DOCUMENT RESUME

ED 072 076 TM 002 315

AUTHOR	Steinke, Elisabeth
TITLE	Computer Aided Tests.
PUB DATE	Aug 70
NOTE	10p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Classification; *Computational Linguistics; *Computer Oriented Programs; German; Information Retrieval; *Item Banks; Languages; *Machine Translation; Programing Languages; Statistical Analysis; *Test Construction
IDENTIFIERS	*COMIT Language

ABSTRACT

An approach to using the computer to assemble German tests is described. The purposes of the system would be: (1) an expansion of the bilingual lexical memory bank to list and store idioms of all degrees of difficulty, with frequency data and with complete and sophisticated retrieval possibility for assembly; (2) the creation of an item-synthesizing center similar to the analyzing center which permit the computer to take apart the German input; (3) the inclusion of statistical data, such as deltas, biserials, percentages, etc.; (4) to determine actual frequency of words as they occur in input; (5) to assemble vertical tests as well as parallel-horizontal tests; and (6) to enable the professional test-production staff to concentrate their efforts on activities that are not likely to be performed by a computer, such as reviewing material, finding appropriate texts, writing questions for passages, etc. "COMIT as and IR Language" by Victor H. Yngve is discussed briefly, and a sample machine translation from "An Introduction to Machine Translation" by Emile Delavenay is reproduced. (DB)

COMPUTER AIDED TESTS

Elisabeth Steinke
August, 1970

Present Classification System For German

There is presently a classification system developed for German at ETS which could eventually be used to assemble German tests with computer assistance.

Items are classified as to:

1. Key Word(s)
2. Sub-Category
3. Case, Person, Number, Gender
4. Matrix
5. Nearer Function
6. Farther Function

When all items (tested, pre-tested or committee approved) are thus classified and encoded, the computer could presumably:

Pull items thus saving valuable human labor
Assure a spread of items in the area of syntax, idioms, phonetics, etc.
Determine difficulty levels according to recorded statistics (already on item cards)
Control overlapping of test items within one test
Avoid testing more than one difficulty per item

However, all items have to be classified and encoded by staff or committee members which is time-consuming. Experience shows also that, despite the classification and encoding instructions, one and the same item may get slightly different treatment by two different people. The time that is saved later on, by computer assisted assembly, may in actuality be equal to the time put in by staff or committee members at the initial phase of classifying.

A Different Approach To Computer Aided Tests

The starting point would be the most recent, advanced, linguistically sound system used for machine translation from a foreign language into English, for example: German to English. Victor H. Yngve describes such a system in his article entitled "COMIT as a IR Language" and calls it a "user-oriented general purpose symbol-manipulation programming language". (see page 8)

In order to machine translate German into English, the computer has already been programmed to analyze, classify, store and retrieve a German input. This input could be anywhere from a title, sentence, paragraph, to a complete work. Such a computer program possesses a complete German-English memory bank. It has devices to handle idioms which are the most difficult phenomena to deal with in machine translation. Upon being fed an input in German, the computer analyzes, classifies, etc., the different parts of speech as to their role in the German sentence.

Since we would not be interested in translation as an end product, the portion of the computer system devoted to the transfer of analyzed German (input) into equivalent English (output) could be modified for our purposes (see page 9) Figure 1, representing a hypothetical translating machine and a sample of machine translation from French into English (see page 10)

(1) One of the purposes would be:

. An expansion of the bilingual lexical memory bank to list and store idioms of all degrees of difficulty, with frequency data and with complete and sophisticated retrieval possibility for assembly.

The following example will illustrate the above function of the computer:

A highly idiomatic German expression occurs: Es ist für die Katz(e)

Literally translated (1)	:	It is for the cat
In correct English	:	It is of no avail
In idiomatic English	:	It is for the birds

The computer classifies this idiom according to noun categories:

Es ist für die Katz(e). The noun Katze has a frequency code number (2) in the computer's lexical memory bank. Katze is a high frequency word occurring in most first year German courses. The idiom is not a high frequency item but it does occur in both spoken and written German and would possibly be learned in a third or fourth year German course.

---

(1)Recently in the English translation of a German novel by a well-known author, the translator rendered the idiomatic expression: ". . . und es war für die Katz" with "and the cat got it", which amounts to total nonsense in English.

