

# COMPUTATIONAL LEXICOGRAPHY

from traditional dictionaries to automated lexicon

Isabelle Warnesson  
IBM-France Scientific Center

## Resúmen

For twenty years, computerized processing in Linguistics has been leading to the creation of a new research field: Computational Linguistics.

Lexicon, syntax and semantics have been the subject of a large research effort. Within the wider context of Language Industries, dictionaries play a fundamental role as obligatory component of any system: speech understanding, speech recognition, text processing and advanced office systems, computerized publishing, text generation systems, natural language interfaces, computer aided translation,... use information stored in dictionaries dedicated to the specific needs of these applications. It is now possible to consider Computational Lexicography as a research and development field by itself.

A dictionary is a human artifact: thus it is subject to errors and inconsistencies. Various mathematical techniques can be used to solve these problems, the computer being the ideal tool.

It is fundamental to be able to:

- improve the ready-made dictionary contents.
- modelize the various lexicographical information.
- highlight the structural problems of these basic dictionaries
- check, disambiguate and correct the contents in order to use the dictionaries with a computer.
- obtain easily interpretable results to update dictionaries.

Some processing on monolingual and bilingual dictionaries will be presented. They have been realised using a new mathematical method called Quadri-decomposition, recently developed at the IBM-France Paris Scientific Center.

COMPUTATIONAL LEXICOGRAPHY

From traditional dictionaries

to

automated lexicons

Isabelle Warnesson

Centre Scientifique IBM-France

3-5 Place Vendôme

75001 Paris France

RELATIONAL AGGREGATION  
 DICTIONARY / DICTIONARIES

NON-LANGUAGE DEPENDENT APPROACH



	M <sub>a</sub>	M <sub>b</sub>	M <sub>c</sub>	M <sub>i</sub>	M <sub>j</sub>	...	M <sub>n</sub>
a					.		
b					.		
c					.		
i	...	...	...	...	<b>C<sub>ij</sub></b>		
j							
...							
n							

SYNONYMS

C<sub>ij</sub> = 1 if M<sub>j</sub> is given as synonym with M<sub>i</sub>

C<sub>ij</sub> = 0 if not.



DEFINITIONS



D<sub>ij</sub> = 1 if M<sub>j</sub> is in the definition of M<sub>i</sub>

D<sub>ij</sub> = 0 if not



	M <sub>a</sub>	M <sub>b</sub>	M <sub>c</sub>	M <sub>d</sub>	M <sub>e</sub>	M <sub>f</sub>	M <sub>j</sub>	...	M <sub>n</sub>
M <sub>g</sub>							.		
M <sub>h</sub>							.		
M <sub>i</sub>	...	...	...	...	...	...	<b>D<sub>ij</sub></b>		
M <sub>n</sub>									



	M <sub>a</sub>	M <sub>b</sub>	M <sub>c</sub>	M <sub>d</sub>	M <sub>e</sub>	M <sub>j</sub>	...	M <sub>o</sub>	M <sub>p</sub>
M <sub>f</sub>						.			
M <sub>g</sub>						.			
M <sub>h</sub>						.			
M <sub>k</sub>						.			
M <sub>i</sub>	...	...	...	...	...	<b>T<sub>ij</sub></b>			
...									
M <sub>q</sub>									

