many short programs for analysis,
translation tests, and so on, all of which
each after a small amount of data has
been processed. The system is designed
to work on structured data, e.g., input
text with grammar-code symbols, sentence-
structure descriptions, and tentative
equivalents attached.

3. GLOSSARY LOOKUP MADE EASY
H. S. Kelly and T. W. Ziehe
Int: H. P. Edmundson, ed., Proceedings
of the National Symposium on Machine
Translation, Prentice-Hall, Englewood

A glossary of forms occurring in a block
of text is compiled by hash addressing.
With this glossary in core storage, the
glossary is read in two parts. In the
first part, each item is compared with
the contents of core storage, and a
record of the items matched is compiled.
In the second part, information about
the matched items is selected by means
of the record previously constructed
and transferred to core, replacing the
items originally there. Finally, this
information is written out on tape in
occurrence order in preparation for the
next linguistic operation.

4. STUDIES IN MACHINE TRANSLATION--10:
RUSSIAN SENTENCE-STRUCTURE DETERMINA-
TION
D. G. Hays and T. W. Ziehe
RM-2558, April 1960, 78 pp.
AD 238 090

Compared with other programs available
at the time, this one was relatively
simple and easy to transfer from one
problem (i.e., input language) to another.
Dependency theory is sketched, and the
isolation of word-order rules of prece-
dence from all other grammatical rules
is explained. Agreement tests are made
using a table, with certain complexities
to save space. Resultant grammatic
interests are the altered descriptions
assigned to units when dependency connec-
tions are made. The special cases of
conjunctions (which the program could
handle), of ellipsis (for which a sub-
routine was to be written), and of rela-
tive clauses (which could be handled)
are outlined in part. Punctuation and
idiom recognition are mentioned.
The routine consists of 2400 instruction
words, running at about 650 text occu-
rances per minute on the IBM 704. Prob-
lems of structure revision--i.e., of
backtracking when a partially completed
structure is found to contain an error--
are considered.
5. CONNECTABILITY CALCULATIONS, SYNTACTIC FUNCTIONS, AND RUSSIAN SYNTAX
D. G. Hays
(Paper completed at EURATOM
Mechanical Translation, vol. 8, no. 1, [August 1964], pp. 32-31.
Code matching is an alternative to table lookup in certain grammatical agreement. This plan requires elaborate descriptions of individual items (e.g., the words in a dictionary) but it avoids the use of large tables or complex routines for the tests. Development of the technique also leads to one clarification of the linguistic concepts of functions, exocentrism, and homography. A format for the description of Russian forms and a program for testing connectability by matching are described. Nine functions are recognized: subjective; first; second; and third complementary; first; second; and third auxiliary; modifying; and predicative. These are the dominative functions; another program must still be written for the coordinative functions: coordination, apposition, etc.

6. TRANSLATION OF ARTIFICIAL LANGUAGES BY COMPILER PROGRAMS
R. F. Rosin
P-1771, September 1959, 15 pp.
A program was written in GAT, the language of a compiler used at the University of Michigan, to translate FORTRAN programs to GAT programs. Rosin suggests several possible specifications that might be used for future translations of both artificial and natural languages.

7. EXPERIMENTS WITH A HEURISTIC COMPILER
H. A. Simon
The experiments were performed during the construction of a compiler using heuristic problem-solving techniques such as those incorporated in the General Problem Solver. The experiments attempt to discover the problems of constructing more powerful programming languages and compilers and to test whether the task of writing a computer program can be considered a "problem" in the sense defined by the General Problem Solver.

8. A PARSING PROGRAM FOR CATEGORIAL GRAMMARS
M. Kay
AD 605 822
A parsing procedure which finds all the structures of a sentence allowed by a categorial grammar has been written for the IBM 7090 computer. High speed and economy of storage of grammatical tags is achieved by first putting grammatical tags into parenthesis-free form and then compiling them into lists of computer instructions.

9. PALL: RAND'S AUTOMATED ADDRESS BOOK
G. E. Bryan
AD 605 456
The PALL program ("Prints Address Lists and Labels") has been in use since early 1963 to produce the various pieces of paper - distribution lists, receipts, mailing labels, and order forms - required in the distribution of RAND publications. It maintains a master file of some 850 organizations approved to receive various categories of RAND reports. In design and operation, the program resembles a two-pass assembler and makes extensive use of modern string-handling, pointer, and dictionary techniques. Section I is a brief discussion of the origin, operation, and advantages of PALL. Section II contains instructions for users of the program, Section III, instructions for operators, Section IV describes program operations, and Section V discusses program formats.

10. ENDOCENTRIC CONSTRUCTIONS AND THE COCKE PARSING LOGIC
J. J. Robinson
AD 614 890
Cocke's Parsing Logic is described with a flowchart. When it is applied to a string like "All the old men on the corner," 16 partial parsings result. Robinson notes that the rules and parsing logic can be made to attach all modifiers on one side before beginning with those on the other side of the head. She gives some other methods tending to avoid waste of computing effort during parsing.

11. LARGE FILES IN LINGUISTIC COMPUTING
M. Kay
AD 615 301
The text and catalog system more fully introduced in RM-4390-PR (the entry 4.1-5 above) is here described briefly and exemplified with applications to a Japanese-English dictionary and Japanese text. In particular, a text encoding scheme for the set of idiograms adapted from Chinese and used in Japanese is presented.

12. THE CATALOG: A FLEXIBLE DATA STRUCTURE FOR MAGNETIC TAPE
M. Kay and T. W. Ziehe
AD 623 938
This Memorandum outlines a generalized storage scheme for large files of highly structured data, or catalogs, and describes their realization on magnetic tape. Each datum, which may be large or small, is assigned to one of a number of data classes of which a user may define any number. The overall organization of a catalog is given by a map which imposes a tree structure on the set of data classes. Catalogs may participate as individual data in other catalogs so that the structure of a file may be recursive. A flexible addressing scheme not only facilitates the retrieval of data and sets of data in response to
11.

13. **THE CATALOG INPUT/OUTPUT SYSTEM**
   - M. Kay, F. Valadez, and T. W. Ziehe

This Memorandum completely defines the format used for catalogs on magnetic tape and describes in detail the routines of the Catalog Input/Output System. Catalog maps, catalog data, and tape labels are written as logical records in a specially designed blocking format. Beginnings and ends of blocks, physical tapes, and catalogs are explicitly marked. The Catalog Input/Output System offers a variety of input/output unit-control operations in addition to the commands for reading and writing data in catalog format. The reading and writing commands are implemented on three levels: Level III handles individual data; Level II reads and writes logical records; and Level I processes blocks of information. Anyone planning to use the system should note that the structure of levels II and III is ordinarily not of interest to the user. Only the simpler Level III structure reflects the user's view of his file in most applications. The Catalog Input/Output System is not intended to supplant specialized programs for extracting the contents of data and analyzing them in accordance with the user's special problems. For this reason, the user will ordinarily attempt to put one unanalyzable item of information into each datum, and prepare special programs, as few as possible, for content processing.

14. **THE TABULAR PARSER: A PARSING PROGRAM FOR PHRASE STRUCTURE**
   - M. Kay
   - AD 635 934

A parsing program for the IBM 7040/44 computers is described which finds all phrase and dependency structures assigned to a sentence by a given grammar. The grammar is prepared for the machine in the form of a set of rule tables which constitute a rich, terse, and perspicious grammatical notation.

15. **COMPUTER ROUTINES TO READ NATURAL TEXT WITH COMPLEX FORMATS**
   - P. Graves, D. G. Hays, M. Kay, and T. W. Ziehe
   - AD 637 303

This Memoranda describes a set of subroutines for the IBM 7040/44 computers for reading textual material with complex formats and coding conventions—questions, libraries, catalog cards, etc.—from any external medium into the high-speed store of the machine. Different kinds of information in the input are recognized by explicit markers, position on the line or page, or syntactic clues given by other items. Less complex material requires only a portion of the system. Information may be recorded according to the user's conventions before being delivered to his program. The routines may be called from either FORTRAN- or MAP-coded programs.

16. **THE CATALOG: A FLEXIBLE STRUCTURE FOR DATA STORAGE**
   - T. W. Ziehe
   - P-3470, November 1966, 22 pp.
   - AD 642 368

An over-all view of the RAND text and catalog system described in RM-4390-PR for processing natural language text for computer storage in large files is presented. The system provides freedom to reorganize the files at will. Organization is based on the dependency tree. Data within any one class may be encoded in any form, disregarding the rest of the file, and the nature of the coding is indicated on the catalog map (index). Catalogs can be nested, and each catalog may be structured separately. A valuable space-saving device is the null datum, in which headings without any data can fulfill their organizing role without occupying any physical space. Computer programs convert text into its internal representation. Formats have been defined for 7-bit magnetic tape and 36-bit-word core storage, and others will be added. Catalog transformation schemes are also planned for the future. (Presented at the Conference on Computers in Humanistic Research, Texas A. & M., November 1966.)

17. **COLLECT: A PROGRAM FOR THE RETRIEVAL OF GRAMMATICAL INFORMATION FROM ANNOTATED TEXT**
   - M. Kay and T. D. Taft

COLLECT is a system for searching a File of Russian sentences annotated with syntactic (and, eventually, morphological) information. The File consists of sentences; the user writes a description of the class of sentences he wishes to examine, and COLLECT edits his request, searches the File (taking several requests simultaneously, if desired), and prints a report. A sentence description consists of descriptions of word-form occurrences— their spellings, grammatical properties, etc.— and their relationships to one another. Small units of descriptions can be composed by conjunction or disjunction, and can be negated.

Although the system was written for this specific File and the kinds of searches linguists can be expected to make, it is flexible enough to permit addition of new kinds of information to the File, enrichment of the request terminology allowed, and modification of report formats. It is also representative of a class of retrieval programs needed for many file-search tasks.

The COLLECT system has been written for the IBM 7044 computer.
4.3. Program Languages

1. THE EVOLUTION OF CONCEPTS AND LANGUAGES OF COMPUTING
   R. D. Elbourn and W. H. Ware
   A review of the evolution of programming languages from the time when all
   programming was done in machine languages, through symbolic coding systems, interpreters, assemblers, generators, and
   compilers, to the recently developed list-processing languages. These lan-
   guages are then applied to game playing, problem solving, and theorem proving.
   Behavior and biological modeling are described. Finally, in anticipation of
   extending the capability of computers to accept, use, and generate natural
   languages, this paper concludes with an introduction to some of the contemporary
   work on formal language theory, including a discussion of six families of abstract
   languages and their practical implementation.

