

Où en sont les travaux du nouveau dictionnaire français-hongrois ?

(État des travaux et documentation)

Miklós PÁLFY

L'initiative d'un nouveau dictionnaire français-hongrois a été rendue nécessaire par le fait que le seul outil lexicographique en usage actuellement, le dictionnaire de Sándor ECKHARDT, est très dépassé dans sa conception et défectueux à bien des égards.

Le protocole de rédaction des nouveaux dictionnaires français-hongrois/hongrois-français a été élaboré en automne 1991¹. Le projet du dictionnaire français-hongrois remonte à l'automne 1992, puis, ce projet a été modifié en 1994². L'aspect informatique a reçu sa forme définitive en été 1993³.

La présente documentation tient compte des résultats obtenus entre juin 1993, début des travaux concrets, et novembre 1994 :

1. Lettres et fichiers : proportions du dictionnaire français-hongrois (projet de 1992, modifié en 1994) ;
2. Aspect informatique : structure des articles, codes et symboles ;
3. Exemple de codage : texte non formaté ;
4. Exemple de codage : texte formaté suivi des données du logiciel ;
5. État des travaux en novembre 1994.

1. Par Vilmos BÁRDOSI, Miklós PÁLFY, Jean PERROT et Tamás SZENDE.

2. Par Miklós PÁLFY.

3. Le système de codage a été élaboré par Mme Júlia PAJZS.

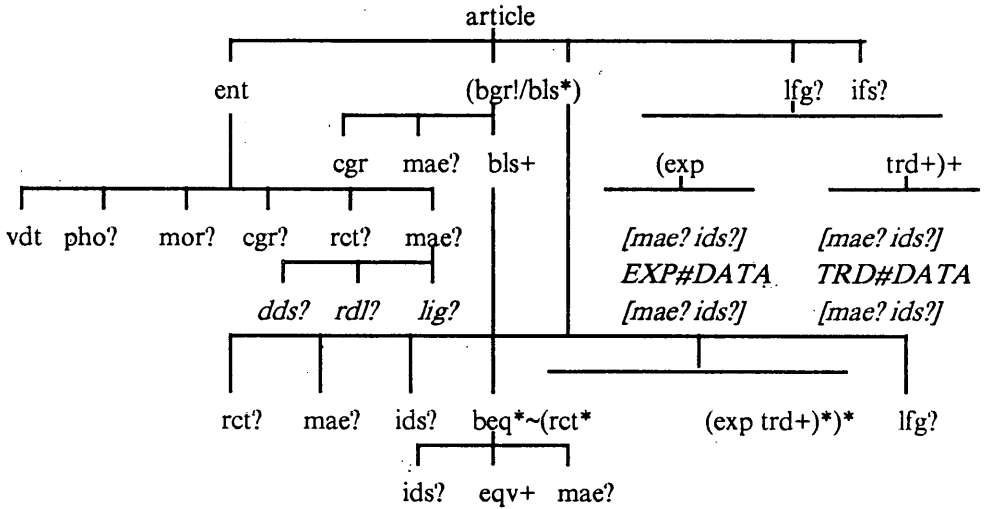
1. Les fichiers du futur dictionnaire français-hongrois
(Projet élaboré en automne 1992, modifié en automne 1994)

	Micro Robert		Robert & Collins		Le futur dictionnaire			
	pages	(%)	pages	(%)	(%)	Fichiers (maximum)	Entrées	
A	97	7,2	59	7,6	7,5	2	3000	A
B	56	4,1	31	4	4	6	1500	B
C	152	11,3	86	11,2	11,25	16	4000	C
D	91	6,8	58	<u>7,5</u>	7,5	11	2750	D
E	108	<u>8</u>	52	6,8	7,5	11	2750	E
F	56	4,1	34	4,4	4,25	6	1500	F
G	39	2,9	22	2,9	3	4	1000	G
H	28	2	12	1,6	2	3	750	H
I	56	4,1	30	3,9	4	7	1750	I
JK	15	1,1	9	1,1	1	1	250	JK
L	37	2,8	20	2,6	3	4	1000	L
M	79	5,9	47	6,1	6	8	2000	M
N	21	1,5	12	1,6	1,5	2	500	N
O	29	2,1	15	1,9	2	3	750	O
P	151	<u>11,3</u>	79	10,2	10,5	14	3500	P
Q	6	0,4	6	<u>0,8</u>	0,5	0,5	125	Q
R	98	7,3	64	<u>8,3</u>	8	11	2750	R
S	90	6,7	50	6,5	6,5	8	2000	S
T	78	5,8	46	6	6	8	2000	T
U	7	0,5	4	0,5	0,5	0,5	125	U
VW								VW
X	46	3,4	28	3,5	3,5	4	1000	X
YZ								YZ