T<sub>ij</sub> = 1 if M<sub>j</sub> is given as translation M<sub>i</sub>

T<sub>ij</sub> = 0 if not

TRANSLATION



# DICTIONARY OF SYNONYMS

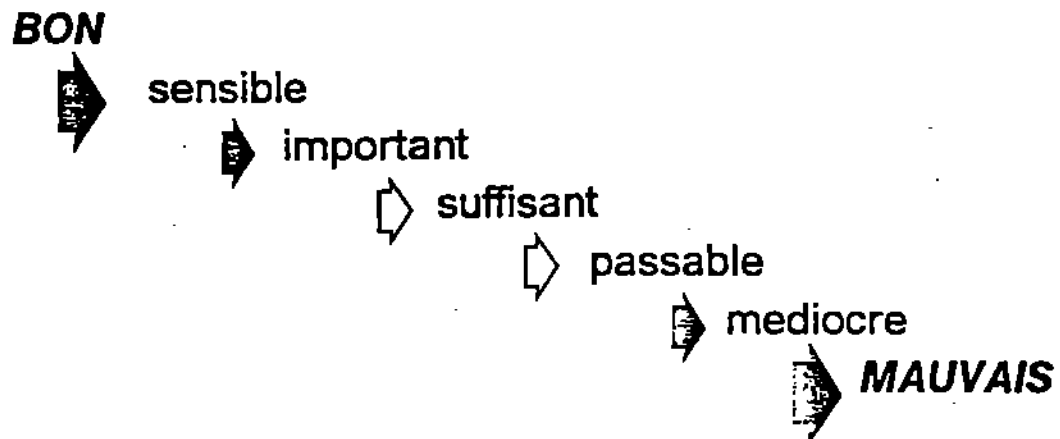
optimization : a need

## ◆ • *in Advanced Text Processing Systems*

- ◆ small clusters
- ◆ semic separate valuation
- ◆ hierarchical thesaurus

## ◆ • *in Traditional Dictionaries*

- ◆ extensive clusters
- ◆ mixing of meanings
- ◆ no hierarchical organization



**chaland**:n

client.

**chalandise**:n

clientèle.

**chalet d'aisance**:n

cabinets.

**chalet de nécessité**:n

cabinets.

**chaleur**:n

<pr>

cuicule, étuve, fournaise, souffeur;

<fig>

amour, animation, ardeur, coeur, cordialité, élan, empressement,  
énergie, enthousiasme, entrain, exaltation, excitation,  
ferveur, feu, fièvre, flamme, fougue, frénésie, impétuosité,  
lyrisme, passion, passion, véhémence, verve, vigueur, vivacité,  
zèle.

**chaîner**:v

<vx>

intéresser.

**chamade**:n

<vx>

appel.

**chamaillieur**:n

disputailleur, querelleur.

**chamaillie**:n

<vx>

dispute, échauffourée.

**chamarré**:j

bariolé.

**chamboulement**:n

chambardement (fam).

**champ de foire**:n

foirail, marché.

**champ de repos**:n

cimetière.

**champ des morts**:n

<lit>

champ du repos, cimetière.

**champignonner**:v

<fam>

proliférer, pulluler.

**championnat**:n

compétition, coupe.

**changement**:n

altération, conversion, évolution, métamorphose, modification, mue,  
imitation, refonte, réformation, réforme, remaniement,  
transformation, variation;

bouleversement, renversement, retournement, revirement, révolution;  
alternance, échange, remplacement, substitution, troc;

<fig>

innovation, mouvement, nouveauté, variété;

remplacement, renouvellement, rénovation.

**changer de crémerie**:v

<arg>

déménager, mettre les bouts (arg); mettre les voiles (arg).

**changer**:vp

évoluer, se convertir, se modifier, se retourner, se transformer,  
tourner bride, varier, virer.

beau1:j

< qqun >

- gracieux, joli, mignon, ravissant, séduisant;
- magnifique, merveilleux, splendide, sublime, superbe;
- accompli, achevé, brillant, consommé, cultivé, éminent, fameux, formidable, grand, haut, magistral, supérieur;
- admirable, digne, estimable, généreux, honn + te, honorable, juste, magnanime, noble, vertueux;
- émouvant, touchant;
- agréable, charmant, gentil, plaisant, séduisant;

< qqch >

- émouvant, touchant;
- adroit, astucieux, habile;
- chic, choisi, élégant, sélect;
- gracieux, harmonieux, joli, mignon;
- esthétique, sculptural;
- magnifique, merveilleux, ravissant, splendide, sublime, superbe;
- fameux, formidable, grand, haut, magistral, parfait, supérieur;
- avantageux, fructueux, lucratif;
- florissant, prospère, riche;
- considérable, énorme, fort, gros, important, imposant;
- intéressant, passionnant;
- adéquat, approprié, bon, heureux;
- agréable, charmant, plaisant, séduisant.



