~~TOP SECRET~~

NATIONAL SECURITY AGENCY
FORT GEORGE G. MEADE, MARYLAND

CRYPTOLOG

JULY 1977

~~THIS DOCUMENT CONTAINS CODEWORD MATERIAL~~

~~TOP SECRET~~

CRYPTOLOG

VOL. IV, NO. 7   JULY 1977

Non - Responsive
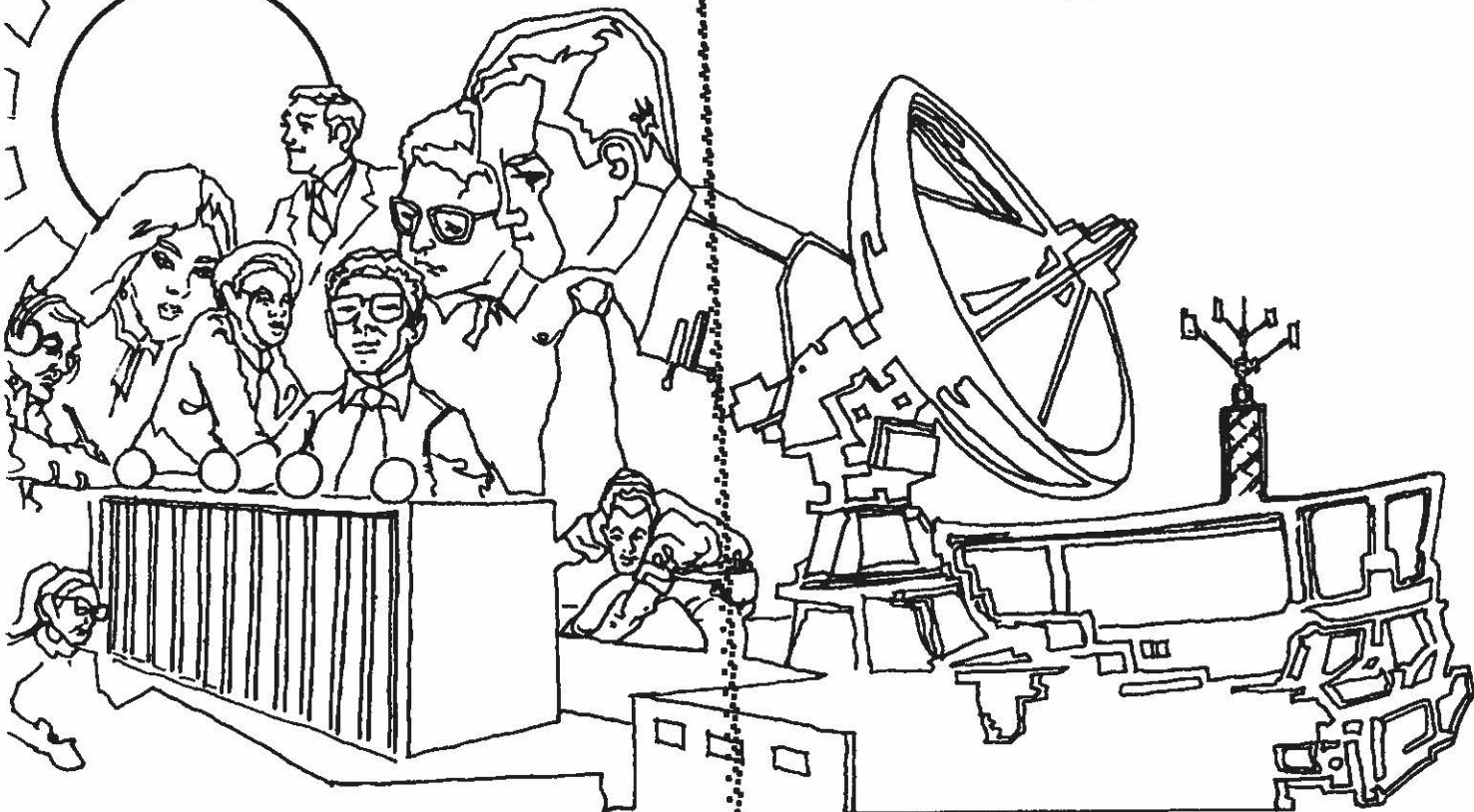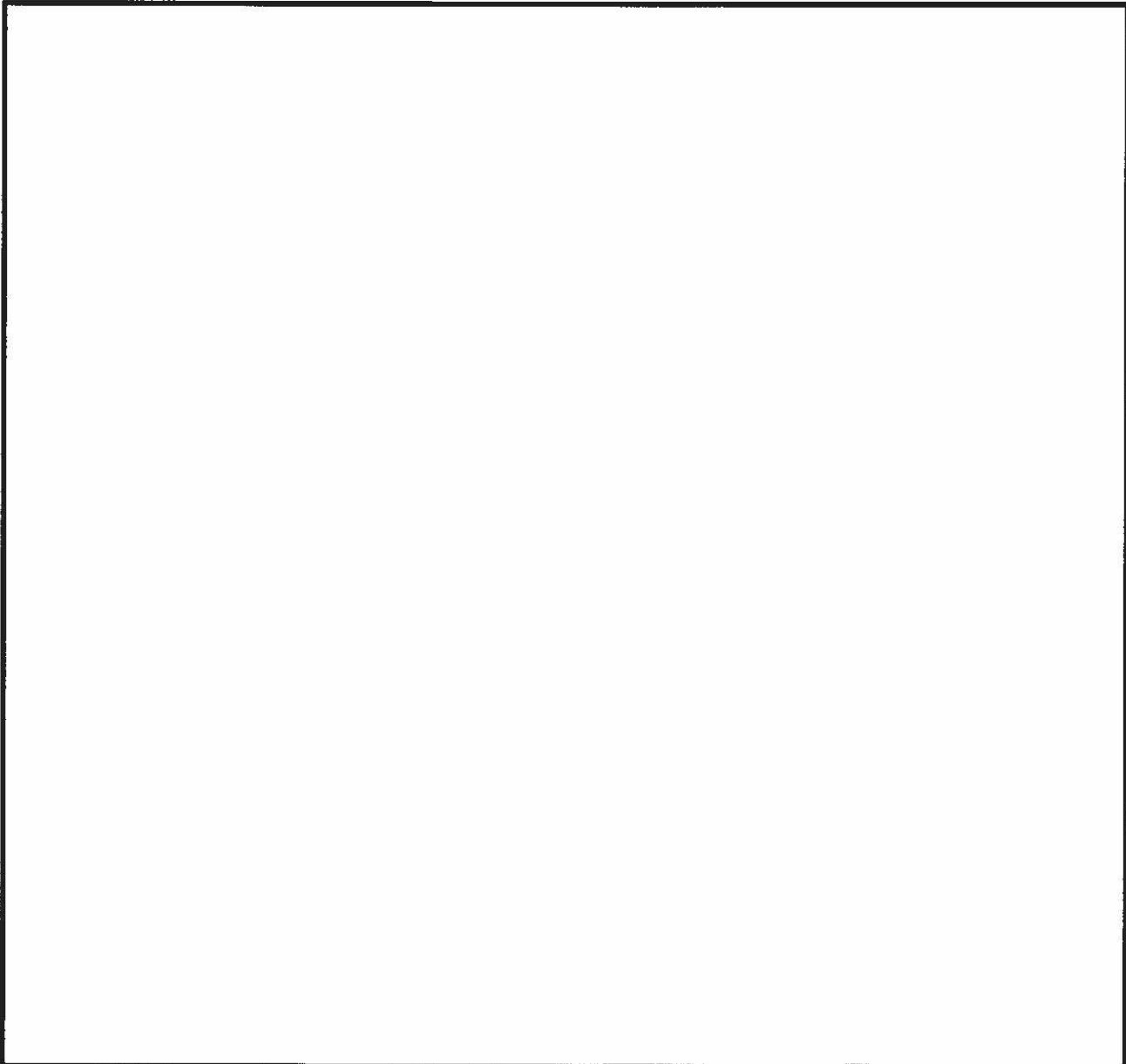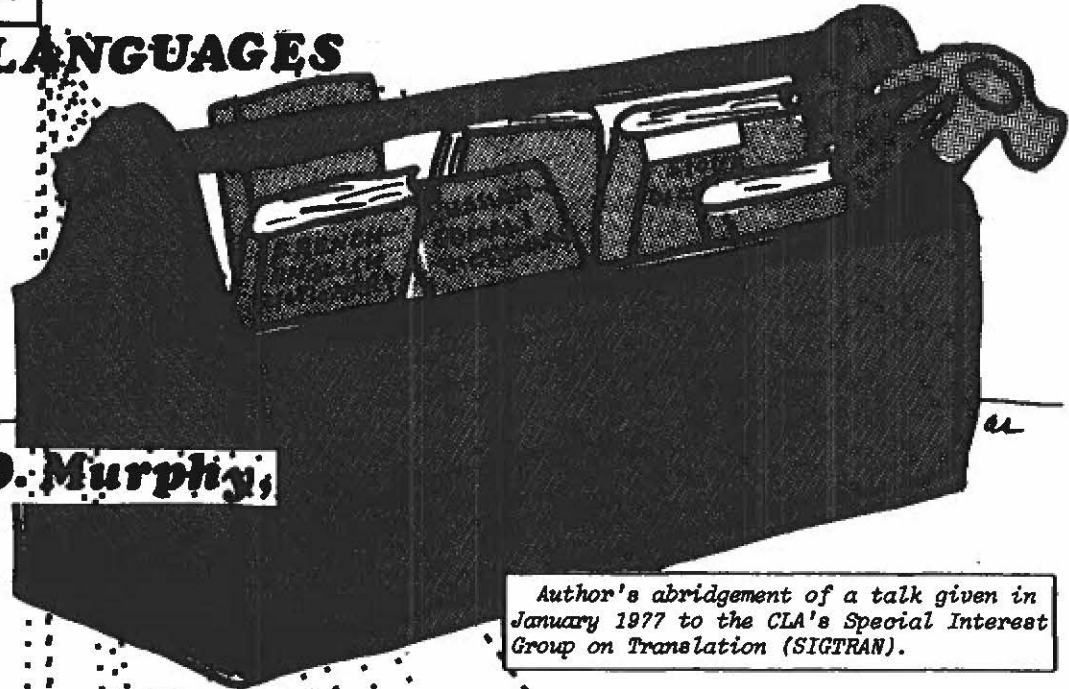
TOOL LANGUAGES