100 % 140 35000

Un fichier est égal à 800 unités (= équivalents ou traductions). 140 fichiers = à peu près 112.000 unités. Cela correspond, grosso modo, à Robert & Collins ou à Micro Robert : donc à 35-36000 entrées environ. Quant à la taille, notre dictionnaire sera donc un dictionnaire « moyen » ; par contre, quant à la quantité d'informations supplémentaires et la « profondeur » des articles, ce sera plutôt un grand dictionnaire.

2. Aspect informatique (sans les articles de renvoi)



article	(ent, (bgr!/bls*), lfg?, ifs?)
ent	(vdt, pho?, mor?, cgr?, rct?, mae?)
bgr	(cgr, mae?, bls+)
bls	(rct?, mae?, ids?, beq*~(rct*, (exp, trd+)*)*, lfg?)
beq	(ids?, (eqv, gen?)+, mae?)
exp	([mae?, ids?], EXP#DATA, [mae?, ids?])
trd	([mae?, ids?], TRD#DATA, [mae?, ids?])

Abréviations

ent	entrée	mae	marques d'emploi
bgr	bloc grammatical	ids	indication de sens
bls	bloc sémantique	beq	bloc équivalent
lfg	locution figée	exp	exemple
ifs	information supplémentaire	trd	traduction
vdt	vedette	eqv	équivalent
pho	phonétique	gen	genre
mor	morphologie	dds	domaine de spécialité
cgr	catégorie grammaticale	rdl	registre de langue
lig	limitation géographique	rct	rection

Explication des signes

/	soit-soit : les deux	+	au moins 1
	<i>s'excluent</i> mutuellement	!	0 ou supérieur à 1
[]	position facultative	*	0 ou plus
~	et/ou	?	0 ou 1

absence de signe : un(e) seul(e)

italique : structure identique aux autres blocs **mae?** ou **exp-trd**

Remarque : si les différentes données sont valables pour tous les sens (tous les blocs sémantiques), elles se trouvent dans l'entrée, pour éviter toute redondance. Ce principe est valable pour les blocs « équivalent » aussi, tout comme pour les équivalents mêmes et leurs traductions.

* * *

Il faut dire encore quelques mots sur « l'arrière-plan » informatique du futur dictionnaire : car apparemment, les possibilités offertes par le logiciel déterminent trop les principes de rédaction (surtout en ce qui concerne la structure des articles). Pourtant, il conviendrait plutôt de dire qu'entre les principes et le système de codage, il y a un rapport de réciprocité, puisque, d'une part, ce système suit de très près les principes de rédaction élaborés en automne 1991, et que, d'autre part, la rigueur informatique du

codage a souvent joué un rôle correcteur dans la constitution définitive de notre « Protocole de rédaction ».

Nous nous servons donc du logiciel « Writer Station » qui est, dans sa conception, une réalisation simultanée de deux choses : d'un traitement de texte et d'une base de données. C'est un logiciel **SGML (Standard Generalized Markup Language)**, c'est-à-dire un traitement de texte qui, à la différence des autres traitements de texte traditionnels, nous permet de :

1) produire des textes (et en général des documents) structurés selon des principes préalablement conçus ;

2) d'utiliser ces documents ainsi structurés dans des systèmes informatiques parfois très différents, et

3) de profiter du système de codage (en tant que base de données) en faveur de certaines recherches de linguistique générale ; à l'aide d'un programme spécial, on peut faire ainsi des recherches complexes dans le corpus organisé : recherches qui relèvent tantôt du domaine de la statistique, tantôt de celui de la théorie des ensembles, etc.

Le nouveau dictionnaire paraîtra donc aussi bien en version électronique qu'en version papier.