2. AN EXPERIMENTAL SYNTAX-DIRECTED DATA STRUCTURE LANGUAGE
   R. E. Lindsay, T. W. Pratt, and K. M. Shavor
   P-3112, April 1965, 48 pp.
   AD 614 782
   Programmers developing systems of the complexity required in artificial intelli-
   gence research are frequently hindered by the rigid programming languages
   available and the time-consuming task of implementing new languages. AMOS
   (for associative memory organizing system) provides a flexible means to struc-
   ture data and experiment with the syntactic forms of program statements while
   lessening the implementation bottleneck. AMOS is a syntax-directed compiler used
   to define languages for constructing a variety of data organizations of which
   FORTRAN-like arrays and IPL-like list structures are special cases. This re-
   search explores the use of syntactic descriptions which are not Backus-Normal
   Form grammars and provides means for defining two-dimensional languages as
   well as the usual linear type. In order to facilitate implementation, the system
   may be conveniently imbedded in any monitor system of common design; AMOS
   operations are manipulations within high-speed storage only.

5. The Russian Language

5.1. Text and Glossaries

1. STUDIES IN MACHINE TRANSLATION: BIB-
   LIOGRAPHY OF RUSSIAN SCIENTIFIC
   ARTICLES
   A catalogue of Russian text available on
   magnetic tape at RAND. Thirty-one corpora
   are included:

<table>
<thead>
<tr>
<th>Subject</th>
<th>Corpora</th>
<th>Pages</th>
<th>Occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physics</td>
<td>12</td>
<td>1507</td>
<td>373,329</td>
</tr>
<tr>
<td>Mathematics</td>
<td>4</td>
<td>471</td>
<td>108,124</td>
</tr>
<tr>
<td>Astrobiology</td>
<td>8</td>
<td>936</td>
<td>260,168</td>
</tr>
<tr>
<td>Cybernetics</td>
<td>6</td>
<td>645</td>
<td>208,565</td>
</tr>
<tr>
<td>Physiology</td>
<td>1</td>
<td>378</td>
<td>111,530</td>
</tr>
<tr>
<td>Total</td>
<td>31</td>
<td>3937</td>
<td>1,001,736</td>
</tr>
</tbody>
</table>

   The physiology text was keyphrased by a
   research project at Georgetown Univer-
   sity several years ago. For each article, the
   bibliography gives title, author, and
   citation, together with length (in pages and occurrences) and RAND identification
   number.

2. A GLOSSARY OF RUSSIAN PHYSICS ON
   PUNCHED CARDS
   K. E. Harper, D. G. Hays, and A.
   Koutoussas
   P-1241, December 1957
   Mechanical Translation, vol. 4, no.
   In 1957, a glossary of 6,000 Russian
   forms, prepared cooperatively by the
   University of Michigan and RAND, was
   offered to qualified researchers.

3. STUDIES IN MACHINE TRANSLATION--12:
   A GLOSSARY OF RUSSIAN PHYSICS
   A. S. Kozak, C. H. Smith, and members
   of the RAND ALDP group
   AD 252 609
   A list of he 23,146 distinct forms of
   6,847 words that occur in 265,417 occur-
   rences of Russian physics text. With
   each form are shown the English equiva-
   lents, grammatic properties, and occur-
   rence frequency derived from the text
   by the RAND cyclic procedure. The forms
   first observed in the last batch of
   30,000 occurrences are included, but no
   postediting results from that corpus are
   embodied in the equivalents and
   grammar-code symbols.

4. HIGH FREQUENCY WORDS AND OCCURRENCE
   FORMS IN RUSSIAN PHYSICS
   A. S. Kozak
   AD 288 133
   This report lists the 1026 highest-fre-
   quency words and the 1013 highest-fre-
   quency forms in the RAND physics glos-
   sary. Each collection is listed by frequency and also alphabetically. (Non-
   Cyrillic items are excluded.) Comparison
   with Jossett's Russian word count,
   based on general prose, shows that 150
   of his 244 most common words do not
   appear among the 244 most common words
   in the RAND text. Moreover, 30 of these
   words do not occur in the quarter-million
   occurrence sample of physics text. The
   1026 words have frequency 35 or higher;
   their total frequency in the text is
202,115, or 86.30% of the occurrences.
The 1013 forms also have frequency 35 or higher, but their total frequency is only 146,308, or 62.47% of all occurrences.

5. A FILE OF RUSSIAN TEXT WITH SYNTACTIC ANNOTATIONS
D. G. Hays, D. B. Gottshall and A. S. Kozak
(Rome Air Development Center)
The file described here is a corpus of more than a million running words of Russian text, mostly coming from the journal literature of physics, mathematics, astrobiology, and physiology, but syntactic examples from a large grammar are included to provide variety for syntactic and possibly semantic studies.
All sentences have been annotated with syntactic structures, using the notation of dependency theory. Each sentence is shown to have an independent word; other words serve specified functions for specified governors. The file is recorded on magnetic tape as a catalog.

5.2. Syntax

1. ORDER OF SUBJECT AND OBJECT IN SCIENTIFIC RUSSIAN WHEN OTHER DIFFERENTIALS ARE LACKING
D. G. Hays
P-1632, March 1959, 6 pp.
Mechanical Translation, vol. 5, no. 3 (December 1958), pp. 111-113.
In a sample of 22,000 occurrences of Russian physics text, 50 instances of true ambiguity were found; in these instances, no morphological properties of subject or object (tense word or modifi-
ers) reduced the ambiguity of either. In these instances, the subject was always after the verb, and the four cases of subject-after-verb seemed idiomatic, involving frozen combinations of verb and object.

2. A RUSSIAN STRUCTURE FOR COMPARISON
D. G. Hays and B. J. Scott
P-1720, June 1959
Three fixed combinations, tot she...chto (which she...what), and takoj she...chto (which one...what), are used to draw comparisons between objects, properties, or circumstances. Using postedited text (120,000 occurrences), Hays and Scott found 27 sentences containing the constructions, and posited three transformations to account for their use. Techniques for identification of the transformations and points of application are proposed.

3. GOVERNORS OF THE CONJUNCTION CHTO
J. H. Pustula
The Russian subordinate conjunction chto (that) introduces noun clauses. In the RAND corpus of about 249,000 occurrences, 51,400 clauses with chto were governed by 85 different verbs, present participles, short-form past participles and adjectives, and nouns, and by 4 idioms. These governors are listed, each with frequency, and classified.

4. SOME LINGUISTIC PROBLEMS OF RUSSIAN GRAPHIC ABBREVIATIONS
W. A. Stewart
Abbreviations are mostly undeclinable, and there is the danger that the morphologi-
cal indicators generally useful in sentence-structure determination. They are sometimes reductions of single words, sometimes of phrases. They exhibit the structural features of the natural language in which they are used, and present more or less the same problems, with the above-noted exceptions.

5. STUDIES IN RUSSIAN MORPHOLOGY--I: THE SUFFIX -AGA
D. S. Worrilow
RN-3235-PR, August 1962, 45 pp.
AD 281 850
About two dozen words in contemporary standard Russian are formed by adding the suffix -aga to nominal, adjectival, and verbal stems. The report considers the formation of derivational stems, the conso-
antal and vocalic alternations which occur when the suffix is added, and the gender of the derived word. Alternations in meaning are discussed, and Worth compares these items with another dozen ending in -aga but not with the suffix (i.e., the ending is homophonous with the suffix). He asserts that the two classes are brought together by speakers of the language, and treated by them as falling into two classes: (i) the referent is usually human and the suffix is expressive, sometime pejorative, sometimes favorable; (ii) the referent is usually nonhuman and the suffix is not expressive.

6. STUDIES ON PREDICATION IN RUSSIAN--II: PREDICATIVE CASE, SHORT FORM ADJEC-
TIVES AND PREDICATIVES
H. Birnbaum
AD 429 524
This is the first, theoretical part of a series of studies on certain aspects of predication (predicative function) and its formal expression in Contemporary Stan-
dard Russian (CSR). This series is pri-
mary concerned with such instances where in Russian the predicate is deno-
ted by means other than a finite verb, i.e., by infinite verbal forms (infinitives and gerunds) as well as nonverbal forms, including the zero morpheme of the copula verb. In its sequel the series makes use of automatic and semi-
automatic procedures applied on corpora of analyzed Russian physics texts avail-
able at the RAND Linguistic Research Project, in addition to evidence from other Russian linguistic materials.
7. COMPLEMENTATION IN RUSSIAN:
   THEORY AND APPLICATION
   A. S. Kozak
   AD 620 932
   Kozak reviews the concept of complementation in Russian, from both the Soviet
   point of view and that adopted at RAND for compilation of a file of scientific
   Russian text with syntactic annotations. Models for different patterns of comple-
   mentation are presented.

8. A DEEP INDEX OF DERIVATIONAL MOR-
   PHOLOGY
   D. S. Worth
   AD 640 962
   A projected "Deep Index" designed to gather, file and retrieve, with the as-
   sistance of the "catalog" system, a large quantity of information concerning the
   extant literature on the derivational morphology of Russian is described.

   The Memorandum describes the type of literature being abstracted, the class-
   ification of information into discrete categories for retrieval, the steps taken
   to prepare data for conversion to machine format, and the transliteration and ab-
   breviation conventions used in the index. An example of Russian scholarly text is
   appended, together with the format in which this text appears after processing.

9. TOWARD EXPLOITATION OF A FILE OF
   RUSSIAN TEXT WITH SYNTACTIC
   ANNOTATIONS
   D. G. Hays and D. S. Worth
   (Rome Air Development Center)
   After summary descriptions of a File of Russian text with syntactic annota-
   tions and the systems of programs called COLLECT usable in searching the File, this
   Memorandum discusses methods of automatic (statistical) classification usable in re-
   ducing data obtained from the File, the use of parsing programs in linguistic research,
   and two of the many kinds of syntactic questions that might be answered by research
   based on the File: modal constructions and sentential apposition.

