

The Economics of Translation
Copenhagen, 1980

Ladies and Gentlemen :

In Edinburgh, in Autumn 1978, I gave an overall view on computer-aided translation, its actual possibilities and its development potential. (1)

In Cranfield, then, in the Summer of 1979, I presented a paper on the applicability of multilingual tools, including thesauri and translation systems, in the field of scientific and technical information. (2)
Both papers relied heavily on the excellent surveys and the evaluation work performed by Georges Van Slype. (3)

Today's paper is again based on a market study performed by PACTEL and Bureau M. Van Dijk, and Georges Van Slype took a large part in its preparation.

This worldwide study of the market of human and computer-assisted translation is Part 5 of a series of studies undertaken by the European Commission in the field of multilingualism.

As you all know, the European institutions in Brussels and Luxembourg are employing close to 2000 officials in their linguistic services : translation, terminology and interpretation. This number should increase by another 5-600 as Greece, Spain and Portugal are applying for membership. This justifies the Commission's efforts to get some relief by introducing modern computer-based methods into its services.

As you also know, Euronet-Diane has gone into operation early this year, involving more than a hundred data bases. 15% of these data bases are multilingual to some extent, and 85% are monolingual (45% in English, 40% in French, Italian and German). This means frustration for potential users in a number of member countries, and it is another justification for Commission initiatives in the field of multilingual devices.

The Commission's action plan now involves the spending of more or less 1 million accounting units (600000 pounds or 6 million FF) per year. Such a huge action deserves careful preparation, and it was in fact planned in five phases :

Phase 1 was a survey of the inter-language communication situation, including the traditional translation activities.

Phase 2 was an inventory of existing methods and devices to overcome the language barrier. This culminated in the Third European Congress in 1977 in Luxembourg. (5)

Phase 3 included a series of evaluation studies, including assessment of cost and quality of the products of existing systems and analysis of the technical feasibility and economic viability of proposed new systems. This phase yielded a number of reports giving detailed cost and quality figures.

Phase 4 involved studies of the development potential of existing systems and devices and of proposed new systems.

Phase 5 was subdivided into a qualitative assessment of existing needs for inter-language communication, and a large action aiming at an exhaustive quantitative analysis of existing and potential needs.

Reports resulting from phases 2, 3 and 4 are referenced in the bibliography annexed to this paper.

The following tables were produced as a result of work in phase 5 and to some extent, in phase 1.

Table 1 : International languages in 1973

Language	Number of persons (in millions) having this as their				
	mother tongue	official language	administr. language	TOTAL	%
English	320	160	875	1 355	33,9
Chinese	800	20	-	820	20,5
Russian	150	-	205	355	8,9
Portuguese	120	145	-	265	6,6
Spanish	210	30	-	240	6,0
French	80	90	55	225	5,6
Arabic	130	-	-	130	3,3
German	105	-	-	105	2,6
Subtotal	1 915	445	1 135	3 495	87,4
Others	-	-	-	505	12,6
TOTAL	-	-	-	4 000	100

Source : BRETON, 1976.

Table 2 : Languages used in international organizations

Name	English	French	Spanish	Russian	Arabic	Chinese	German	Portuguese	I, FL, DK
United Nations	E	F	S	R	A	C			
World Health Organization	E	F	S	R	A				
International Labour Office	E	F	S	R	(A)		D		
ESCO	E	F	S	R	A				
Food and Agricultural Organization	E	F	S		A	(C)	(D)		
European Community	E	F					D		I, NL, DK
OECD	E	F					(D)		
WTO	E	F							
ERN	E	F					D		
Arabic League	E	F			A				
Organization of American States	E	(F)	S					F	
Organization of African States	E	F			A				

Table 3 : Translation of books (by country)