**John D. Murphy,
P16**

Most of our translation problems
involve two languages: the foreign
language and English. If we have a
piece of foreign text (for example,
_____ or Russian, or _____ or whatever),
we use the appropriate two-language dictionary
_____ English, Russian-English, etc.
There are, however, many instances in which we
cannot translate directly from the foreign
language to English, but must go through an
intermediary language. This is comparable to
the situation in which you might want to talk to
a person who speaks Polish. You do not know
Polish. But both of you can make yourselves under-
stood in passable German, so you use German as
what we shall call here a "tool language."

The use of tool languages is very common in
language research, especially when dealing
with uncommon languages in which there is
little literature in English. For example, one
of the best currently available _____ dic-
tionaries is _____-French. The most recent
bilingual _____ dictionary is _____ Russian.
A somewhat different type of problem is en-
countered when the target language is small,
obscure, and little studied, but belongs to a
family that has a least one well-known and
prestigious member.

A situation very similar to the above arises
in studying the nonstandard dialects of major
languages. From a practical, if not from a
theoretical, point of view, it would certainly
be inadvisable to study Neapolitan dialect
without some background in standard Italian,
or to study Munich dialect without knowing
standard German. At present a class is being
conducted in _____ a variant of _____
used _____ All of the students have
an excellent grounding in standard _____ At
least one course is being planned to teach an
_____ to students who have a background
in standard _____

What follows below is a discussion of tool
languages of the first type, that is, those
languages which do not have a close genetic re-
lationship to the language of interest. The
three most important of these are French,
German, and Russian.