beau01:j

< qqun06 >

gracieux(00), joli(00), mignon(00), ravissant(00), séduisant(00).

beau02:j

< qqun06 >

magnifique(00), merveilleux(00), splendide(00), sublime(00), superbe(00).

beau03:j

< qqun06 >

accompli(00), achevé(00), brillant(00), consommé(00), cultivé(00), éminent(00), fameux(00), formidable(00), grand(00), haut(00), magistral(00), supérieur(00).

beau04:j

< qqun06 >

admirable(00), digne(00), estimable(00), généreux(00), honn + te(00), honorable(00), juste(00), magnanime(00), noble(00), vertueux(00).

beau05:j

< qqun06 >

émouvant(00), touchant(00).

beau06:j

< qqun06 >

agréable(00), charmant(00), gentil(00), plaisant(00), séduisant(00).

beau07:j

< qqch13 >

émouvant(00), touchant(00).

beau08:j

< qqch13 >

adroit(00), astucieux(00), habile(00).

beau09:j

< qqch13 >

chic(00), choisi(00), élégant(00), sélect(00).

beau10:j

< qqch13 >

gracieux(00), harmonieux(00), joli(00), mignon(00).

beau11:j

< qqch13 >

esthétique(00), sculptural(00).

beau12:j

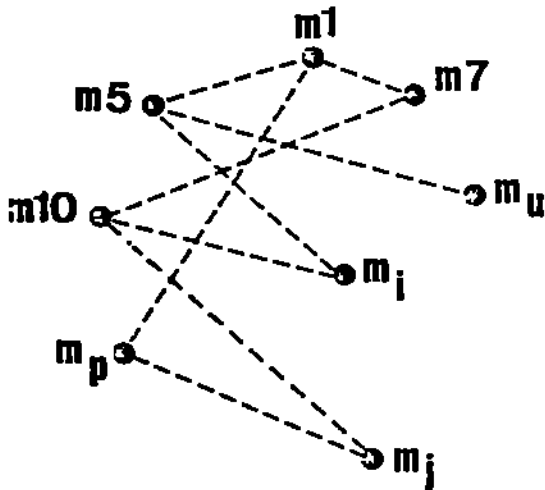
< qqch13 >

magnifique(00), merveilleux(00), ravissant(00), splendide(00), sublime(00), superbe(00).