10. STUDIES IN RUSSIAN MORPHOLOGY: II.
    VOWEL-ZERO ALTERNATIONS IN DERIVATION
    D. S. Worth
    In one type of vowel-zero alternation, a flexional-level vowel-zero morphopheme
    is stabilized as either a full vowel or a zero in the course of derivation. In an-
    other type, a full vowel in the derivational base alternates with zero in the de-
    rived word, or vice versa. Worth suggests the existence of a vowel-zero morphopheme
    on the derivational level as an explanation for certain of these alternations. As
    background, he sketches a generative framework within which to consider specific
    techniques of word formation, and briefly surveys flexional vowel-zero alternations.

5.3. Semantics

1. PAIRS OF RUSSIAN WORDS WITH HIGH COR-
   RELATION
   D. G. Hays
   Using the material prepared by H. H.
   Josselson for The Russian Word Count,
   co-occurrence (on the same page) was
   tallied for every possible pair of words
   on a list of 2000 common words in 4037
   pages of text. The 2731 most highly
   correlated pairs are listed.

2. PROCEDURES FOR THE DETERMINATION OF
   DISTRIBUTIONAL CLASSES
   K. E. Harper
   AD 257 287
   (Air Force Office of Scientific Re-
   search; AFOSR-TN-149)
   In: 1961 International Conference on
   Machine Translation Languages and
   Applied Language Analysis, Her Majes-
   ty's Stationery Office, London, 1962,
   pp. 687-700.
   A distributional class is a collection of
   words each of which can bear a specified
   syntactic relation to any member of some
   other set. The second set used in the
   definition may be a morphological or
   syntactic class, or one formed a priori,
   that is, an intuitively relevant class-
   such as the class of "abstract" nouns, or
   hence, perhaps, a semantic class. Some
   distributional classes are examined: that
   of verbs having only animate actor nouns
   as dependents, that of governors of nouns
   naming physical particles, and that of
   governors of rezkij (sharp). Several
   problems raised by the procedure are
   considered.

3. TRANSFORMATIONAL CRITERIA FOR THE
   CLASSIFICATION OF PREDICATIVE GENITIVE
   CONSTRUCTIONS IN RUSSIAN
   D. S. Worth
   AD 257 288
   (Air Force Office of Scientific Re-
   search; AFOSR-TN-148)
   In: 1961 International Conference on
   Machine Translation Languages and
   Applied Language Analysis, Her Majes-
   ty's Stationery Office, London, 1962,
   pp. 725-737.
   Some Russian constructions containing
   genitive substantives must also include a
   modifier, such as an adjective or a
   substantive. For example, "Sapog bol'shogo razmera" ("A boot of large size")
   cannot be reduced to "Sapog razmera" ("A boot of size"). Several
   varieties of such constructions are de-
   termined by transformational possibili-
   ties; Worth finds that every variety
   contains implies a predication: "A
   boot, of which the size is large."

4. MACHINE TRANSLATION OF RUSSIAN PREPO-
   SITIONS
   K. E. Harper
   Using postedited text, Harper classified
   the governors and dependents of Russian
prepositions according to their effect on translation of the preposition. Some
idioms are recognized (e.g., v zaklyuchenii - in conclusion), and some fixed
combinations (e.g., sostoyat' v (to consist of)) that cannot be called idioms in
the narrow sense because word order is not fixed. 42 prepositions that occurred
about 650 times in about 240,000 occurrences are equivalent each. 12
prepositions, occurring about 12,000 times in about 120,000 occurrences, could
be translated accurately in the sample text upon identification of 60 fixed
combinations and a dozen word classes.

5. THE POSITION OF PREPOSITIONAL PHRASES IN RUSSIAN

K. E. Harper
AD 405 501

In automatic sentence-structure determination, selection of the governors of
prepositional phrases is reported as exceedingly difficult. Harper first sets
aside combinations in which the preposition is strongly governed, i.e., those
in which there is no special relation between the preposition and its governor;
these combinations can be identified, and the preposition is almost always depend-
don the corresponding governor when the two occur together in a clause—but
only about 1 occurrence in 5 is strongly governed. Examining various subsamples
of the RAND-parsed Russian physics text, he finds: (i) Verbs and nouns are the
most frequent governors of prepositions, adjectives about 1/10, other parts of
speech practically not at all. (ii) When a noun and a verb compete for the
preposition, the noun itself depending on the verb, then (a) projectivity makes
the preposition depend on the nearest element in N-V-P and P-V-N, (b) by ob-
servation, the preposition almost always depends on the nearest verbs in P-N-V
and V-P-N, and the exceptions are syntactically recognizable, and (c) in the order N-P-V
or V-N-P, the noun governs the preposi-
tion in about 90% of the occurrences.

(iii) In case (c) above, when the verb
governor, usually either the preposi-
tional phrase is recognizably adverbial
(e.g., in reality), or the noun that
does not govern the preposition is in an
adverbial construction (e.g., itself the
object of an adverbial prepositional
phrase). There remains for further study
the case of a prepositional phrase fol-
lowing a sequence of nouns, i.e., N-N...
-N-P, where each noun depends on the one
before.

6. PREPOSITIONAL PHRASES AND AUTOMATIC
PARSING

K. E. Harper
AD 615 319

A concordance was prepared for machine-
processed Russian text, showing the syn-
tactic governors of prepositional
phrases when the word preceding the
preposition was a noun (P-1-N). The concor-
dance showed that under this condition,
Russian prepositions strongly tend to
serve either an adnominal function or an
adverbial function. These properties are
not inherent in the individual prepo-
sition, but are environmentally deter-
mined. The resultant groupings of prepo-
sitions is presented as information
that can be used in a program for sentence-
structure determination.

7. MEASUREMENT OF SIMILARITY BETWEEN
NOUNS

K. E. Harper
AD 615 319

A study was made of the degree of "simi-
larity" between pairs of Russian nouns,
as expressed by their tendency to occur in
sentences with identical syntactic roles.
A similarity matrix was prepared for forty
nouns; for each pair of nouns, the num-
ber of shared (i) adjectival dependents,
(ii) noun dependents, and (iii) noun
governors was automatically retrieved
from machine-processed text. The simi-
larly coefficient for each pair was
defined to be the ratio of the total of
such shared words to the product of the
frequencies of the two nouns in the text.
The 780 pairs were ranked according to
this coefficient. The text comprised
120,000 running words of physics text
processed at The RAND Corporation; the
frequencies of occurrence of the forty
nouns in this text ranged from 42 to 328.

The results suggest that the sample of
text is of sufficient size to be useful for
the intended purpose. Many noun
pairs in this similarity matrix (synonymy,
antonymy, derivation from distribution-
ally similar verbs, etc.) are charac-
terized by high similarity coefficients;
the converse is not observed. The role
of various syntactic relationships as
criteria for measurement is discussed.

8. STUDIES IN INTER-SENTENCE
CONNECTION

K. E. Harper
AD 626 572

Portions of the Russian physics text
processed at The RAND Corporation were
subjected to systematic analysis to
determine the extent of repetition in
adjacent sentences. Recurrence of
the words in all pairs of contiguous
sentences in the text (2,467 pairs)
was recorded in a machine printout;
sentence pairs for which word recurrence
was not automatically recorded were
visually inspected for other types of
recurrence (through lexical stems, pro-
nouns, synonyms, and paraphrases). The
extent of the different types of recur-
rence is reported, and features of the
recurring items are discussed. Sentence
pairs characterized by nonrecurrence
(12 percent of the total) are examined,
and the relevance of inter-sentence re-
currence and nonrecurrence to automatic
syntactic analysis and abstracting is
suggested.
9. STUDIES ON PREDICATION IN RUSSIAN-II: ON THE PREDICATIVE USE OF THE RUSSIAN INFINITIVE
H. Birnbaum
AD 626 207

Birnbaum here continues the study begun in RM-3774 (the entry 5.2-0 above). He examines the predicative use of the Russian infinitive against the background of the entire range of semantic functions it serves in contemporary standard Russian.
In the colloquial language, the infinitive increasingly competes with the finite verb. Among the uses examined are: semi-predicative, wherein object infinitives occur in combination with the dative complement; predicative infinitives in one and two membered sentences and subordinate clause equivalent phrases; constructions with the methods of transformational analysis of these predicative constructions are briefly discussed and exemplified. The Memorandum concludes with a few suggestions about automatic translation.

10. SOME COMBINATORIAL PROPERTIES OF RUSSIAN NOUNS
K. E. Harper
AD 638 924

The study, based on 280 nouns which occurred at least 50 times in 180,000 of parsed Russian text, investigates the "valence" of nouns (tendency to occur as the syntactic governor of a noun in the genitive case). The valence of most nouns is intermediate. Factors influencing valence are: 1. a noun's position in a noun string (noun strings of equal length one or two account for almost per cent of all noun strings), 2. the potential of the dependent genitive noun to form an adjective (partial interchangeability is observed), 3. valence is reduced in certain stable lexical combinations, 4. valence is decreased in non-initial occurrences of a noun in an article. Among the nouns of high valence are derogative nouns and names of physical properties (abstract nouns), and among those of low valence are abbreviations for units of measurement, proper names, nouns employed in set phrases, and concrete nouns. Considerable variation in valence was observed for nouns of the same semantic or derivational classes, but sample studies suggest that these differences are well correlated with certain other morphological or word features. In general, the results suggest that (1) for some nouns, valence may be attributed to certain inherent semantic properties, and (2) for other nouns, valence may be highly correlated with other "contextual" features both syntactic and lexical.

6. THE ENGLISH LANGUAGE

1. PRELIMINARY CODES AND RULES FOR THE AUTOMATIC PARSING OF ENGLISH
J. J. Robinson
AD 295 651

The author relies on a computer program that obtains all possible immediate constituent structures for a sentence, given a grammar. The subject of this paper is a preliminary codification of English grammar, including the establishment of categories and rules for combinations of categories, an encoding scheme, and an analysis of the results obtained when the parsing program and grammar were applied to selected, not always simple, sentences. Over 500 condensed rules, equivalent to more than 2,500 rules written without condensation, are listed. 65 structures for 16 sentences, ranging up to 37 occurrences in length, are presented by reproduction of machine listings.

2. PARSE: A SYSTEM FOR AUTOMATIC SYNTACTIC ANALYSIS OF ENGLISH TEXT - PARTS I & II
J. J. Robinson and S. Marks
AD 621 311 and AD 621 311

This Memorandum explains PARSE, the RAND system for the automatic syntactic analysis of English sentences. The system can analyze and label the structures of a variety of sentences, including those with coordinate, subordinate, relative, indicative, and interrogative clauses.