French is an essential tool in working in
many languages _____
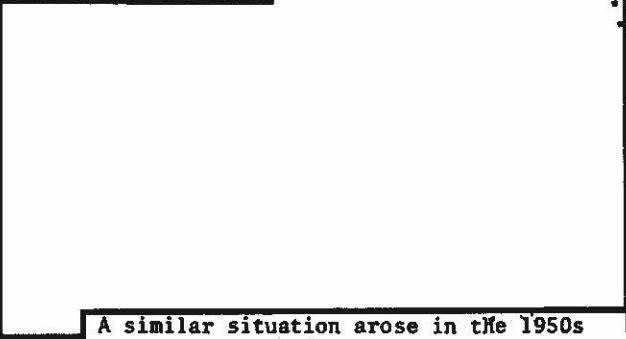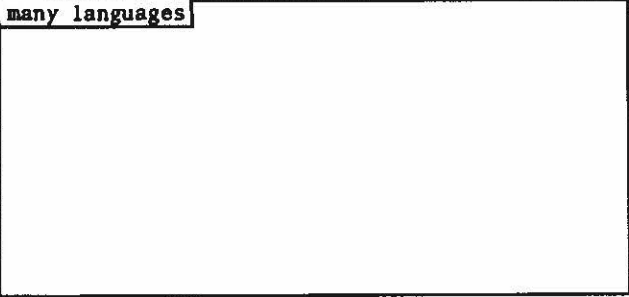
A similar situation arose in the 1950s

French used to be of great importance in the study of the languages of what was once French Indo-China, but today English-language works on Vietnamese, Cambodian, and Lao are far superior to those available in French. Finally, French is of considerable importance in the study of several languages of the Near East, particularly ⎕

Despite the fact that Germany lost her colonies after World War I, German continues to be an important tool language. ⎕

The third major tool language is, of course, Russian. It is not at all surprising to learn that the Russians have produced a large number of excellent bilingual dictionaries in the languages of Eastern Europe such as Hungarian, Bulgarian, etc., but it is truly startling to learn the extent of their coverage of the languages of Asia, particularly India and Southeast Asia. One can⎕ examine a large number of impressive dictionaries of Hindi, Bengali, Tamil, Sinhalese, Thai, Burmese, Cambodian, Chinese, and other languages. In many instances the Russian dictionaries are much larger and far more recent than any West European equivalents. In addition, many of the Russian dictionaries contain an added feature which greatly enhances their value: a supplement containing a concise but very adequate summary of the chief grammatical features of the language under study. Also, if the foreign language has a complex writing system, this is analyzed in detail. For many languages, such as Bulgarian, Armenian, and others, there exists a specialized type of dictionary, called a phraseological dictionary, which, as the term implies, is more concerned with phraseology than isolated words. Finally, most of the Soviet-produced dictionaries are beautifully printed and bound and, by Western standards, incredibly cheap. What has been said above applies to the languages of Europe and Asia, not Africa. Russian dictionaries of African languages are few (Amharic, Somali, and Luganda) and, except for the Amharic, not very impressive. But this situation will doubtless only be temporary.

Apart from these three major tool languages, a few other European languages have, to a lesser extent, functioned in this capacity. In the early days of the Agency's ⎕ the only useful dictionary and grammar were in Italian. There is now a good ⎕ dictionary, but it

still has not completely displaced ⎕ a magnificent work which, in the author's words, was ⎕ Portuguese is important for the study⎕

Dutch dictionaries and grammars are useful for the study of Malay, Javanese, Indonesian, and other languages of the South Pacific.