Country of publication	1971	1972	1973	1974
Germany (West)	4 293	2 767	6 458	6 638
Soviet Union	4 729	3 957	4 402	4 270
Spain	3 147	3 181	4 489	3 780
Japan	228	2 178	2 284	2 485
France	1 992	2 128	1 934	2 485
Italy	2 285	1 861	2 095	1 905
U.S.A.	2 285	2 174	1 968	1 866
Denmark	3 037	1 441	1 415	1 866
Netherlands	1 970	1 838	1 816	1 846
Sweden	1 245	1 435	1 599	1 710
Yugoslavia	1 460	969	1 482	1 526
Czechoslovakia	1 292	1 276	1 437	1 499
United Kingdom	727	731	682	1 469
Brasil	1 037	173	1 684	1 311
Turkey	801	979	991	1 263
Norway	1 039	1 473	1 004	1 168
Hungary	1 053	1 037	1 083	1 163
Subtotal 17 countries	34 620	29 598	35 341	38 250
66 other countries	8 356	8 019	11 697	8 006
TOTAL	42 976	37 617	47 038	46 256

Source : Unesco statistical yearbooks

Table 4 : Translation in the UN and the EC

UN Statistics (Geneva), 1978

Target language	Translation (in 1000 words)
Arabic	2 522
Chinese	2 081
English	5 388
French	13 728
Russian	11 142
Spanish	9 203

EC Statistics (Brussels), 1978

Target language	Translation (in 1000 words)
French	16 104
German	20 701
Italian	16 097
Dutch	14 717
English	22 314
Danish	14 149
Other	1 160
TOTAL	105 242

Table 5 : Translation requests for ITC (Delft)

Y E A R S	Requests	Source language distribution					Response			
		Russian	Other East Europe	Japanese	Chinese	Other	Trans- lations located	Trans. Supplied	not located	response rate
57	9.722	6.375	1.409	1.419	459	60	3.737	1.638	5.985	38,5%
58	10.072	6.338	1.601	1.859	214	60	2.602	832	7.470	25,8%
59	9.937	6.014	1.413	1.864	49	592	1.819	326	8.118	18,4%
70	11.505	5.898	1.394	3.216	131	866	2.127	423	9.378	18,5%
71	11.950	6.796	1.432	3.041	263	418	2.277	857	9.673	19,1%
72	8.091	4.692	1.119	2.004	119	157	1.204	292	6.887	14,9%
73	8.331	4.979	1.451	1.678	15	208	955	144	7.376	11,5%
74	8.095	4.665	1.341	1.892	26	171	1.180	176	6.915	14,6%
75	8.559	5.029	1.363	1.987	54	126	1.915	127	6.644	22,4%
76	8.214	4.746	1.374	1.885	45	164	1.933	148	6.281	23,6%
77	6.801	3.661	1.154	1.819	44	123	1.429	174	5.372	21,0%
78	6.774	3.514	1.198	1.839	49	174	1.415	127	5.359	20,9%

Table 6 : Translation cost and tariffs

Countries and organisations	Date	Cost, in FF, per 100 words	
		into own language	into foreign language
<u>Belgium:</u> Chambre Belge des Traducteurs (European languages) (non-Eur. languages)	1978	20-28 36-46	22-29 65-83
<u>Canada:</u> free-lance translators high-quality translation	1979	21-25 54-64	
<u>France:</u> Société Française des Traducteurs (European languages) (exotic languages) Unesco	1979 1977	31 64 50	46-52 96-109
<u>Germany:</u> Bundessprachenamt free-lance	1979	65-109 76	
<u>United Kingdom:</u> technical translations	1979	15-30	
<u>Sweden:</u> Translators' Union	1979	70	
<u>Switzerland:</u> Association des Interprètes et Traducteurs	1979	33	
<u>U.S.A. :</u> American Translation Association free-lance	1979	8,50 15-25	

Table 7 : CCE Translation Statistics 1978 , Bruxelles
(number of pages)