PARSE consists of three major components:
1) a glossary of English word-forms for which grammatical rules have been designated;
2) a table of grammatical rules listing the permitted combinations of adjacent grammar codes and assigning a code to the combination; and 3) a program for applying the rules systematically and exhaustively to English text coded from the glossary. The computer program was submitted to SHARE during 1965, together with a table of rules and a glossary.

The Memorandum is bound in two parts; part I gives linguistic details and describes the programs; part II contains the glossary of word-forms.

The system employs a parsing logic, devised by John Cocke of IBM, which disengages the grammar from the routines that apply it to text, so that changes in grammar will not affect the program. This feature is not only highly desirable; it appears to be essential in view of the complex task of developing rules adequately detailed to label intricate patterns of actual texts.

The logic also applies the grammar exhaustively: every construction recognized by the grammar is displayed for every string submitted for analysis. This feature provides data for evaluating the grammar and locating the points where modifications should be made. It also reveals numerous unsuspected syntactic and semantic ambiguities in many of the sentences analyzed, and furnishes a framework for determining how sentences are understood and disambiguated through linguistic and extra-linguistic contexts.
Finally, the paper suggests an interpretation of the whole library problem as one where the request is considered as a clue on the basis of which the library system makes a concatenated statistical inference in order to provide as an output an ordered list of those documents which most probably satisfy the information needs of the user.

2. AUTOMATIC INDEXING: AN EXPERIMENTAL INQUIRY

M. E. Maron
RN-2601, August 1960, 35 pp.
(Also published as P-2180, February 1961)
AD 245 175
Journal of the Association for Computing Machinery, vol. 8, no. 3

An empirical test of a statistical technique is reported. The technique uses differential word frequencies in pre-established categories.

3. A LOGICIAN'S VIEW OF LANGUAGE DATA PROCESSING

M. E. Maron
P-2279, August 1961, 47 pp.

The nature of logic is described in relation to both natural language and the problem of language-data processing. Maron analyzes the problem of information identification and retrieval from this point of view.

4. INFORMATION RETRIEVAL: A LOOK AT THE LOGICAL FRAMEWORK AND SOME NEW CONCEPTS

M. E. Maron

Weighted index tags allow for computing a measure of the relevance of a document to an information request. Statistical techniques aid in automatically relating and associating documents on the basis of their subject content. However, further work is needed (i) to allow the machine to identify documents on the basis of its "knowledge" of the complete text and (ii) for better understanding of the processes behind intelligent problem solving.

5. PROBABILITY AND THE LIBRARY PROBLEM

M. E. Maron
Behavioral Science, vol. 8, no. 3 (July 1963), pp. 250-257.

Automatic information retrieval systems are needed because of the growing rate of scientific publication. Besides ordinary libraries, such systems can be used in patent offices, medical practice, law, military intelligence, etc. The system should process incoming documents automatically, disseminate them on standing orders, and store the content of the documents for later requests. Some machines designed for the purpose are mentioned. The problem of determining the relevance of a document to a request is analyzed probabilistically.
6. A PROPOSAL FOR THE INDIRECT RETRIEVAL OF UNPUBLISHED TECHNICAL MATERIAL
R. L. Patrick
P-20616, August 1962, 4 pp.
The author suggests that ASTIA search its Field of Interest Register if its bibliographic files fail to produce a stipulated minimum number of items in response to a given request.

7. THE SOVIET CLASSIFICATION SCHEME FOR LITERATURE
F. J. Krieger
AD 260 021
Technical and nontechnical literature are both rigorously compartmentalized in the Soviet Union. Krieger compares the national bibliographic schemes of the United States and the USSR.

8. AUTOMATIC PARSING AND FACT RETRIEVAL: A COMMENT ON GRAMMAR, PARAPHRASE, AND MEANING
J. J. Robinson
AD 432 036
The aim of automated parsing is to enable a computer to manipulate information received in the form of natural English sentences. Through the refinement of parsing grammars, insights are to be gained into the nature of meaning, as exemplified, for instance, in the creation of a paraphrase. These insights can then be applied to the development of fact retrieval systems which receive requests for information and output responses in natural English.

9. MECHANIZED DOCUMENTATION: THE LOGIC BEHIND A PROBABILISTIC INTERPRETATION
M. E. Maron
P-2898, April 1964
AD 437 781
This paper discusses the problem of information storage and retrieval in a library. The concepts of probability are required to frame the logic of the problem properly, since the transition from a user's request to the resulting retrieved documents must be schematized as an inverse probability inference. Statistical association techniques are required, for the library computer must be designed to use all the clues and inference techniques available to it.

10. THE LOGIC OF INTERROGATING A DIGITAL COMPUTER
M. E. Maron
P-3006, November 1964, 23 pp.
AD 608 369
An attempt to clarify the problem of how a computing machine must be organized to deal with language in order to respond to interrogation as if it understood meanings and relevance. The paper is divided into three parts: (1) an outline of the origins and scope of the information sciences and the impetus they have given to this search; (2) a look at two subproblems, literature searching and data retrieval, to suggest how these problems should be framed; and (3) a consideration of the question of comprehension with the aim of specifying how some aspects of knowing can be discussed in a mechanical way and related to the information-flow organization required to generate comprehension-like behavior.

11. RELATIONAL DATA FILE: A TOOL FOR MECHANIZED INFERENCE EXECUTION AND DATA RETRIEVAL
R. Levien and M. E. Maron
AD 625 409
Levien and Maron describe the background and status of a RAND project dealing with automatic data storage and retrieval. Their research emphasizes the development and testing of logical techniques for data retrieval and inference-making. These techniques are being implemented in the form of computer routines, and tested on a large corpus of factual information concerning the field of cybernetics.

A major theoretical decision controlling the design of a data retrieval system of this sort concerns the nature of the language that will be used. The Levien-Maron system differs from some other question-answering systems in that it does not store information in ordinary or natural language. They discuss their reasons for this decision, examining some of the logical, and linguistic aspects of the overall problem, and explaining the selection of an artificial language, an interpreted relational calculus, for use as the information language.

Levien and Maron summarize the theory of relations briefly, showing how relations can be represented in a computer and how some operations on given relations can be computed so as to derive new relations.

They illustrate the operation of their system with some typical data retrieval requests and how they can be executed once incoming data have been mapped into the artificial information language. It also mentions some of the statistical operations, such as correlations and trend analyses, that can be executed.

The key problems of inference are the different kinds of inference that can be made, what role they play in a data retrieval system, and what problems must be solved in order to mechanize inference-making. A key feature of the system is the ability of the user to
frame an inference schema and instruct the machine to execute it. This technique allows the man and the machine to collaborate in deriving a conclusion required to answer a given request for information.

A major part of the Memorandum is devoted to techniques for practical realization of the data file. Input (acquisition, organization, and preprocessing of data before it enters the computer) is a major problem. Resolutions of input problems of data sources, acquisition, extraction, etc., are outlined. Input aids extract data from source documents and enter them onto forms according to precise rules. The computer then translates from forms to relational language sentences. This section discusses the motivation and structure of the input forms.

Output problems are treated against the background of sources and types of output requests that the file will be expected to satisfy. The problem of communicating these requests to the file, and procedures for processing the requests, are discussed.

Techniques for storage and processing are discussed; e.g., how data will be structured in memory, the organization of dictionaries and thesauri, and the steps in the programming process.

The Memorandum concludes with a consideration of the data file as a whole, examining its principal features, application, possible extension, and relation to other data retrieval research. The authors review the question of literature searching and show how the data retrieval system embodies features that permit very effective literature searching. They summarize the hypothesis that context data of the type this system is designed to handle offer a promising approach to more effective literature searching. They present some of the next steps that will be undertaken to extend the capability of the system, and finally summarize and evaluate their work and contrast it with other activities in this field.

12. THE TRANSFORMATION OF SENTENCES FOR INFORMATION RETRIEVAL
   J. J. Robinson
   AD 674 090

Robinson discusses some of the difficulties that must be overcome before automatic retrieval of information expressed in natural language text becomes practical. The major linguistic task is to provide detailed, analytic, recognition grammars with transformational components adequate to deal with the complexities of surface structures of natural sentences.

13. RELATIONAL DATA FILE I: DESIGN PHILOSOPHY
   M. E. Maron
   AD 636 299

Maron's paper and Levien's (see 7-14) are descriptions of the Relational Data File presented at the Third Annual National Colloquium on Information Retrieval, 1966. Some of the key features of the design philosophy of the system are: information is stored in the form of millions of sentences, processed as necessary to answer questions; all information is uniformly represented in binary relational sentences in a formalized information language; the system is designed to interpret and execute inferences when the answer to a query is merely implicit in the stored data; since literature searching can be accomplished when some of the stored data is about information, sophisticated searches can be conducted through inference techniques dealing with context data.

14. RELATIONAL DATA FILE II: IMPLEMENTATION
   R. E. Levien
   AD 636 311

Related to Maron's (see 7-13), Levien's paper describes the basic features of the implementation of the system: Data are entered by means of special input forms; programs written in FOREMAN, a special form retrieval and manipulation language, derive binary relational sentences from the input tape; the computer code dictionary encodes each sentence as four computer words and full or partial copies of the sentences are entered into four separate files on disk; only subsections of each file are transferred between disk and core; the language at the user's disposal has elementary commands corresponding to the basic operations of arithmetic, set theory and logic, allowing the user to implement simple retrieval, augment and correlate, and also execute complex inference schemes.

15. A COMPUTER SYSTEM FOR INFERENCE EXECUTION AND DATA RETRIEVAL
   R. E. Levien and N. E. Maron
   AD 642 120

This Memorandum is an up-to-date brief and informal description of Relational Data File (see 7-11, 7-13 and 7-14) and the problems that arise in the implementation of such a logical analysis system and how they have been resolved in the now partly operational Relational Data File. Among the specific problems that are solved are the logical and linguistic problems of representing the basic data, file problems of organizing the data for rapid and direct access, programming problems of designing a language for file interrogation, and hardware problems of providing sufficiently large, rapid access storage.
In the Relational Data File, facts are stored in the form of sentences in an artificial information language. Each sentence expresses a binary relationship between two entities. The sentences are directly accessible by content. An intensional file that stores facts about relationships permits many data sentences to be derived when needed rather than stored explicitly. The file of sentences is organized in quadruplicate in a disk memory so that access time is minimized. A programming language has been designed that enables a user to express, in addition to direct retrieval requests, inferential processes that are required to derive implied conclusions. The system is designed for use on an IBM 7044 with a 1301 disk, which holds 232,000 sentences in the information language.