To my knowledge, Spanish has never been important as a tool language, but potential political developments could alter this situation drastically. Since the death of Franco, restrictions on the use of Catalan and Basque have been gradually eased, and newspapers and other printed matter are now being published in both languages. There are 6 million Catalan speakers in northeastern Spain around Barcelona and three-quarters of a million Basques in the north around Bilbao. Both groups have for centuries been tenaciously dedicated to the preservation and perpetuation of their language and culture, and both have exhibited revolutionary and extreme separatist tendencies. In the event of a fragmentation of Spain, both Catalan and Basque would be of interest to us, and both would have to be studied in large measure through the medium of Spanish. In Spanish America there are two major Indian languages: Quechua, the language of the Incas, spoken by about 7 million people in Peru, Bolivia, and Ecuador; and Guarani, spoken by most of the 2½ million inhabitants of Paraguay. It is very doubtful, however, that either of these will pose a problem for us now or in the foreseeable future.

The tool languages thus far discussed have all been European, but Chinese and Japanese must be reckoned as "sleepers" which sooner or later will become fully as productive as the languages of Europe. And when this happens, God help us.

In the previous discussion, most of the emphasis was on bilingual dictionaries, but there are also many excellent reference grammars, courses, and readers in languages other than English. Most of these are far more finite and amenable to translation than the dictionaries, but they do present problems that do not occur in ordinary translation. To take an example, a German-language ⎕ grammar with which I am familiar quite naturally compares and contrasts features of ⎕ with German, so, obviously, any English-language rendition must not only be a translation but an adaptation to make it suitable for readers of English. In other cases an otherwise valuable work might be unduly verbose and diffuse, so what is needed is not only to translate and adapt it, but to submit it to a process analogous to dehydration. But all such translations are a makeshift; the ideal is to have linguists who are familiar with both the tool language and the language of interest.

The point that I should like to stress in conclusion is that there is a sharp variation in our ability to use the various tool languages.
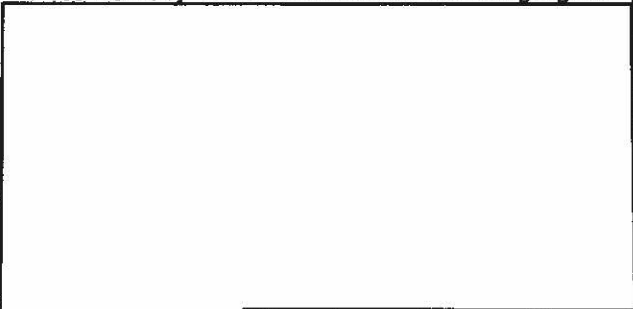
There is another possible reason. Many high school and college graduates have, in the course of their education, acquired some knowledge of French or (less often) German because these languages are standard subjects in many high schools and colleges. But almost nobody studies Russian unless he plans to become a specialist in Russian.

It must be stressed that the ability to use a bilingual Russian dictionary requires a rather good knowledge of Russian, not just a smattering. The readings given in a dictionary often consist of a series of isolated words with no context at all. Furthermore, even with an abundance of time and money, it is almost impossible to translate a bilinguial dictionary unless those doing the translating have an excellent knowledge of both languages (e.g., Burmese and Russian), something which rarely happens. To illustrate this point I should like to cite the word [   ] which occurs in a [   ] German glossary I have been working on. The word is equated with five German words, all of which are without context. On consulting a German-English dictionary one finds that each German word is equated with several English words. So [   ] may mean: [   ]
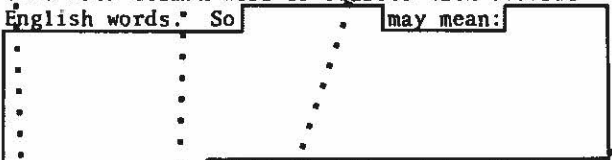
This may be an extreme case, but there are enough such extreme cases to make things mighty rough for the "translator." One wonders how machine translation would fare with problems of this type.

The solution of problems arising from our inability to handle tool languages would seem to lie either in the direction of providing Russian training, perhaps of a very special type, to linguists specializing in Asian or African languages, or of diverting competent Russian linguists to the study of Asian and African languages. There are also very obvious implications for job descriptions, PQEs, and recruitment.