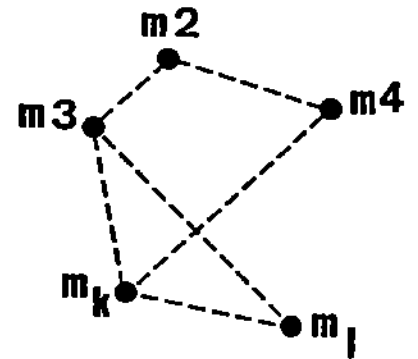
beau13:i

# SYNONYMICAL CONNECTED COMPONENTS

*Decomposition of the problem*

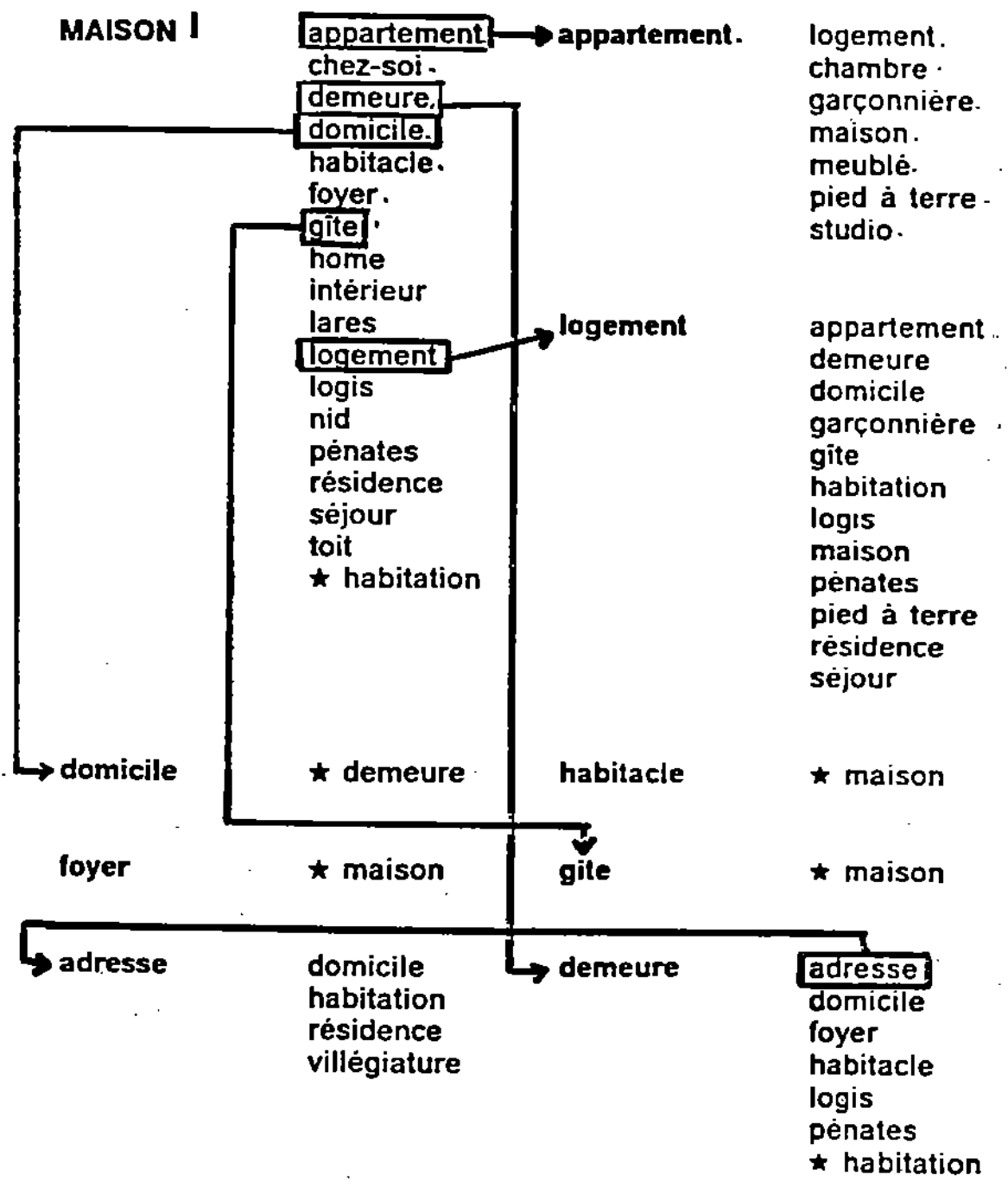


	m1	m5	m7	m10	mi	mi	mu	mr
m1	1	1	1	0	0	0	0	1
m5	1	1	1	0	0	1	0	1
m7	1	0	1	1	0	0	1	0
m10	0	0	1	1	1	1	0	0
mi	0	1	0	1	1	1	0	0
mj	0	0	0	1	1	1	0	1
mu	0	1	1	0	0	0	1	0
mr	1	0	0	0	0	1	0	1