16. A FORMAL SYSTEM FOR THE LOGICAL ANALYSIS OF TEMPORAL RELATIONSHIPS BETWEEN INTERVALS OF TIME
R. Mattison

The system is an applied first-order predicate logic rich enough to express all the usual ordering relations between intervals. The formal language is a first-order language with one binary predicate (interpreted as the "wholly before" that holds between intervals) and two unary operations (interpreted as "the beginning of" and "the ending of" applied to intervals). The semantic completeness of the logical axioms is presupposed, while the syntactic completeness of the temporal axioms is proved. The author also gives a mathematical model of intervals of time which is also a model of the temporal axioms. This model and the temporal axioms are adapted to the needs of the RAND cybernetics data-research project, but philosophically more satisfying axioms are given in an appendix.

17. A SIMPLE SCHEME FOR FORMALIZING DATA RETRIEVAL REQUESTS
F. M. Tonge

The user of the cybernetics data file may be ignorant of computer jargon or the details of file organization. If he is to present data retrieval requests in a "natural" language, those requests must be translated. Tonge gives a simple procedure using rules that depend upon the position of words in the request and simple semantic information about words. It presents its translation to the user for modification or correction prior to file search. Tonge gives the rules and several illustrations, and mentions some deficiencies of the present scheme.

8. Content Analysis

1. AUTOMATIC LANGUAGE-DATA PROCESSING IN SOCIOLOGY
D. G. Hays

Basic processes of ALDP are described briefly. Techniques based on Osgood's "evaluative assertion analysis" are outlined. The need for research on the relations between linguistic and social variables is emphasized.

2. AUTOMATIC CONTENT ANALYSIS: SOME ENTRIES FOR A TRANSFORMATION CATALOG
D. G. Hays

Automatic content analysis would be possible if three steps could be performed automatically: (i) Sentence-structure determination, (ii) structure simplification and regularization, (iii) derivation of sociological observations from the simplified and regularized text. Hays describes the RAND approach to (i) very briefly; step (ii) is discussed at greater length. The third step is examined in terms of Osgood's "evaluative assertion analysis," and some shortcomings of the method, notably its shallowness, are noted.

3. PROCESSING NATURAL LANGUAGE TEXT
D. G. Hays
P-3461, October 1960, 7 pp.
AD 640 658

Two applications of the computer to natural language processing are discussed: for publication of scientific findings and for content analysis by social and behavioral scientists. A documentation system is outlined that starts at the author's typewriter and uses the computer as an editorial aide and as final typist; to help with typing-setting; and to produce lists, bibliographies, files, etc. These tasks require little linguistics and are already feasible, although interchange of results will require adopting a standard procedure. Content analysis, on the other hand, presupposes sophisticated linguistic analysis, especially in semantics, which is still lacking. The present successful applications in documentation promise to support further work in linguistics. (Prepared for a Seminar on Computational Linguistics at M.I.T., October 1966).

9. Psycholinguistics

1. AN INFORMATION PROCESSING THEORY OF VERBAL LEARNING
E. A. Feigenbaum

A theory of some elementary forms of human symbolic learning: memorization, discrimination, association, and attention direction. The theory is concerned with mental activity at the level of the
processing of information symbols, which are the basic units manipulated. The
precise statement of the theory is given in the language of a digital computer,
specifically as a set of programs in
IPL-V called EPAM (Elementary Perceiver and Memorizer). The author deals gener-
ally with information structures and
processes for discrimination and associa-
tion learning, and specifically with be-
havior in the standard rote-learning
task. The behavior of the model is com-
pared with that of human subjects.

2. THE SIMULATION OF VERBAL LEARNING
BEHAVIOR
E. A. Feigenbaum
Proceedings of the Western Joint
pp. 120-132.
An abbreviated description of EPAM and
its behavior.

3. PERFORMANCE OF A READING TASK BY AN
ELEMENTARY PERCEIVING AND MEMORIZING
PROGRAM
E. A. Feigenbaum and H. A. Simon
Behavioral Science, vol. 8, no.
3 (January 1962), pp. 72-76.
Experiments show that the mechanisms
postulated in EPAM for rote memory tasks
are adequate for simulating, at least
macroscopically, the processes used by
human beings in acquiring the ability
to read and understand printed words.
The authors provide a summary descrip-
tion of the EPAM program, mention the
main processes it uses in rote memory
tasks, and describe how these processes
are used in learning to read.

4. A THEORY OF THE SERIAL POSITION
EFFECT
E. A. Feigenbaum and H. A. Simon
The theory presented is sufficient to
predict (qualitatively and quantitatively)
the shape of the serial error curve. In
addition, other rote-learning phenomena
are explained. The theory postulates a
serial information processing mechanism
that learns (on the average) one item
from a serial list every k seconds, and
uses an anchor-point processing strategy
for organizing its learning effort over
time. Two ways described to make predic-
tions from the postulates are by a com-
puter programmed to process information
and by a simple mathematical model.

5. A NET TO SIMULATE MORSE-CODE LEARN-
ING
I. Barr
AD 429 107
The design of a neural net which can
learn and recall six letters of Morse
Code is presented. The net, called
MCN, can learn these coded equivalents
in any order. By cascading the MCN with
a sequence-recall net, a new net is
created, called T, which is able to en-
code words after being taught the indi-
vidual letter codes. The method of
combining these two nets and a modifi-
cation of the referenced sequence-recall
net are also discussed.

6. STUDIES IN INFORMATION PROCESSING
THEORY: SIMILARITY AND FAMILIARITY
IN VERBAL LEARNING
H. A. Simon and E. A. Feigenbaum
AD 430 739
A presentation of results obtained by
simulating various verbal learning ex-
periments with the Elementary Perceiving
and Memorizing Program (EPAM), an infor-
mation processing theory of verbal learn-
ing. Predictions were generated for
experiments manipulating intra-list simi-
larity (Underwood), inter-list similar-
ity (Bruce), and familiarity and meaning-
fulness. The stimulus materials were
nonsense syllables, learned in paired-
associate fashion. The predictions made
by the model are generally in good agree-
ment with the experimental data. The
fit of the EPAM predictions to the
Chentoff data is particularly signi-
cificant since it lends support to the hy-
pothesis that the mechanism by which a
high degree of meaningfulness of items
facilitates learning is the high fa-
miliarity of these items.

7. CONSTRUCTION OF A SIMULATION PROCESS
FOR INITIAL PSYCHIATRIC INTERVIEWING
N. L. Gilbreath, R. E. Bellman, M. B.
Friend, and L. Kurland
P-2933, June 1964, 13 pp.
AD 602 976
A description of the construction of a
simu-
tion of an initial psychiatric in-
terview, regarded as an example of an
adaptive, multi-stage decision process.

8. ON THE CONSTRUCTION OF A SIMULATION
OF THE INITIAL PSYCHIATRIC INTERVIEW
R. E. Bellman, M. B. Friend, and
L. Kurland
(National Institute of Health)
AD 602 649
A description of the construction of a
simulation of an initial psychiatric
interview. It can be regarded as an
example of an adaptive, multi-stage de-
cision process.

9. COMPUTER SIMULATION OF HUMAN BEHAVIOR
E. A. Feigenbaum
P-2905, May 1964, 14 pp.
AD 603 075
This paper discusses computer simulation
of human behavior with particular empha-
sis on human cognitive behavior (learn-
ing and problem-solving processes).
Included are remarks about the simula-
tion of this behavior, the digital comput-
er.
10. GENERALIZATION OF AN ELEMENTARY PERCEIVING AND MEMORIZING MACHINE
E. A. Feigenbaum and H. A. Simon
P·25, March 1965, 16 pp.
Proceedings of IFIP Congress 62,
North Holland, 1963.

An exploration of the Elementary Perceiv-
er and Memorizer (EPAM) III model, which
is concerned with simulating the behav-
or observed in psychological experiments
on meaningfulness in verbal learning.

In this model, EPAM information processes
and structures are generalized to deal
with stimulus objects of arbitrary com-
plexity. Discrimination processes dis-
criminate objects on the basis of pro-
peries of the objects themselves or on
the basis of properties of constituent
subobjects. The learning processes of
the EPAM III model provide an associa-
tive mechanism by means of which earlier
learning is brought to bear in a useful
way on later learning.

11. LINGUISTIC RELATIVITY AND THE
LANGUAGE LEARNING PROCESS
R. M. Schwarzc
AD 644 929
A five-stage analysis of the language-
learning process, and an investigation of
whether this analysis supports or contra-
dicts the Whorfian hypothesis of linguis-
tic relativity. The syntactic construc-
tions of a language influence the types
of conceptual relationship that the child
perceives. Words and phrases grouped to-
gether by the conventions of the language
have a determining effect on conceptual
representations associated with them.

This effect is somewhat mitigated by the
transformational phase of language learn-
ing, which opens up the full range of
stylistic devices available in a language,
thus extending the number of concepts and
concepts that may gain expression in that
language. However, earlier perceptual
and conceptual habits will probably per-
sist throughout a person's adult life un-
less he is forced to change by some new
experience, such as the mastery of a dif-
ferent language.

12. STEPS TOWARD A MODEL OF LINGUISTIC
PERFORMANCE: A PRELIMINARY SKETCH
R. M. Schwarzc
AD 646 598
The model of linguistic performance out-
lined here is based on the author's
theoretical approach to linguistics de-
scribed in 9-11. The task of formulating
a model of linguistic performance is dis-
cussed and an approach to embodying the
model as a computer program is proposed.

The methodology of current linguistic
theory is criticized for several of its
features that render it inapplicable to
a realistic model of performance, and
remedies for these deficiencies are pro-
posed. The syntactic and conceptual data
structures, inference rules, generation
and understanding mechanisms, and learn-
ing mechanisms proposed for the model are
all described. The learning process is
formulated as a series of five stages,
and the roles of nonlinguistic feedback
and inductive generalization relative to
these stages are described. Finally, the
implications of a successful performance
model for linguistic theory, linguistic
applications of computers, and psychological
theory are discussed.

