

ASSESSMENT OF SYSTRAN

by Jacques CHAUMIER

Bureau Marcel van Dijk

As part of its action programme on multilingualism, the Commission of the European Communities has had the SYSTRAN translation system developed.

SYSTRAN has been assessed from two angles:

- a) performance and cost;
- b) translation quality.

The latter assessment is still being carried out and has been completed on approximately 50% of the test sample.

The test was carried out on translations from English to French dealing with food science and technology.

The test sample comprised:

- 20 abstracts from Food Science and Technology Abstracts (approximately 3 000 words);
- 2 articles from scientific journals (approximately 4 000 words);
- 4 Community documents (approximately 4 200 words).

The dictionary used contained:

approximately 25 000 words or inflections of words
4 500 expressions.

Two runs were carried out on the IBM computer at the Commission's Computer Centre, with updating of the dictionary after the first run.

See Fig. 1 for the machine times recorded.

At the same time the Commission's translation department translated the sample texts manually (translation plus revision) and revised the automatic translation produced by the first computer run.

The various cost items were assessed for the two processes (automatic translation and manual translation). See fig. 2 for the overall translation costs.

The time taken to revise the automatic translation, as shown in these figures, ought probably to be corrected, since this part of the work was carried out at abnormally high speed.

The resulting summary of costs can be seen in fig. 3.

From an economic point of view, it seems that SYSTRAN

- is competitive with manual translation plus revision of normal texts carried out in a large translation department such as that of the Commission;
- is not competitive with manual translation, without revision, of normal texts done by freelance translators (Bfrs 150 to 250 per 100 words).

The tests also proved that SYSTRAN is perfectly reliable from the point of view of the data processing system.

Quality is assessed at the following two levels:

- overall assessment: overall intelligibility;
- linguistic assessment.

Intelligibility was measured for all the 507 sentences of the sample in the six possible texts:

- original text;
- manual translation;
- automatic translation (first run);
- automatic translation (second run);
- revised version of manual translation;
- revised version of automatic translation.

Each sentence was graded according to the following scale:

- 0 poor or unusable
- 1 rather poor
- 2 fairly good
- 3 good.

See fig.4 for the results obtained for 362 sentences (= 71% of the sample). Since the level of intelligibility depends heavily on the reader, the possibility of introducing a suitable weighting coefficient is being studied. Linguistic assessment is still being carried out. It involves a sample of a quarter of the sentences from all the texts submitted for translation.

This assessment is based on the following criteria:

- structural analysis of the sentences:
 - recognition of the noun group
 - subject
 - verb group
 - complements (of noun and verb).
- translation of sections of the text:
 - recognition and translation of verbs
 - negatives
 - nouns
 - articles
 - adjectives
 - etc....

With regard to sentence structure, figure 5 shows the results obtained for 62 sentences (=48.92% of the quarter sample).

Lastly, a final test was carried out on the sample to determine the extent of revision on both manual and automatic translations. Fig.6 shows the results, which refer to the number of words added, deleted or changed in relations to the number of words of the translation.

The following run times were noted:

		Elapsed time	C.P.U.Time	Number of words dealt with by SYSTRAN
1st test 18/10/76	Single step	17'85"	4'26" 47	11 340
	1st step	5'83"	1'19" 52	2 662
	2nd step	6'04"	1'18 51	2 390
	3rd step	5'82"	1'07" 03	2 020
	4th step	9'97"	1'19" 79	2 401
	5th step	2'16"	0'36" 83	837
	6th step	2'67"	0'32" 55	495
	7th step	2'92"	0'32" 76	535
	Total of 7 steps	35'41"	5'66" 99	11 340
2nd test 5/11/76	Single step	16'04"	4'29" 44	11 232

fig. 1

Cost items	Cost price per 100 words (in Bfrs)						%					
	Abstracts		Texts		Abstracts		Texts		Manual	SYSTRAN	Manual	SYSTRAN
	Manual	SYSTRAN	Manual	SYSTRAN	Manual	SYSTRAN	Manual	SYSTRAN				
Acquisition	-	95	-	95	-	-	29.1	-	-	-	29.6	
Translation	800	33	296	33	74.3	10.1	10.1	57.6	10.3	10.3	10.3	
Typing	37	-	38	-	3.4	-	-	7.4	-	-	-	
Revision	203	161	142	155	18.9	49.4	49.4	27.6	48.3	27.6	48.3	
Typing	37	37	38	38	3.4	11.4	11.4	7.4	11.8	7.4	11.8	
TOTAL	1.077	326	514	321	100.0	100.0	100.0	100.0	100.0	100.0	100.0	

fig. 2

COST PRICE PER 100 WORDS (IN Bfrs.)

ABSTRACTS		TEXTS	
Manual	SYSTRAN	Manual	SYSTRAN
1 077	487	514	476

fig. 3

INTELLIGIBILITY (MAXIMUM = 3.00)

	OT	MT	AT 1	AT 2	RMT	RAT
Overall (362 sentences)	2.90	2.79	1.40	1.95	2.94	2.85
Abstracts (145)	2.82	2.63	1.42	1.99	2.97	2.76
Scientific documents (48)	3	2.77	1.46	1.98	2.83	2.83
Community texts (67)	2.87	2.90	1.42	1.85	2.88	2.84

fig. 4

Average number per sentence

Noun groups		
number		2.27
clearly defined		2.18
percentage		96.04%
Verb groups		
number		1.27
tense recognized		1.14
percentage		89.76%
case recognized		1.17
percentage		92.13%
Verb complements		
number		2.05
recognized		2.01
percentage		98.05 %
Invariable words		
number		1.88
correctly translated		1.68
percentage		89.36%

Fig. 5

Extent of revision

fig. 6.

	Original text	Manual translation	Revision of manual translation	Automatic translation	Revision of Automatic translation
Overall	1.111	1.446 4.84%	70	1.364 26.56%	362
Doc.1 (Abstracts)	408	600 4.50%	27	538 26.58%	143
Doc.2 (scientific)	474	572 6.64%	38	565 25.84%	146
Doc.5 (Community)	93	127 0%	0	115 26.09%	30
Doc.7 (Community)	136	147 3.40%	5	146 31.62%	43