Il peut y avoir des différences structurelles assez grandes entre les différents dictionnaires électroniques. La toute première distinction qu'il convient de faire, c'est celle qui existe entre les dictionnaires qu'on peut lire à l'aide d'ordinateurs (*machine-readable dictionaries*)⁴, et ceux qui sont structurés en tant que bases de données (*computerized dictionaries*) et qui nous permettent de faire des recherches complexes et simultanées des données, des ensembles arbitrairement définis et des références les plus diverses. C'est ce deuxième type de dictionnaires électroniques que le nouveau dictionnaire français-hongrois représentera.

4. Júlia PAJZS, *Számítógép és lexikográfia*, in : *Linguistica, Series A, Studia et Dissertationes*, 4, p. 9. MTA Nyelvtudományi Intézet, 1990.

3. Exemple de codage : *texte non formaté (détail)*

<ART><ENT><VDT>connaître </VDT><MOR>(57) </MOR></ENT><BGR>
 <CGR>v tr </CGR><BLS> <EXP>~ qn/qc </EXP><TRD>(meg)ismer vkit/vmit;
 </TRD><EXP>~ qn de vue/nom/réputation </EXP><TRD>látásból/névről/hírből
 ismer vkit; </TRD><EXP>apprendre à ~ qn </EXP><TRD>lassan megismer vkit;
 </TRD><EXP>je l'ai connu tout petit </EXP><TRD>egész kicsi volt, amikor
 megismertem; </TRD><EXP>connaissez-vous ce pays? </EXP><TRD>ismeri ezt az
 országot?/járt már ebben az országban? </TRD><LFG><EXP># il ne connaît pas la
 pitié </EXP><TRD>nem ismer irgalmat; </TRD><EXP>sa patience ne connaît pas de
 bornes </EXP><TRD>türelme nem ismer határokat </TRD></LFG></BLS><BLS>
 <IDS>(date/nom/adresse) </IDS><BEQ><EQV>tud: </EQV></BEQ><EXP>je ne
 connais pas son nom/adresse </EXP><TRD>nem tudom a nevét/címét </TRD>
 ><LFG><EXP># il connaît ce texte par cœur </EXP><TRD>kívülről tudja ezt a
 szöveget </TRD></LFG></BLS><BLS> <IDS>(endroit/résultat/décision)
 </IDS><BEQ><EQV>tud/ismer: </EQV></BEQ> <EXP>connaissez-vous un bon
 restaurant près d'ici? </EXP><TRD>ismertek/tudtok egy jó vendéglőt a közelben?;
 </TRD><EXP>q

4. Exemple de codage : *texte formaté*

connaître (57) I v tr 1 ~ QN/QC (meg)ismer vkit/vmit: ~ qn de vue/nom/réputation
 látásból/névről/hírből ismer vkit; **apprendre à ~ qn** lassan megismer vkit; **je l'ai connu**
tout petit egész kicsi volt, amikor megismertem; **connaissez-vous ce pays?** ismeri ezt
 az országot?/járt már ebben az országban? ■ **il ne connaît pas la pitié** nem ismer
 irgalmat; **sa patience ne connaît pas de bornes** türelme nem ismer határokat 2
 (date/nom/adresse) tud: **je ne connais pas son nom/adresse** nem tudom a nevét/címét
 ■ **il connaît ce texte par cœur** kívülről tudja ezt a szöveget 3 (endroit/résultat/décision)
 tud/ismer: **connaissez-vous un bon restaurant près d'ici?** ismertek/tudtok egy jó
 vendéglőt a közelben?; **quand est-ce qu'on pourra ~ les résultats?** mikor lehet
 megtudni az eredményeket? 4 (nouvelle) hall; tud: **tu connais la nouvelle?** hallottad/tu-
 dod, mi újság?; **tu connais la dernière?** tudod, mi a legújabb hír? 5 (être compétent)
 tud/ismer/ért vmit; ért vmihez; jártas vmiben: **il connaît son métier** tudja/ismeri/érti a
 szakmáját; **il connaît son affaire** érti a dolgát; **il connaît plusieurs langues étrangères**