10. Character Readers

1. A DIGITAL SIMULATION OF AN AIDED
ADAPTIVE CHARACTER READING MACHINE
P. Baran and G. Estrin
The pattern-recognition system here
simulated on a digital computer uses an
initial man-machine learning phase.
Transformations on a deformed set of 48
samples of each of ten numerals are used
to form separation filters, while a
second set of 480 similarly varied num-
erals serve as the "unknown" characters
that are examined. Measured probability
density distributions of the inked areas
of all characters are established, and
a weighted stencil or filter is created
to distinguish each character relative
to the possible set of characters. This
experiment demonstrates the extent to
which the actual value of the best
"score at match" between the unknown
and each character in the set provides con-
fidence in recognition. Whenever the
best score is too low, it is possible to
call for more complex processes to aid
recognition, permitting the construction
of recognition systems of greater
accuracy than the basic reading mechan-
ism.

2. AN AIDED ADAPTIVE CHARACTER READER
FOR MACHINE TRANSLATION OF LANGUAGES
P. Baran and G. Estrin
Here the authors show that the eleme-
tary model of P-1989 may aid in con-
structing a fast-input device for a lan-
guage translation machine if it was
able to make use of frequency distribu-
tion characteristics of the dictionary.
A possible implementation with a raw
character reading rate up to 500 charac-
ters a second appears feasible.

3. THE RAND TABLET: A MAN-MACHINE
GRAPHICAL COMMUNICATION DEVICE
M. R. Davis and T. O. Ellis
AD 444 103
(Advanced Research Project Agency)
Proceedings of the Fall Joint Computer
Conference, Spartan Books, Baltimore,
1964, pp. 325-331.

This Memorandum describes a low-cost,
two-dimensional graphic input tablet
and stylus developed at the RAND Corpor-
ation for conducting research on man-
machine graphic communications. The
tablet is a printed-circuit screen com-
plete with printed-circuit capacitive-
coupled encoders with only 40 external
connections. The writing surface is a 10" x 10" area with a resolution of 100 lines per inch in both x and y. Thus, it is capable of digitizing 100 discrete locations with excellent linearity, allowing the user to "write" in a natural manner. The system does not require a computer-controlled scanning system to locate and track the stylus. Several institutions have recently installed copies of the tablet in research environments. It has been in use at RAND since September 1963.

4. ABSTRACTION AND PATTERN CLASSIFICATION
   R. E. Bellman, R. E. Kalaba, and L. A. Zadeh
   AD 607 500
   A preliminary discussion of a general framework for the treatment of pattern-recognition problems. The authors define "fuzzy" set, then show how this may be used in a sequential experimental procedure to ascertain whether a symbol is a member of a particular set or not.

5. COMPUTER RECOGNITION OF ON-LINE, HANDWRITTEN CHARACTERS
   M. I. Bernstein
   AD 451 231
   (Advanced Research Project Agency)
   Discussion of a method for recognizing single, hand-written characters using an on-line graphical input device such as a digitizing pantograph, light pen, or RAND Graphic Input Tablet, as the primary information source. Basically, the method consists of filtering and smoothing the input stream to eliminate as much redundancy as possible. Direction of the stylus movement is quantized into one of eight directions, allowing each stroke of a character to be described as a series of connected straight-line segments. By eliminating various measures on the stroke, the description is size-, position-, and rotation-independent. To restore some rotational orientation and to discriminate between open, closed, and multi-stroke characters, end-point comparisons are added to the description.

6. ON THE DEVELOPMENT OF EQUITABLE GRAPHIC I/O
   T. O. Ellis and W. L. Sibley
   AD 637 781
   The desire for direct interaction between man and machine has led to the study of computer interpretation of free-hand motions of a stylus and the real-time responses to these motions. The operating environment of CR-AIL--Graphic Under Language Under development at RAND is discussed. The system utilizes elements of pictorial and verbal languages.

11. TRANSLATIONS

1. EXACT METHODS IN LINGUISTIC RESEARCH
   O. S. Akhmanova, I. A. Mel'chuk, R. N.
   Frumkina, E. V. Paducheva
   (Translated by D. G. Hays and D. V.
   Mohr)
   R-397, September 1963
   University of California Press,
   Berkeley, 1963
   AD 420 656
   An attempt to illuminate several scientific results in the application of exact methods to linguistic research. Often termed "mathematical linguistics," the essence of this new approach and its content consists not of creating some special kind of "linguistics," but rather of perfecting, of making accurate, reliable, and modern the methods of conventional linguistic research. This Report deals with: (a) questions of general linguistics which relate the following discussions of concrete methods of exact study and description of linguistic phenomena to earlier linguistics; (b) the place and role of machine translation in contemporary linguistics in both a theoretical and a practical sense; (c) possible applications of statistical methods to linguistic research, together with a discussion of basic principles and concepts of statistical analysis; (d) possible applications of information theory to language study. The Report does not treat linguistic applications of "nonquantitative" mathematics--in particular, mathematical logic. This work was translated in order to show the direction and scope of the best Russian research in modern linguistics.

2. TWO OPERATORS FOR DETERMINING AGREEMENT FOR AUTOMATIC SYNTACTIC ANALYSIS
   I. A. Mel'chuk
   (Translated by D. V. Mohr)
   RM-3190, June 1962
   Two interrelated programs are presented for the automatic analysis of Russian sentences in any language characterized by grammatical agreement among certain word classes that must be tested to derive the proper syntactic structure for each sentence being processed.

3. PROBLEMS OF ALGORITHMIC COMPOSITION OF SUBJECT INDEXES
   (Brief Survey of Foreign Literature)
   B. V. Takushin
   Scientific-Technical Information
   (Nauchno-rekhlicheskaia Informatsiya)
   (Translated by J. B. Gazley and
   W. B. Holland
   LT-65-102, March 15, 1960, pp. 22-25
   Major trends in developing methods of machine assignment of subject indexes are examined. Two approaches are discernible: a) the use of subject descriptors based on key words extracted from the document (usually from the title); b) the use of predetermined subject categories under which documents are indexed. It is noted that the systems that rely on statistical characteristics of the appearance of key words in the documents are the more promising.
DEPOSIT LIBRARIES IN THE UNITED STATES

Alabama
Auburn: Auburn University*

Arizona
Tempe: Arizona State University*
Tucson: University of Arizona*

California
Berkeley: University of California
Goleta: General Motors Corporation*
Irvine: University of California*
Los Angeles: University of California University of Southern California*
Pasadena: California Institute of Technology
Riverside: University of California*
Stanford: Stanford University

Colorado
Denver: Denver Public Library

Connecticut
New Haven: Yale University

Washington, D.C.
The Library of Congress

Florida
Gainesville: University of Florida

Georgia
Atlanta: Georgia Institute of Technology

Hawaii
Honolulu: University of Hawaii (partial deposit, beginning with R-287, RM-1606, and P-787)

Illinois
Chicago: University of Chicago
Des Plaines: Borg-Warner Corporation*
East St. Louis: Parks College of Aeronautical Technology*
Urbana: University of Illinois

Indiana
Lafayette: Purdue University

Iowa
Ames: Iowa State University of Science and Technology
Iowa City: State University of Iowa

*Subscription library; access to collection may be restricted.
Kansas
  Manhattan: Kansas State University*

Kentucky
  Lexington: University of Kentucky

Louisiana
  New Orleans: Tulane University*

Maryland
  Baltimore: The Johns Hopkins University
  College Park: University of Maryland*

Massachusetts
  Amherst: University of Massachusetts*
  Cambridge: Harvard University
           Massachusetts Institute of Technology
  Chestnut Hill: Boston College*

Michigan
  Ann Arbor: University of Michigan
  Detroit: Detroit Public Library*
  East Lansing: Michigan State University*
  Kalamazoo: Western Michigan University*

Minnesota
  Minneapolis: University of Minnesota

Missouri
  Columbia: University of Missouri*
  Kansas City: Linda Hall Library
  St. Louis: Washington University

Nebraska
  Lincoln: The University of Nebraska

Nevada
  Las Vegas: Nevada Southern University*

New Jersey
  Princeton: Princeton University

New York
  Buffalo: State University of New York*
  Ithaca: Cornell University
  New York: Columbia University
           The New York Public Library
  Stony Brook, L. I.: State University of New York*
  Syracuse: Syracuse University*

North Carolina
  Durham: Duke University

* Subscription library; access to collection may be restricted.
<table>
<thead>
<tr>
<th>State</th>
<th>Location</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ohio</td>
<td>Cleveland</td>
<td>Case Institute of Technology*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cleveland Public Library</td>
</tr>
<tr>
<td></td>
<td>Columbus</td>
<td>Ohio State University</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>Norman</td>
<td>The University of Oklahoma</td>
</tr>
<tr>
<td>Oregon</td>
<td>Corvallis</td>
<td>Oregon State University</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>Pittsburgh</td>
<td>Carnegie Library of Pittsburgh*</td>
</tr>
<tr>
<td></td>
<td>University Park</td>
<td>Pennsylvania State University*</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>Providence</td>
<td>Brown University</td>
</tr>
<tr>
<td>Tennessee</td>
<td>Memphis</td>
<td>Memphis State University*</td>
</tr>
<tr>
<td></td>
<td>Nashville</td>
<td>Joint University Libraries</td>
</tr>
<tr>
<td>Texas</td>
<td>Austin</td>
<td>The University of Texas</td>
</tr>
<tr>
<td></td>
<td>College Station</td>
<td>Texas A &amp; M University*</td>
</tr>
<tr>
<td></td>
<td>Houston</td>
<td>Rice University</td>
</tr>
<tr>
<td>Utah</td>
<td>Provo</td>
<td>Brigham Young University*</td>
</tr>
<tr>
<td></td>
<td>Salt Lake City</td>
<td>University of Utah</td>
</tr>
<tr>
<td>Virginia</td>
<td>Blacksburg</td>
<td>Virginia Polytechnic Institute*</td>
</tr>
<tr>
<td></td>
<td>Charlottesville</td>
<td>University of Virginia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington</td>
<td>Renton</td>
<td>The Boeing Company*</td>
</tr>
<tr>
<td></td>
<td>Seattle</td>
<td>University of Washington</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>Madison</td>
<td>University of Wisconsin</td>
</tr>
<tr>
<td>Wyoming</td>
<td>Cheyenne</td>
<td>Wyoming State Library*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Subscription library; access to collection may be restricted.
DEPOSIT LIBRARIES ABROAD

Australia
Sydney: The University of Sydney (partial deposit beginning with R-287, RM-1606, and P-787)

Canada
Ottawa: Defence Research Board*
National Research Council of Canada

England
Boston Spa, Yorkshire: National Lending Library for Science and Technology

France
Paris: S.E.M.A. Documentation*
Société Francaise de Recherche Opérationnelle

Germany
Bochum-Querenburg: Universitätsbibliothek der Ruhruniversität*
Ebenhausen/Isartal: Stiftung Wissenschaft und Politik*
Frankfurt a.M.: Stadt- und Universitäts-Bibliothek
Munich: Industrieanlagen-Betriebsgesellschaft Einsteinstrasse*

India
Kanpur: Indian Institute of Technology*
New Delhi: Defence Science Laboratory*
Indian Council of World Affairs*
Indian Institute of Public Administration*

Israel
Haifa: Technion - Israel Institute of Technology*
Tel-Aviv: Mamram Computer Center*

Japan
Tokyo: National Diet Library (complete deposit except for R-245)

Puerto Rico
Rio Piedras: Universidad de Puerto Rico*

Sweden
Nykoping: Akiebolaget Atomenergi

*Subscription library; access to collection may be restricted.
Author Index

BARAN, P., AND G. ESTRIN
A digital simulation of an aided adaptive character reading machine, 10-1.