Source language	Target language							
	F	D	I	N	E	DK	other	TOTAL
French	445	56 487	45 067	41 323	62 706	38 021	4 201	248 250
German	16 934	544	4 497	3 768	12 684	3 500	69	41 996
Italian	7 530	3 166	13	1 344	4 040	762	27	16 882
Dutch	6 863	2 908	920	52	5 255	625	3	16 626
English	26 415	18 259	13 354	11 510	580	13 766	338	84 222
Danish	3 458	1 180	494	678	2 885	18	-	8 713
Other	2 772	261	42	193	1 106	5	-	4 379
TOTAL	64 417	82 805	64 387	58 868	89 256	56 697	4 638	421 068

Table 8 : Language Knowledge in England
(3249 British Lending Library users)

Language	All users (N= 3249)		Users below 26 (N= 684)	
	Reading Knowledge	with dictionary use	Reading Knowledge	with dictionary use
French	12,0 %	71,8 %	6,9 %	73,4 %
German	4,3	40,6	2,5	31,1
Russian	0,8	5,7	0,4	2,9
Japanese	0,2	0,7	0,3	0,4
Spanish	1,5	5,0	1,0	4,0
Italian	1,0	5,2	0,6	2,7
Dutch	0,5	1,5	0,1	1,0

Reading knowledge, by subject field

Language	Science and technology N=2458	Social science N=461	Humanities N=190	Other N=114	All subjects N=3223
French	10 %	11 %	37 %	12 %	12 %
German	3	5	15	2	4
Russian	0	2	4	1	1
Japanese	0	0	0	0	0
Spanish	1	3	5	3	2
Italian	0	1	6	1	0
Dutch	0	1	1	0	0

Source: ELLEN (1979)

A number of statements can be derived from these tables, some unexpected.

From table 1 : Among the 3000 languages and 6000 dialects spoken in the world, English has a very special status, since it is spoken by 34% of the world population, and seven other languages, spoken by another 54% of the world population, can be considered as the world's main languages.

From table 2 : Two languages only are in use with all international organisations and can be termed truly international languages : English and French. Four others are extensively used in some, but not all these organisations : Spanish, Russian, Arabic and German.

From table 3 : A very limited percentage (less than 10%) of all published books are re-published in another language. More books are translated from English into German, Russian and Spanish than from these languages into English.

The corresponding figure for journal articles is less than 1%.

From table 4 : More text is translated in the European institutions than in the UN agencies. This is due to the fact that much of the Common Market work has a legal bearing in its member states.

From table 5 : The number of requests addressed to the International Translations Centre, as well as its success in locating existing translations, are diminishing as national services are improved or created and are in position to satisfy their own users. Most of the translations I are from Russian into Western languages.

From table 6 : The average cost of human translation can be said to be around 50 FF = 350 FB = 5 £ = 12 \$ per 100 words, in spite of widely differing figures for the United States (and Israel), where abundant translating manpower faces a low-need situation.

From table 7 : The EC translation department in Brussels, which is the largest in the world, produces more than 400000 pages in Brussels (and another 100000 in Luxembourg). While more than one half of these have French as source language, all six official languages have equal importance as target languages.

From table 8 : Language knowledge of British research personnel is alarmingly low less than 20% can read a scientific text in another language.

From all this, from the literature on human and computer-aided translation and the Commission's own experience in the field of information handling I conclude that there are a thousand language barriers and there is a real, large, fundamental unsatisfied need for translation

- in a variety of language couples,
- in various fields of knowledge,
- for a variety of document types, and
- in various ranges of cost and quality.

On the other side, there are a number of methods, devices and systems in existence or, under development, which are likely to correspond to this variety of needs.

- Multilingual thesauri, managed and updated using increasingly efficient software, are the solution to the documentalist's inter-language problems.

- multilingual terminology banks, made accessible through international communication networks, will greatly enhance the productivity of the human translator.

- The new pocket translator will be fitted with enlarged dictionary chips and satisfy even the more demanding tourist.

- Interactive terminals, supplied with dictionaries and elementary syntactic routines, will enable less specialized persons to produce translations of acceptable quality.