(C)

WHICH TAPE HAS THE INTELLIGENCE?

PROJECT CRISELOR: A Study of Selection and Rejection
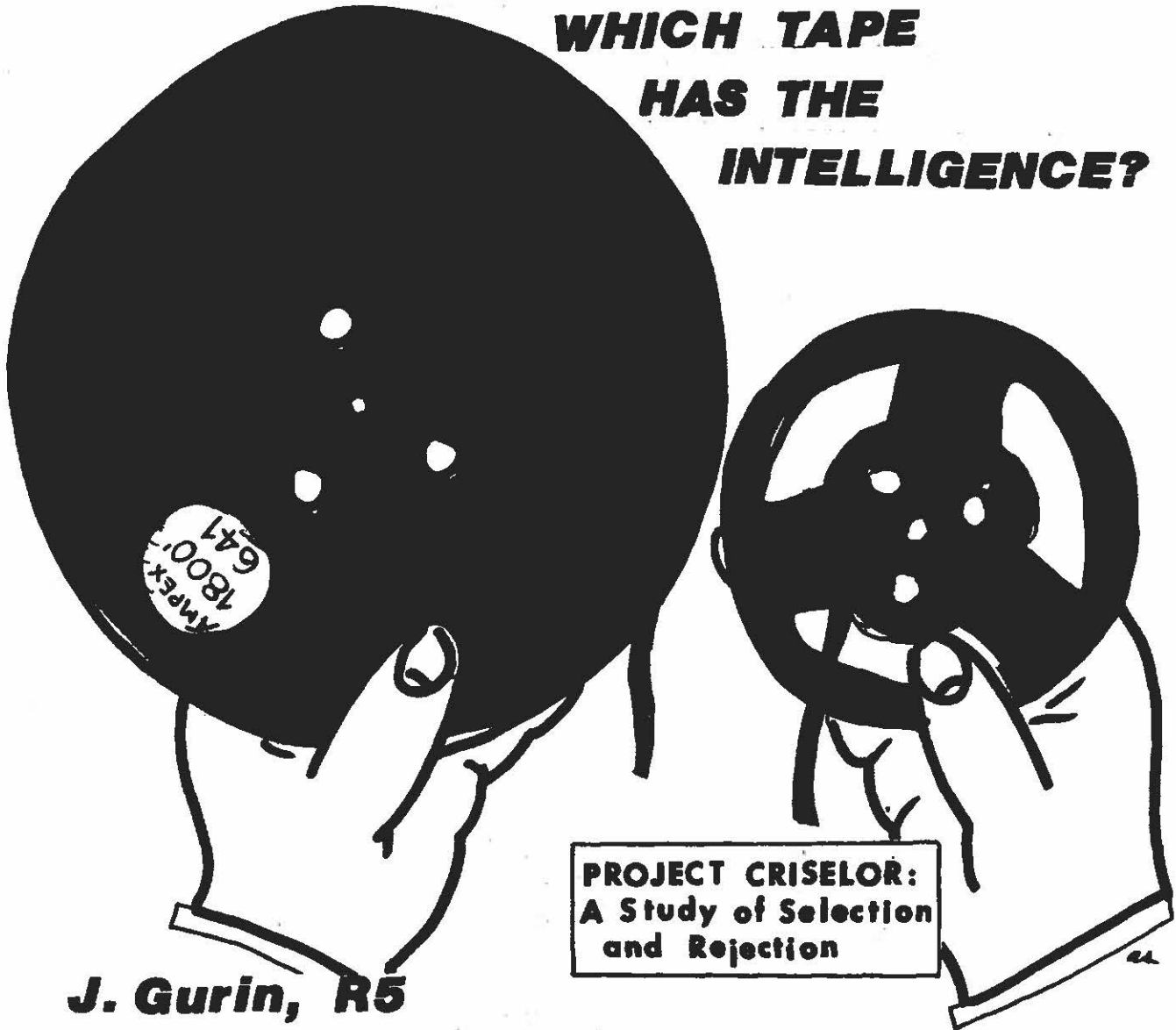
J. Gurin, R5