An aided adaptive character reader for machine translation of languages, 10-2.

BARR, L.
A net to simulate Morse-code learning, 9-5.

BELLMAN, R. E., M. B. FRIEND, AND L. KURLAND
On the construction of a simulation of the initial psychiatric interview, 9-8.

BELLMAN, R. E., R. E. KALABA, AND L. A. ZEIDAH
Abstraction and pattern classification, 10-4.

See also N. L. Gilbreath.

BERNSTEIN, M. I.
Computer recognition of on-line, hand written characters, 10-5.

BIRNBAUM, H.
Predicative case, short form adjectives, and predicatives, 5.2-6.

On the predicative use of the Russian infinitive, 5.5-9.

BRYAN, G. E.

DAVIS, M. R., AND T. O. ELLIS
The RAND Tablet: A man-machine graphical communication device, 10-3.

DREYFUS, H. L.
Alchemy and artificial intelligence, 1-21.

EDMUNDS, H. P.
Linguistic analysis in machine translation research, 1-8.

EDMUNSON, H. P., AND D. G. HAYS
Research methodology, 3-3.

EDMUNSON, H. P., K. E. HARPER, AND D. G. HAYS
Survey and critique, 1-5.

EDMUNSON, H. P., K. E. HARPER, D. G. HAYS, AND B. J. SCOTT
Manual for postediting Russian text, 3-5.

EDMUNSON, H. P., D. G. HAYS, E. K. RENNEN, AND R. I. SUTTON
Manual for keypunching Russian scientific text, 4.1-1.

EDMUNSON, H. P., D. G. HAYS, AND R. I. SUTTON
Resume of machine codes and card formats, 4.1-2.

EDMUNSON, H. P., D. G. HAYS, E. K. RENNEN, AND D. V. NOHR
Manual for pre-editing Russian scientific text, 4.1-3.

ELBOURN, R. D., AND W. H. WARE
The evolution of concepts and languages of computing, 4.3-1.

ELLIS, T. O., AND W. L. SIBLEY
On the development of equitable graphic I/O, 10-6.

See also M. R. Davis.

ESTRIN, G.
See P. Baran.

FEIGENBAUM, E. A.
Soviet cybernetics and computer sciences, 1960, 1-3.

An information processing theory of verbal learning, 9-1.

The simulation of verbal learning behavior, 9-2.


FEIGENBAUM, E. A., AND H. A. SIMON
Performance of a reading task by an elementary perceiving and memorizing program, 9-3.

A theory of the serial position effect, 9-4.

Generalization of an elementary perceiving and memorizing machine, 9-10.

See also H. A. Simon.

FRIEND, M. B.
See N. L. Gilbreath; R. E. Bellman.

GAILMAN, H.
Dependency systems and phrase structure systems, 2.1-3.

GAZLEY, J. B., AND W. B. HOLLAND
Problems of algorithmic composition of subject indexes, 11-3.

GILBREATH, N. L., R. E. BELLMAN, M. B. FRIEND, AND L. KURLAND
Construction of a simulation process for initial psychiatric interviewing, 9-7.

GOLD, E. M.
Language identification in the limit, 3-8.

GOTTSHALL, D. B.
See Dr. G. Hays.

GRAVES, P. A., D. G. HAYS, M. KAY, AND T. W. ZIEHE
Computer routines to read natural text with complex formats, 4.2-15.
GRUENBERGER, F. J.
Benchmarks in artificial intelligence, 1-4.

HALLIDAY, N. A. K.
Some aspects of the thematic organization of the English clause, 6-9.

HARPER, K. E.
Soviet research in machine translation, 1-12.
Machine translation, 1-17.
Procedures for the determination of distributional classes, 5.3-2.
Machine translation of Russian propositions, 5.3-4.
The position of prepositional phrases in Russian, 5.3-5.
Prepositional phrases and automatic parsing, 5.3-6.
Measurement of similarity between nouns, 5.3-7.
Studies in inter-sentence connection, 5.3-8.
Some combinatorial properties of Russian nouns, 5.3-10.

HARPER, K. E., AND D. G. HAYS
The use of machines in the construction of a grammar and computer program for structural analysis, 1-6.

HARPER, K. E., D. G. HAYS, A. S. KOZAK, AND B. J. SCOTT
Bibliography of Russian scientific articles, 5.1-1.

HARPER, K. E., D. G. HAYS, AND D. V. MOHR

HARPER, K. E., D. G. HAYS, AND A. KOUTSOUDAS
A glossary of Russian physics on punched cards, 5.1-2.

HARPER, K. E., D. G. HAYS, D. S. WORTH, AND T. W. ZIEHE
Six tasks in computational linguistics, 1-10.
See also H. P. Edmundson.

HAYS, D. G.
Automatic language-data processing, 1-4.
Linguistic research at The RAND Corporation, 1-7.
Automatic computers in machine-translation research, 1-9.

Report of a summer seminar on computational linguistics, 1-10.
Readings in automatic language processing, 1-24.
Introduction to computational linguistics, 1-25.
Grouping and dependency theories, 2.1-1.
Basic principles and technical variations in sentence-structure determination, 2.1-2.
Dependency theory: A formalism and some observations, 2.1-4.
An annotated bibliography of publications on dependency theory, 2.1-5.
Research procedures in machine translation, 3-1.
An introduction to computational procedures in linguistic research, 3-2.
On the value of a dependency connection, 3-4.
Connectability calculations, syntactic functions, and Russian syntax, 4.2-5.
Order of subject and object in scientific Russian when other differentiations are lacking, 5.2-1.
Pairs of Russian words with high correlation, 5.3-1.
Automatic language-data processing in sociology, 8-1.
Automatic content analysis: Some entries for a transformation catalog, 8-2.
Acquisition, archiving, and interchange, 4.1-6.
Computational linguistics: Research in progress at The RAND Corporation, 1-22.
Processing natural language text, 8-3.

HAYS, D. G., AND R. MA

HAYS, D. G., AND T. W. ZIEHE
Russian sentence-structure determination, 4.2-4.

HAYS, D. G., AND B. J. SCOTT
A Russian structure for comparison, 5.2-2.
HAYS, D. G., AND D. V. MOHR
Exact methods in linguistic research, 11-1.

KELLY, H. S., AND T. W. ZIEHE
Glossary lookup made easy, 4.2-3.

KAYS, D. G., B. HENISZ-DOSTERT, AND
M. L. RAPP

KOCHEM, M., D. M. MAC KAY, M. E. MARON,
H. SCRIVEN, AND L. UHR
Computers and comprehension, 2.2-4.

KOUTSOUDA, A.
See K. E. Harper.

KOZAK, A. S.
High frequency words and occurrence forms in Russian physics, 5.1-4.
Complementation in Russian: Theory and application, 5.2-7.

KOZAK, A. S., C. H. SMITH, AND MEMBERS OF ALDP GROUP
A glossary of Russian physics, 5.1-3.

KRIEGER, F. J.
The Soviet classification scheme for literature, 7-7.

KUINS, J. L.
See N. E. Maron.

KURLAND, L.
See N. L. Gilbreath; R. E. Bellman.

LAND, D. N.
A simple proof of a theorem on
self-synchronizing automata, 2.2-6.

LEVIN, R. E.
Relational data file II: Im-
plementation, 7-14.

LEVI, R. E. AND N. E. MARON
Cybernetics and its development in
the Soviet Union, 1-15.
Relational data file: A tool for
mechanized inference execution and
data retrieval, 7-11.
A computer system for inference
execution and data retrieval, 7-15.

LIEBEN, S. L.
The queens grammar, 2.2-9.

LINDSAY, R. K., T. W. PRATT, AND
K. M. SHAIVOR
An experimental syntax-directed data
structure language, 4.3-2.

MAK, R.
See D. G. Hays.

MAC KAY, D. M.
Linguistic and non-linguistic "under-
standing" of linguistic tokens, 2.2-5.
See also M. Kochen.

MARKS, S. L.
See J. J. Robinson; T. W. Ziehe.
MARON, M. E.
On cybernetics, information processing, and thinking, 1-13.

Automatic indexing: An experimental inquiry, 7-2.

A logician's view of language data processing, 7-3.

Information retrieval: A look at the logical framework and some new concepts, 7-4.

Probability and the library problem, 7-5.

Mechanized documentation: The logic behind a probabilistic interpretation, 7-9.

The logic of interrogating a digital computer, 7-10.

Relational data file I: Design philosophy, 7-13.

MARON, M. E., AND J. L. KUHNS
On relevance, probabilistic indexing and information retrieval, 7-1.

See also R. E. Levien; M. Kochen.

MATTISON, R.
A formal system for the logical analysis of temporal relationships between intervals of time, 7-10.

See also H. P. Edmundson; K. E. Harper; D. G. Hays.

MOHR, D. V.
Two operators for determining agreement for automatic syntactic analysis, 11-2.

See also H. P. Edmundson; K. E. Harper; D. G. Hays.