(2)Katze: Frequency, see J. Pfeffer, Basic Spoken German Word List, Prentice Hall, Inc. 1964, page 27

The idiom Es ist für die Katz(e) would be stored as follows:

Katze, idiomatic usage: Es ist für die Katz(e)

Das ist für die Katz(e)	–	alternate form of idiom
It is for the cat	–	literal translation (nonsense)
It is of no avail	–	correct English
It is for the birds	–	idiomatic English

Having English equivalencies in the lexical memory bank is valuable because certain test items may be tested with English stems or options.

Other idiomatic usages of the key word Katze would be stored under that noun category also, for example:

Die Katze im Sack kaufen – to buy a pig in a poke
Wie die Katze um den heissen Brei gehen – to beat about the bush
Die Katze aus dem Sack lassen – to let the cat out of the bag
Die Katze lässt das Mausen nicht – what is bred in the bone will out in the flesh

Our first purpose then would be to expand the bi-lingual lexical memory bank to hold frequency data and to accomodate listings of idioms in such a way that they can be pulled just like any lexical entry.

(2) Another purpose would be:

The creation of an item-synthesizing center similar to the analysizing center which permits the computer to take apart the German input.

For example: the existing system is able to analyze the following test item linguistically:

Hast du daran gedacht, . . . eine Einladung zu schicken?

(A) unserem neuen Direktor
(B) unseren neuen Direktor
(C) unseres neuen Direktors
(D) unser neuer Direktor

Order/words	Key Word	Frequency Category	Sub-Category	Type	Frame Tested
1	Hast	Verb	Auxiliary 1-4	Perf. Tense Question –	
2	Du	Pronoun	Fam. address	Pattern (1-4)	
3	daran	Da-word	da + an		
4	gedacht	Participle	denken – an	mixed	
5	unserem	Possessive	unser	dat. case; sing. masc./neuter	dat. sing. masc. 5-7
6	neuen	Adjective	neu	weak form	dat. weak
7	Direktor	Noun	irreg. Genitive	masc. sing.	dat. case sing.
8	eine	Article	indef. art.	nom./acc. sing.	
9	Einladung	Noun		fem. sing.	
10.	zu	Prep./sign	infin. sign	10-11	
11.	schicken	Verb	weak	Infinitive	

With proper instructions given, the existing system has analyzed the above input automatically and will be able—given a new set of instructions—to synthesize an unlimited number of parallel items with a given frequency code and difficulty range (*). Thus parallel item production is assured for x number of tests.

(*)Difficulty range will be treated under paragraph three.

It is possible to imagine how "frames" for all major areas of language analysis can be programmed so that the computer can fill them from its expanded lexical bank to produce unlimited amounts of items according to definite specifications.

Example:

Noun: frequency data – case – number – gender specifications – testing specification

Adjective: frequency data – case – number – gender specifications – testing specification

Verb: frequency data – Tense – number – Voice – specifications – testing specification

(3) A third purpose of the computer would be:

The inclusion of statistical data such as shown on the item card below:

Hast du daran gedacht, . . . eine Einladung zu schicken?

(A) unserem neuen Direktor
(B) unseren neuen Direktor
(C) unseres neuen Direktors
(D) unser neuer Direktor

M NO. 50 CARD NO. 1 SERIES. 385 TEST. GERMAN FORM. NCBX49 BASE N. 300 DATE TABULATED 04

EDUCATIONAL TESTING SERVICE

ITEM ANALYSIS (650 OUTPUT)

RESPONSE CODE	LOW N_1	N_2	N_3	N_4	HIGH N_5
OMIT	6	3	2		1
A	13	28	38	50	56
B	13	11	12	7	2
C	10	4	1	1	
D	11	6	3	2	1
E					
TOTAL	53	52	56	60	60

FORM	BASE N	OMIT	A	B	C	D	E	M_{TOTAL}	Δ_8 SCALE	Δ_8	CRITERION
NCBX49	300	12	185*	45	16	23		13.2	BOARD	12.2	15 60
TEST CODE	ITEM NO	M_0	M_A	M_B	M_C	M_D	M_E	P_{TOTAL}	P_+	Δ_0	r_{bis}
GERMAN	50	10.5	14.7	11.1	8.8	10.3		.94	.66	11.6	.66

Statistical data such as deltas, biserials, percentages, etc., could become functional in the production and/or assembly of items. These data which importantly complement mere frequency data from the lexical bank. would be stored and retrieved upon instruction—just how cannot be discussed here.