A certain excitement seems to have been generated by a modest project, devised in the Office of Research (R5) and under way in A6. It purports to discover how decisions are made either to retain or to reject intercepted conversations. No doubt the excitement is related to the desire, or even anxiety, of those who are seeking some way in which the machine may help solve the problem of too much traffic for too few transcribers. Project CRISELOR (*CR*iteria for *SEL*ection *Or R*ejection), under way for several months, is attempting to pin down the precise factors which enter into the human decision-making act, in the hope that much will be learned that will be of value in the eventual automation of some part of the process.

The problem of discovering how we select or reject for intelligence reporting is certainly not new, but perhaps it has never been scrutin- ized quite as carefully as this. Selection/rejection is, of course, a characteristic process in any situation in which material is being sifted for that proportion which is most relevant. If the volume is large enough and time is short, we resort to shortcuts, looking for indicators which will permit correct decisions without spending a great deal of time in making the judgments. In some cases the shortcuts are deliberately chosen and are easily identifiable, but often we do not know exactly how our choices are made. How does one decide that a conversation is of potential value? Who makes the decision that causes one class of conversations to be retained and another discarded? What part of the decision process is likely to be routine, mechanical, and precisely described, and what part inspirational, instinctive, and difficult to predict? How much of the process is left to the analyst and transcriber and how

much is dictated by rigid priority statements,
handed down from above?

It may seem odd that we have never examined
our selection mechanism in this way until now,
but perhaps the need was never felt so acutely.
Whether we succeed eventually in automating some
significant portion of the process or not, we
should gain some useful insight into the process
and benefit in some other ways as a result of
this investigation.

Non - Responsive

NSA·crostic No. 8
By A.J.S.

> The quotation on the next page was taken from a published work of an NSA-er. The first letters of the WORDS spell out the author's name and the title of the work.

DEFINITIONS

A. Medieval English style of architecture in which vertical lines predominate

127 116 174 144 59 152 146 207 5 37 82 187 50

B. Antiquated

120 71 131 85 103 147 30 25 17 201 193 69

C. Experience

20 149 226 137 79 96 6

D. Info

136 126 98 35 163 154 209

E. Something that is supposed to have a settling influence on straying husbands ("It's an old trick, but it just might work!")

74 219 42 27 128 185 171 124 88 83

F. Prevaricator

167 100 192 173 22 217

G. Excessively ornate or intricate

60 133 101 10 178 29

H. State (abbrv)

222 205

I. One from Column B (3 wds)

51 216 223 75 70 153 4 176 197 87

J. When the guy in the fancy French restaurant pointed to this on the menu and the waiter gasped, "Eez zat *all*, m'sieur?", he answered, "And a cuppa coffee!"

7 106 91 132 39 138 140 181 72

K. Low

111 208 168

L. Decide a matter in dispute

198 11 212 145 64 105 95 117 34

M. Engage in fondling and kissing

134 55 107 189

N. Where you'd go in Boston for real *pasta fazool* (2 wds)

224 210 129 108 45 142 3 121

O. Miser

220 52 125 184 180 115 112

P. Polynesian wood or stone image

158 214 151 76

Q. Why it's difficult for Middle East archeologists to agree on classification criteria (7 wds)

43 215 109 177 130 57 195 190 119 148 143 86 41

90 23 175 28 156 218 48 1 14 166 94 186

8 169 54 104 65 199

R. Character in Mozart's *The Magic Flute* (his girl friend has a very similar name)

204 18 61 32 194 92 139 81

S. Silly, peevish person (Brit)

113 155 56 162

T. Ancient language of eastern Asia Minor

$$\overline{159}\ \overline{53}\ \overline{170}\ \overline{80}\ \overline{196}\ \overline{183}\ \overline{16}$$