	m2	m3	m4	mk	ml
m2	1	1	1	0	0
m3	1	1	0	1	1
m4	1	0	1	1	0
mk	0	1	1	1	1
ml	0	1	0	1	1

# SEARCH OF CONNECTED COMPONENTS





# RAW DATA MATRIX

	Maison	Appartement	Chez-soi	Demeure	Domicile	Habitacle	Foyer	Gîte	Home	Intérieur	Lares	Nid	Pénates	Résidence	Séjour	Toit	Chambre	Garçonnière	Habitat	Logement	Pied-à-terre	Meublé	Studio	Adresse	Logis	Abri	Refuge	Siège	Endroit	Retraite	Village	Maison	Relais	Hotel	
Maison	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
Appart.	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	
Chez-soi	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
Demeure	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
Domicile	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
Habitac.	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
Foyer	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
Gîte	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
Home	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Intérieur	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
Lares	1	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	
Nid	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
Pénates	1	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	
Résidence	1	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	
Séjour	1	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	
Toit	1	1	1	1	1	0	0	1	1	0	0	1	0	1	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	
Chambre	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Garçon.	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	
Habitat.	1	1	1	1	1	0	0	1	1	0	0	1	0	1	1	0	1	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	
Logement	1	1	0	1	1	0	0	1	0	0	0	0	1	1	1	0	0	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	
Pied-à-t.	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	
Meublé	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
Studio	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
Adresse	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	
Logis	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Abri	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	
Refuge	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	
Siège	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Endroit	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Retraite	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	
Village	1	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	
Maison	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	
Relais	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Hotel	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

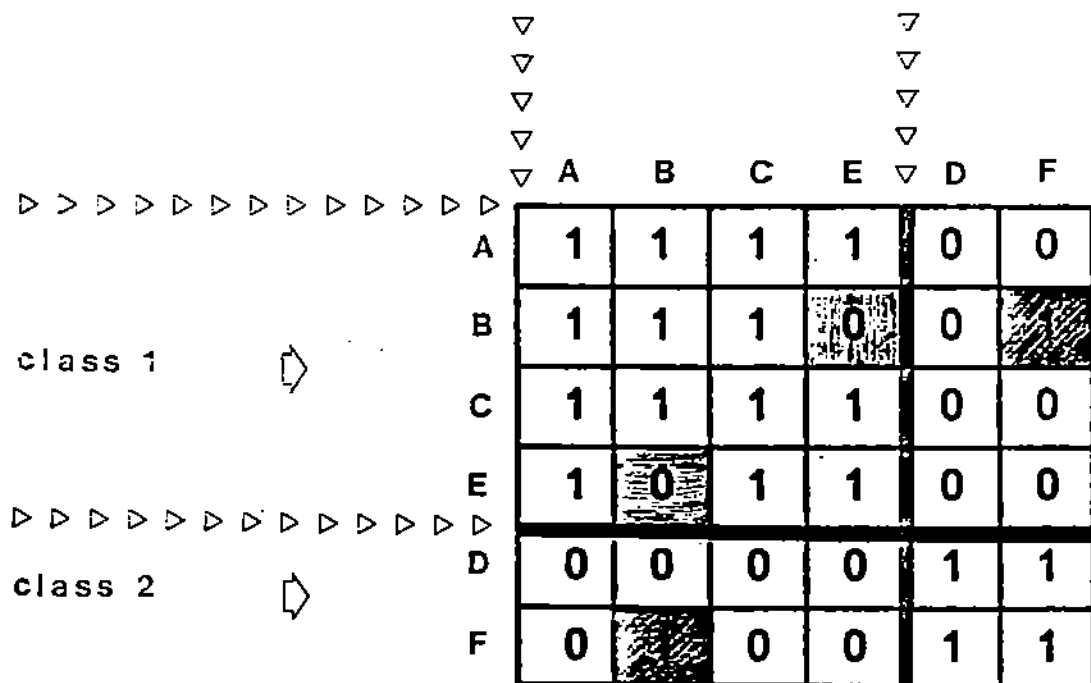
# EXAMPLE

**C**

**Y**

	A	B	C	D	E	F
A	1	1	1	0	1	0
B	1	1	1	0	0	1
C	1	1	1	0	1	0
D	0	0	0	1	0	1
E	1	0	1	0	1	0
F	0	1	0	1	0	1

	A	B	C	D	E	F
A	1	1	1	0	1	0
B	1	1	1	0	1	0
C	1	1	1	0	1	0
D	0	0	0	1	0	1
E	1	1	1	0	1	0
F	0	0	0	1	0	1



C matrix permuted according to Y

- A1** Internal adjustments
- A2** External adjustments
- E1** Internal errors
- E2** External errors