több idegen nyelvet is ért/tud/ ismer; több idegen nyelven is tud/ért ■ ~PRON INDÉF À QC
il ne connaît pas grand'chose à l'informatique nem nagyon/sokat ért a számítástechnikához; **il n'y connaît rien** semmit sem ért hozzá 6 (*avoir*) (vmilyen nehéz helyzetet) átél; vmilyen sorsra jut: ~ **des moments difficiles** nehéz pillanatokot él át; **ils ont connu des temps meilleurs** láttak már jobb napokat is; **le pays connaît une crise économique grave** az ország súlyos gazdasági válsággal küzd; ~ **un sort tragique** tragikus sorsra jut 7 (*succès*) (hírnévre/ismertségre) szert tesz: **ce roman a connu un succès formidable** ennek a regénynek óriási sikere volt 8 (*qualités/circonstances*)~ QC À QN tud vmit/vmiről; tudomása van vmiről: **je ne lui connaissais pas ces talents** nem is tudtam ezekről a képességeiről; **je ne lui connais pas de défauts/d'ennemis** nem tudok arról, hogy hibái/ellenségei volnának ■ **tu ne connais pas ta chance!** nem is tudod, milyen szerencsés vagy; **il ne connaît pas sa force** nincs tudatában annak, hogy milyen erős 9 **faire** ~ (*idée/décision; sentiment*) (*elképzelést/döntést*) ismertet/közöl vkivel; vkinek a tudomására hoz; (érzelmet) kifejezésre juttat 10 (*pièce/livre*) (*színdarab/könyv*) **faire** ~ q ismertté tesz vkit II v *pron* I ismeri saját magát: **connais-toi toi-même** ismerd meg önmagadat 2 **ils se connaissent depuis longtemps** régóta ismerik egymást 3 **se faire** ~ megismerteti magát; ismertségre/hírnévre tesz szert 4 S'Y ~ ■ **là-dedans, il s'y connaît** ehhez aztán ért; **elle s'y connaît en électricité** ért a villanyszereléshez

5. Sont prêts les fichiers suivants (*état de novembre, 1994*)

(D 1) D-DEB	A 1 (A)	E 1 (E)	C 1
EQV 316	EQV 270	EQV 430	EQV 422
TRD 453	TRD 528	TRD 373	TRD 330
769/185 VDT	798/150 VDT	803/192 VDT	752/248
(D 2) DEC	A 2 (ABRUT)	E 2 (ÉCOS)	C 2
EQV 467	EQV 288	EQV 501	EQV 500
TRD 340	TRD 518	TRD 309	TRD 303
807/247 VDT	806/156 VDT	810/221 VDT	803/287
(D 3) DED-DEJ	A 3 (ACCOUD)	E 3 (ÉL)	C 3
EQV 344	EQV 285	EQV 489	EQV 450
TRD 370	TRD 514	TRD 252	TRD 359
714/206 VDT	799/140 VDT	741/207 VDT	809/239

(D 4) DEL-DEM	A 4 (ACTU)	E 4 (EMBROUS)	C 4
EQV 454	EQV 323	EQV 481	EQV 398
TRD 447	TRD 485	TRD 347	TRD 402
901/240 VDT	808/198 VDT	828/228 VDT	800/199

(D 5) DEN-DEQ	A 5 (AFFAL)	E 5 (ENCHAN)	C 5
EQV 428	EQV 373	EQV 470	EQV 304
TRD 397	TRD 419	TRD 337	TRD 501
825/234 VDT	792/199 VDT	807/167 VDT	805/152

(D 6) DER-DESI	A 6 (AGR)	E 6 (ENGIN)	C 6
EQV 447	EQV 314	EQV 450	EQV 291
TRD 399	TRD 520	TRD 297	TRD 530
846/228 VDT	834/181 VDT	747/185 VDT	821/163

(D 7) DESO-DEU	A 7 (ALGER)	E 7 (ENTON)
EQV 458	EQV 327	EQV 504
TRD 437	TRD 484	TRD 364
895/188 VDT	811/168 VDT	868/210 VDT

(D 8) DEV-DIM	A 8 (AMAT)	E 8 (ÉPINETTE)
EQV 449	EQV 389	EQV 530
TRD 371	TRD 442	TRD 429
820/235 VDT	832/223 VDT	959/299 VDT

(D 9) DIN-DISP
EQV 378
TRD 375
753/166 VDT

(D 10) DISQ-DOP
EQV 525
TRD 445
970/236 VDT

(D 11) DOR-DYS
EQV 420
TRD 458
878/226 VDT

***Total** : 25 % dans 1 an et demi = 27.245 unités/6.841 vedettes en 33 fichiers. D'après la lettre D (2.400 vedettes au lieu des 2.750 du projet 1992/1994), on peut s'attendre à 100.000 unités et à 33.000 vedettes pour le manuscrit total.*