U. Piece of music for practicing some special point of technical execution

$$\overline{213}\ \overline{66}\ \overline{165}\ \overline{12}\ \overline{172}$$

V. Plurality of wives or husbands

$$\overline{49}\ \overline{73}\ \overline{179}\ \overline{122}\ \overline{26}\ \overline{2}\ \overline{84}\ \overline{191}$$

W. Echo

$$\overline{164}\ \overline{97}\ \overline{33}\ \overline{182}\ \overline{62}\ \overline{38}\ \overline{157}\ \overline{15}\ \overline{40}\ \overline{206}\ \overline{21}$$

X. More irritable

$$\overline{46}\ \overline{110}\ \overline{19}\ \overline{67}\ \overline{160}\ \overline{118}$$

Y. Divert

$$\overline{221}\ \overline{9}\ \overline{203}\ \overline{44}\ \overline{77}\ \overline{99}$$

Z. Hot wind from Libya that blows on the Northern Mediterranean coast

$$\overline{13}\ \overline{47}\ \overline{161}\ \overline{68}\ \overline{135}\ \overline{188}\ \overline{36}$$

Z_1. Specific scientific phenomenon named usually for its discoverer

$$\overline{78}\ \overline{123}\ \overline{211}\ \overline{225}\ \overline{202}\ \overline{58}$$

Z_2. Craving for candy (2 wds)

$$\overline{93}\ \overline{141}\ \overline{31}\ \overline{63}\ \overline{150}\ \overline{200}\ \overline{24}\ \overline{114}\ \overline{89}\ \overline{102}$$

1 Q	2 V	3 N	4 I		5 A	6 C	7 J	8 Q	9 Y	10 G	11 L	12 U	13 Z		14 Q
15 W	16 T		17 B	18 R	19 X	20 C	21 W		22 F	23 Q	24 Z_2	25 B	26 V	27 E	
28 Q	29 G		30 B	31 Z_2	32 R	33 W	34 L		35 D	36 Z	37 A	38 W	39 J		40 W
41 Q		42 E	43 Q		44 Y	45 N	46 X	47 Z	48 Q		49 V	50 A	51 I	52 O	53 T
54 Q	55 M		56 S	57 Q	58 Z_1	59 A	60 G	61 R	62 W	63 Z_2	64 L	65 Q	66 U	67 X	68 Z
69 B		70 I	71 B	72 J		73 V	74 E	75 I	76 P	77 Y	78 Z_1	79 C		80 T	81 R
82 A	83 E		84 V	85 B		86 Q		87 I	88 E	89 Z_2		90 Q		91 J	92 R
93 Z_2	94 Q	95 L	96 C	97 W		98 D	99 Y	100 F	101 G	102 Z_2		103 B	104 Q	105 L	106 J
107 M	108 N	109 Q	110 X		111 K	112 O		113 S	114 Z_2		115 O	116 A	117 L		118 X
119 Q	120 B	121 N	122 V		123 Z_1	124 E	125 O		126 D	127 A	128 E	129 N	130 Q	131 B	132 J
133 G	134 M		135 Z	136 D	137 C	138 J	139 R		140 J	141 Z_2	142 N	143 Q	144 A		145 L
	146 A	147 B	148 Q	149 C	150 Z_2		151 P	152 A	153 I	154 D		155 S	156 Q	157 W	158 P
159 T	160 X	161 Z		162 S	163 D	164 W	165 U	166 Q		167 F	168 K	169 Q		170 T
171 E	172 U		173 F	174 A	175 Q	176 I	177 Q		178 G	179 V	180 O	181 J	182 W	183 T	
184 O	185 E		186 Q	187 A	188 Z	189 M		190 Q	191 V		192 F	193 B	194 R	195 Q	
196 T	197 I		198 L	199 Q	200 Z_2	201 B	202 Z_1	203 Y	204 R	205 H	206 W	207 A	208 K	209 D	
210 N	211 Z_1		212 L	213 U	214 P	215 Q	216 I		217 F	218 Q	219 E	220 O	221 Y	222 H	223 I
224 N	225 Z_1	226 C													A.J.S.

(Solution next month)