NEEDHAM, R. M.
Automatic classification in linguistics, 3-10.

The termination of certain iterative processes, 3-9.

NEWELL, A.
New areas of application of computers, 1-1.

PATRICK, R. L.
A proposal for the indirect retrieval of unpublished technical material, 7-6.

PETERSON, V.
Technical writers: Educated or trained?, 1-18.

PRATT, T. W.
See R. K. Lindsay.

PUSTULA, J. H.
Governors of the conjunction chto, 5-2-3.

RAPPE, M. L.
See D. G. Hays.

RENNER, E. K.
See H. P. Edmundson.

ROBINSON, J. J.
Endocentric constructions and the Cocker parsing logic, 4.2-10.

Preliminary codes and rules for the automatic parsing of English, 6-1.

Automatic parsing and fact retrieval: a comment on grammar, paraphrase, and meaning, 7-8.

The transformation of sentences for information retrieval, 7-12.

ROBINSON, J. J., AND S. L. MARKS
PARSE: A system for automatic syntactic analysis of English text - Parts I & II, 6-2.

ROSIN, R. F.
Translation of artificial languages by compiler programs, 4.2-6.

SCOTT, B. J.
See K. E. Harper; H. P. Edmundson; D. G. Hays.

SCHUTZENBERGER, M. P.
On products of finite dimensional stochastic matrices, 2.2-8.

Some remarks on acceptable sets of numbers, 2.2-7.

SCHWARTZ, R. M.
Steps toward a model of linguistic performance: A preliminary sketch, 9-12.

Linguistic relativity and the language learning process, 9-11.

SCRIVEN, M.
See M. Kochen.

SHAVOR, K. M.
See R. K. Lindsay.

SIBLEY, W. L.
See T. O. Ellis.

SIMON, H. A.
Experiments with a heuristic compiler, 4.2-7.

SIMON, H. A., AND T. A. VAN WORMER
Some Monte Carlo estimates of the Yule distribution, 3-6.

SIMON, H. A., AND E. A. FEIGENBAUM
Similarity and familiarity in verbal learning, 9-6.

SMITH, C. H.
See A. S. Kozak.
STEWART, W. A.
Some linguistic problems of Russian
graphic abbreviations, 5.2-4.

SUTTON, R. I.
See H. P. Edmundson.

TAFT, T. D.
See M. Kay.

TONGE, F. M.
A simple scheme for formalizing data
retrieval requests, 7-17.

UHR, L.
See M. Kochen.

VALADEZ, F.
See M. Kay.

VAN NORMER, T. A.
See H. A. Simon.

WARE, W. H.
See R. B. Elbourn.

WORTH, D. S.
Suprasyntactics, 2.2-3.
The suffix -aga, 5.2-5.
Transformational criteria for the
classification of predicate genitive
constructions in Russian, 5.3-1.
A deep index of derivational morphology,
5.2-8.
Vowel-zero alternations in derivation,
5.2-10.
See also K. E. Harper; D. G. Hays.

ZADEH, L. A.
See R. E. Bellman.

ZIEHE, T. W.
The Catalog: A flexible structure for
data storage, 4.2-16.
See also P. A. Graves; K. E. Harper; D. G.
Hays; M. Kay; H. S. Kelly.

ZIEHE, T. W., AND S. L. MARKS
The nature of data in language analy-
sis, 4.2-1.
## PUBLICATION NUMBER INDEX

### RAND MEMORANDA

<table>
<thead>
<tr>
<th>Publication Number</th>
<th>Volume</th>
<th>Page</th>
<th>Number</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>2060</td>
<td>3-3</td>
<td>3625</td>
<td>5.3-5</td>
<td>4540</td>
</tr>
<tr>
<td>2061</td>
<td>4.1-1</td>
<td>3753</td>
<td>10-5</td>
<td>4582</td>
</tr>
<tr>
<td>2063</td>
<td>1-5</td>
<td>3774</td>
<td>5.2-6</td>
<td>4645</td>
</tr>
<tr>
<td>2064</td>
<td>4.1-2</td>
<td>3850</td>
<td>9-5</td>
<td>4654</td>
</tr>
<tr>
<td>2065-1</td>
<td>4.1-3</td>
<td>3889</td>
<td>1-16</td>
<td>4793</td>
</tr>
<tr>
<td>2066-1</td>
<td>4.1-4</td>
<td>3892</td>
<td>2.2-5</td>
<td>4828</td>
</tr>
<tr>
<td>2068</td>
<td>3-5</td>
<td>3979</td>
<td>9-6</td>
<td>4920</td>
</tr>
<tr>
<td>2538</td>
<td>4.2-4</td>
<td>4005</td>
<td>7-8</td>
<td>4933</td>
</tr>
<tr>
<td>2601</td>
<td>7-1</td>
<td>4044</td>
<td>9-8</td>
<td>4986</td>
</tr>
<tr>
<td>2646</td>
<td>2.1-1</td>
<td>4065</td>
<td>2.2-4</td>
<td>5077</td>
</tr>
<tr>
<td>2655</td>
<td>5.1-3</td>
<td>4087</td>
<td>2.1-4</td>
<td>5085</td>
</tr>
<tr>
<td>2712</td>
<td>3-4</td>
<td>4122</td>
<td>10-3</td>
<td>5143</td>
</tr>
<tr>
<td>2713</td>
<td>5.3-2</td>
<td>4136</td>
<td>3-8</td>
<td>5150</td>
</tr>
<tr>
<td>2714</td>
<td>5.3-3</td>
<td>4156</td>
<td>1-15</td>
<td>5168</td>
</tr>
<tr>
<td>2799</td>
<td>1-3</td>
<td>4224</td>
<td>3-7</td>
<td>5188</td>
</tr>
<tr>
<td>2803</td>
<td>1-10</td>
<td>4258</td>
<td>4.2-9</td>
<td>5194</td>
</tr>
<tr>
<td>2916</td>
<td>3-1</td>
<td>4283</td>
<td>4.2-8</td>
<td>5209</td>
</tr>
<tr>
<td>3161</td>
<td>2.2-3</td>
<td>4307</td>
<td>10-4</td>
<td>5210</td>
</tr>
<tr>
<td>3190</td>
<td>11-2</td>
<td>4383</td>
<td>5.3-6</td>
<td>5214</td>
</tr>
<tr>
<td>3235</td>
<td>5.2-5</td>
<td>4390</td>
<td>4.1-5</td>
<td>5223</td>
</tr>
<tr>
<td>3325</td>
<td>7-7</td>
<td>4477</td>
<td>5.3-9</td>
<td>5224</td>
</tr>
<tr>
<td>3339</td>
<td>6-1</td>
<td>4479</td>
<td>2.1-5</td>
<td>5243</td>
</tr>
<tr>
<td>3383</td>
<td>5.1-4</td>
<td>4523</td>
<td>1-19</td>
<td>5252</td>
</tr>
<tr>
<td>3610-1</td>
<td>5.1-1</td>
<td>4532</td>
<td>5.3-7</td>
<td>5279</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>RAND MEMORANDA (CONT.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5345</td>
<td>1-23</td>
<td>1989</td>
<td>10-1</td>
<td>3006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1990</td>
<td>10-2</td>
<td>3069</td>
</tr>
<tr>
<td>PAPERS</td>
<td>2142</td>
<td>1-1</td>
<td>3101</td>
<td>4.2-10</td>
</tr>
<tr>
<td>187</td>
<td>2.2-1</td>
<td>2197</td>
<td>4.2-1</td>
<td>3112</td>
</tr>
<tr>
<td>1218</td>
<td>5.3-1</td>
<td>2206</td>
<td>5.2-4</td>
<td>3136</td>
</tr>
<tr>
<td>1241</td>
<td>5.1-2</td>
<td>2235</td>
<td>9-2</td>
<td>3242-1</td>
</tr>
<tr>
<td>1251</td>
<td>3-3</td>
<td>2279</td>
<td>7-2</td>
<td>3243</td>
</tr>
<tr>
<td>1321</td>
<td>1-9</td>
<td>2315</td>
<td>2.1-3</td>
<td>3244</td>
</tr>
<tr>
<td>1328</td>
<td>1-8</td>
<td>2327</td>
<td>1-11</td>
<td>3408</td>
</tr>
<tr>
<td>1588</td>
<td>1-6</td>
<td>2339</td>
<td>2.2-2</td>
<td>3411</td>
</tr>
<tr>
<td>1624</td>
<td>3-5</td>
<td>2349</td>
<td>4.2-7</td>
<td>3413</td>
</tr>
<tr>
<td>1632</td>
<td>5.2-1</td>
<td>2358</td>
<td>9-3</td>
<td>3415</td>
</tr>
<tr>
<td>1720</td>
<td>5.2-2</td>
<td>2375</td>
<td>9-4</td>
<td>3436</td>
</tr>
<tr>
<td>1771</td>
<td>4.2-6</td>
<td>2455</td>
<td>7-3</td>
<td>3439</td>
</tr>
<tr>
<td>1817</td>
<td>9-1</td>
<td>2471</td>
<td>7-4</td>
<td>3461</td>
</tr>
<tr>
<td>1866</td>
<td>8-2</td>
<td>2526</td>
<td>4.3-1</td>
<td>3476</td>
</tr>
<tr>
<td>1896</td>
<td>1-12</td>
<td>2555</td>
<td>9-10</td>
<td>3497</td>
</tr>
<tr>
<td>1900</td>
<td>1-7</td>
<td>2586</td>
<td>1-2</td>
<td>3500</td>
</tr>
<tr>
<td>1909</td>
<td>4.2-3</td>
<td>2599</td>
<td>3-6</td>
<td></td>
</tr>
<tr>
<td>1910</td>
<td>2.1-1</td>
<td>2616</td>
<td>7-5</td>
<td>REPORT</td>
</tr>
<tr>
<td>1926</td>
<td>4.2-2</td>
<td>2879</td>
<td>1-13</td>
<td>397</td>
</tr>
<tr>
<td>1941</td>
<td>5.3-4</td>
<td>2892</td>
<td>1-14</td>
<td></td>
</tr>
<tr>
<td>1962</td>
<td>8-3</td>
<td>2898</td>
<td>7-9</td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>5.2-3</td>
<td>2905</td>
<td>9-9</td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>2.1-2</td>
<td>2933</td>
<td>9-7</td>
<td></td>
</tr>
</tbody>
</table>