## THE MODEL

⇒ ★ Max (A1 + A2)

$$\text{Max} \sum_{i=1}^n \sum_{j=1}^n [C_{ij}Y_{ij} + \bar{C}_{ij}\bar{Y}_{ij}]$$

●  $Y$  being a partition

$$\begin{array}{l} - \text{Reflexivity} \\ - \text{Symmetry} \\ - \text{Transitivity} \end{array} \left\{ \begin{array}{l} Y_{ii} = 1 \\ Y_{ij} = Y_{ji} \\ Y_{ij} + Y_{jk} - Y_{ik} \leq 1 \end{array} \right.$$

or

⇒ ★ Min (E1 + E2)

$$\text{Min} \sum_{i=1}^n \sum_{j=1}^n [\bar{C}_{ij}Y_{ij} + C_{ij}\bar{Y}_{ij}]$$

●  $Y$  being a partition

$$\begin{array}{l} - \text{Reflexivity} \\ - \text{Symmetry} \\ - \text{Transitivity} \end{array} \left\{ \begin{array}{l} Y_{ii} = 1 \\ Y_{ij} = Y_{ji} \\ Y_{ij} + Y_{jk} - Y_{ik} \leq 1 \end{array} \right.$$

# C TABLEAU PERMUTED ACCORDING TO THE OPTIMAL PARTITION Y

	MAISON	CHEZ SOI	HABITACLE	FOYER	GITE	INTERIEUR	NID	TOIT	LOGIS	DEMEURE	DOMICILE	HABITATION	RESIDENCE	SEJOUR	ADRESSE	APARTEMENT	CARTEMIERE	LOGEMENT	PIED-A-TERRA	STUDIO	ABRI	REFUGE	RETRAITE	ASILE	ENDROIT	LIEU	HOME	LARAS	PANATES	CHAMBRE	MEUBLE	SIÈGE	VILLÉGIATURE	HALTE	RELAIS	HOTEL						
HABITACLE																																										
FOYER																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										
LOGIS																																										
TOIT																																										

# RESULTING OPTIMAL PARTITION

● Class 1	➔	maison nid toit habitation logis intérieur habitacle gîte chez-soi foyer	● Class 2	➔	pension palace meublé garni auberge hotel
			● Class 3	➔	studio pied-à-terre logement garçonnière appartement
● Class 4	➔	emplacement endroit site position lieu situation	● Class 5	➔	résidence siège
● Class 7	➔	demeure domicile	● Class 6	➔	séjour villégiature
○ Class 9	➔	adresse	○ Class 8	➔	lares
○ Class 11	➔	pénates	○ Class 10	➔	place