- The TITUS system quite adequately translates controlled-syntax abstracts from tropical agriculture and the textile industry simultaneously into several languages.

- The TAUM system efficiently translates English-Canadian weather bulletins into French for the Province of Quebec.

- The SYSTRAN system economically produces impressive amounts of translations from Russian into English for the U.S. Air Force in Dayton (Ohio).

One basic rule underlies the choice operated by an array of users : They chose the most economic device for whatever they have to achieve.

The immediate conclusion is that all existing devices have a useful role to play, and that no single system is likely to supersede them all.

Since the actual cost depends on

- (a) the quality level required (or the delay allowed) and
- (b) the amount of text to be translated, it will be necessary
 - (a) to develop reliable methods and criteria for assessing translation quality, and
 - (b) to perform an exhaustive market study to determine for which languages and document types the volume of text to be translated justifies the development of a translation system or, better, of additional dictionaries for one or more existing translation systems.

The actual cost of translating text by computer is well known by the owners of existing systems. Some keep their cost data confidential, but the Commission's evaluations of the Systran system were widely publicized.

The cost of making text machine-readable is around BF 100 (15 FF or 1,5 £) per 100 words through keypunching or magnetic encoding. The cost of reformatting existing text material using a tailor-made interface programme diminishes as the volume of text increases.

The cost of pre-editing existing text prior to the input is of the same order of magnitude. The combined cost of input + pre-editing can be lowered through the use of text processing equipment.

The cost of machine translation by Systran is approximately BF 35 (5 FF or 0,5 £) per 100 words if the text sample is not smaller than 10 000 words. The cost is significantly lower for larger batches.

The complex system of the future might require a slightly higher computer capacity, but since computer cost goes down all the time, the translation costs are not expected to undergo significant variations.

The cost of post-editing by human linguists was around BF 150 (20 FF or 2 £) per 100 words in the Systran evaluation tests. This can be considered as a maximal cost, corresponding to a situation where the revisor critically examines every word in the translated text. As the translation quality improves and the revisors acquire some experience, including experience of text processors, the post-editing cost is expected to go down to less than half that amount.

Typing, which now costs BF 35 (5 FF or 0,5 £) per 100 words, will soon be replaced by print-out from text processors at virtually no cost at all.

The main advantage of text processors is not so much the reduction of cost of computer-aided translation, but the reduced time requirement, especially if the text processor is linked to the mainframe computer. A high-quality translated text can thus be obtained within a few hours, the main time component remaining the input time.

The very low cost of machine translation itself makes it particularly suitable for applications where

- the text to be translated is in machine-readable form
- only an unsophisticated raw translation is required.

This applies to information gathering by scientists and engineers having no knowledge of the source language of their documents. It also applies to information gathering from large bibliographic data bases through the Euronet-DIANE network, where the user can receive a translation of the abstracts retrieved in a search in his mother tongue within a few hours.

In conclusion it can be stated that machine translation is very cost-effective today; for a number of applications requiring speedy handling without excessive quality requirements.

It breaks even with human translation for any application (such as published papers or legal documents) requiring high-quality products. But psychological barriers remain to overcome for any application involving pre- or post-editing by trained linguists.

Bibliography

- (1) Li.N. Rolling : Computer-aided translation today and tomorrow.
Proc. 39th FID Congress, Edinburgh, Sept. 1978. pp.164-169
(re-published as euro-article in German and French)
- (2) L.N. Rolling : The second birth of machine translation, a timely event
for data base suppliers and users.
7th Cranfield Conference, July 1979, 7p.
- (3) G. Van Slype : Critical study of methods for evaluating the quality
of machine translation.
CCE report, Nov. 1979, 180p.
- (4) Bureau M. Van Dijk and P.A. : Mieux traduire pour mieux communiquer -
étude prospective du marché de la traduction. 1980, 250p.
- (5) C.C.E. : Overcoming the Language Barrier.
Proc. 3d European Congress on Information Systems and Networks.
1977, 2vol., 700p.