(4) A further use of the computer would be:

To determine actual frequency of words, as they occur input, such as graded teaching texts (high school and college levels). These data could be used as comparative frequency data to frequency codes in the lexical bank of the computer (B. Q. Morgan's German Frequency Word Book) [1]. Other frequency data could be sought by feeding the computer input from nongraded material such as magazine articles, newspaper clippings, literary works of all types, etc. Scanning of texts on the basis of word frequency data (for difficulty levels) could thus be accomplished by the computer. It would serve as a pre-screening device for materials reviewed by test production personnel.

(5) A computer equipped with word and idiom frequency data and statistical data could also assemble:

vertical tests as well as parallel-horizontal tests (see (2) pp. 3-5). Vertical tests would differ from parallel or horizontal tests in that they are not assembled to reproduce similar difficulty levels but produce tests which start with easy items and become increasingly difficult in the different sections.

Logically the computer could also produce parallel batteries of such tests.

(6) At the stage of computer development and cost right now it seems unlikely a computer equipped with a symbol-manipulation programming language could be used for so small a field as German. However, if future prospects are as rosy as Emile Delavenay states then "Translating machines will soon take their place beside gramophone and colour reproductions in the first rank of modern techniques for the spread of science and culture" [2].

---

[1] J. Alan Pfeffer: Basic (Spoken) German Word List, Prentice-Hall, Inc. 1964

[2] Emile Delavenay: An Introduction to Machine Translation, F. A. Praeger, Publishers, New York, 1960

If the cost factor is reduced and ways can be found to make such a language serve more than one foreign language in item production and test assembly (instead of serving the purpose of translation), the professional staff working on the production of tests (staff members, Committee members, item writers and others) could concentrate their efforts on the more rewarding tasks of reviewing material, finding appropriate texts, writing questions for passages and similar activities which are not likely to be performed by a computer.

(7) The German Quarterly[1] reports that

> "A research group led by the German scholar Hans Eggers of the University of Saarbrücken has succeeded for the first time in syntactically analyzing modern German on a broad basis by means of an electronic computer. The results have been published in a report entitled Electronische Syntaxanalyse (Tübingen: Max Niemeyer Verlag, 1969). Eggers and his team spent several years in developing a computer program which enables them to take any German sentence, identify it grammatically and syntactically classify it."

A study of the Saarbrücken research mentioned above might yield a basis for computer aided item generating.

---

[1] Vol. XLIII, #4, November, 1970, pp. 837-38.

Appendixed Sources

Victor H. Yngve: "COMIT as an IR Language"[1]

Many of the features that make COMIT a good all-around symbol manipulation language also render it well suited to various types of information retrieval programs. Presented here is a general discussion of this unique and different programming language and an examination of some of its applications. . . . COMIT is available for the IBM 709 or 7090; a 704 version is partially checked out and could easily be put into shape. The system consists of a two-pass compiler that translates the COMIT language into a machine-oriented notation which is then run interpretively. Compiler and interpreter total about 16,000 instructions. Although COMIT was only recently released through Share, it has been in experimental use for some time and a number of problems have been programmed and run with the system. A list of some of the problem areas in which COMIT programs have been written or are being written is as follows: mechanical translation routines, information retrieval research, vocabulary analysis, text processing-editing, random generation of sentences, automatic milling machine programming, sociological data reduction, simulation of human problem solving, simulation of games, theorem-proving and mathematical logic, logico-semantic investigations, electrical network analysis.

COMIT promises to be especially useful for information retrieval problems for two reasons. Two of the most central built-in features are a simple scheme for dictionary search and a simple scheme of search using criteria such as class inclusion and context. The dictionary search scheme offers automatic alphabetization of the dictionary entries and a high-speed binary search at run time. The other search scheme, using criteria such as class inclusion and context, is a linear scan through a defined portion of the data looking for a condition of exact match or of inclusion. This "workspace search" can be easily used for searches based on descriptors or other more complicated schemes using local or distant contest. . . .

[1] Communications of the ACM, Vol. 5, #1, January 1962, pp. 19-21.

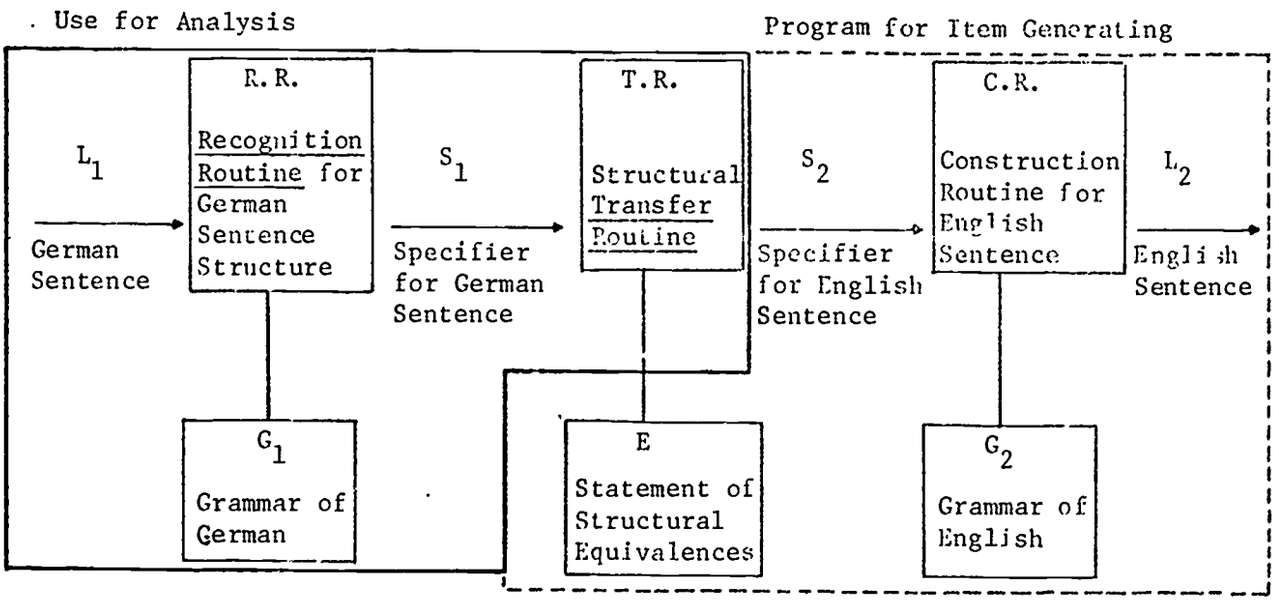
COMIT as an IR Language (Continued)

The workspace constituent therefore is the basis entity in which data is stored. A constituent may represent a letter or a word of text; it may represent an algebraic symbol; or it may represent a retrieval item consisting of a document name, a document number, and a set of descriptors. A retrieval item is typical of a COMIT constituent and can serve an an illustration of the internal structure of a constituent. A constituent consists of a symbol and optional subscripts. There may be one numerical subscript and any number of logical subscripts. . . .

COMIT is decidedly not, however, one of the programming languages that allows one to "program in English". But it is a programming language that takes advantage of intuitive feelings of naturalness that stem from the user's fluency in a natural language. . . .

A Framework For Syntactic Translation, V.H. Yngve, Mechanical Translation 1957, V.4 No.3 pp.59-65

Syntactic Translation 63



A Framework for Mechanical Translation

Figure 1

Sample Machine Translation

Émile Delavenay: An Introduction to Machine Translation[1]

LA RECHERCHE A PROGRESSEI DE FACON SPECTACULAIRE DEPUIS $FIG 1955, . . DES TRADUCTIONS UTILES SONT FAITES PAR DES MACHINES ET LEUR NOMBRE IRA EN CROISSANT, LEUR QUALITEL S AMEILIORERA CONSTAMMENT.

THE RESEARCH HAS PROGRESSED IN A SPECTACULAR MANNER SINCE 1955 /COLON/ USEFUL TRADUCTIONS ARE DONE BY MACHINES AND THEIR NUMBER WILL INCREASE CONTINUOUSLY, THEIR QUALITY WILL IMPROVE ITSELF CONSTANTLY.

Fig. 5. A Specimen of Machine Translation

(a) A Foreword to this book, as typed out in its original French in the course of its mechanical translation on I. B. M. 784 computer. This Foreword was written for the sole purpose of being so translated. See page 119 for an explanation of figures in words and other conventional symbols.

(b) Reproduction of the actual machine-translation of the same Foreword, as typed out by the I. B. M. 704 computer in Paris on 19th June 1959. The French-to-English translation programme used was conceived and designed by Mr. A. F. R. Brown of Georgetown University for the translation of texts on chemistry and nuclear energy. A fuller explanation will be found on page 119.

(1) Frederick A. Praeger, Publishers, New York