# CLUSTERS REPRESENTATIVES

\*\*\*\* ANALYSE PAR MOT ( EN % ) \*\*\*\*

## VENTILATION PAR CLASSE

CLASSE	LIENS		
	CREES+DETRUITS	CREES	DETRUITS
<b>CLASSE 1</b>			
•• 1 PENSION	0.00	0.00	0.00
•• 2 PALACE	0.00	0.00	0.00
•• 3 MEUBLE	0.00	0.00	0.00
•• 4 GARNI	0.00	0.00	0.00
•• 5 AUDERGE	0.00	0.00	0.00
➔ 6 HOTEL	2.78	0.00	3.23
<b>CLASSE 2</b>			
•• 26 GARCONNI	0.00	0.00	0.00
•• 8 PIED-A-T	2.78	25.00	0.00
•• 7 STUDIO	5.56	50.00	0.00
➔ 27 APPARTEH	5.56	0.00	6.25
•• 25 LOGEMENT	16.67	25.00	15.63
<b>CLASSE 3</b>			
•• 9 SIEGE	0.00	0.00	0.00
➔ 19 RESIDENC	13.89	0.00	14.29
<b>CLASSE 4</b>			
•• 10 VILLYGIA	0.00	0.00	0.00
➔ 18 SEJOUR	11.11	0.00	11.43
<b>CLASSE 5</b>			
•• 12 SITE	0.00	0.00	0.00
•• 11 SITUATIO	0.00	0.00	0.00
➔ 16 IMPLACEM	2.78	0.00	3.23
•• 13 POSITION	2.78	20.00	0.00
➔ 15 LIEU	2.78	0.00	3.23
•• 17 ENDROIT	8.33	20.00	6.45
<b>CLASSE 6</b>			
14 PLACE	8.33	-1.00	8.33
<b>CLASSE 7</b>			
•• 22 MID	0.00	0.00	0.00
•• 33 CHEZ-SOI	0.00	0.00	0.00
•• 32 CITE	2.78	0.00	3.70
•• 30 INTERIEU	5.56	22.22	0.00
•• 21 TOTI	8.33	33.33	0.00
•• 31 HABITACL	11.11	22.22	7.41
➔ 29 LOGIS	16.67	0.00	22.22
➔ 28 MAISON	16.67	0.00	22.22
•• 24 ...	17.77	...	...

## CONCEPTUAL LEXICON

- specific
- generic
- related

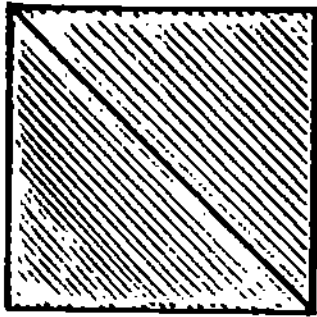
# RELATIONAL ANALYSIS

**Raw Data**

**Processed data**

**Result**

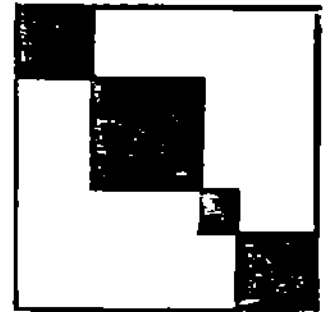
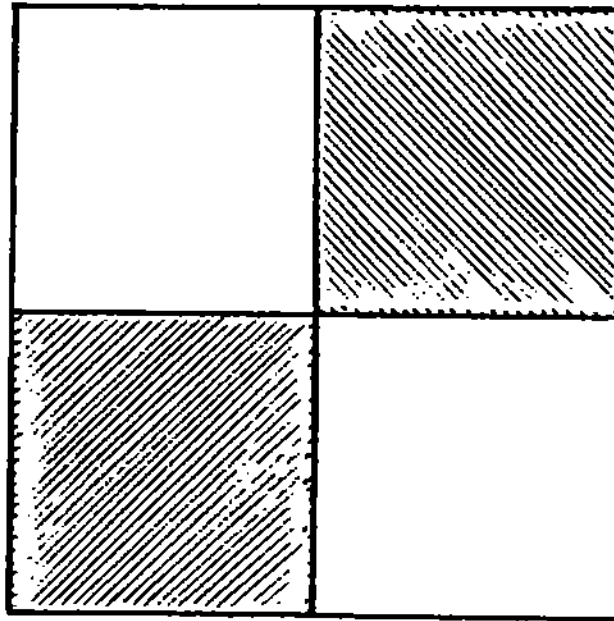
I



SQUARED &  
SYMMETRIC

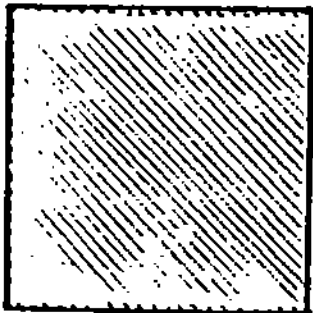
I

I



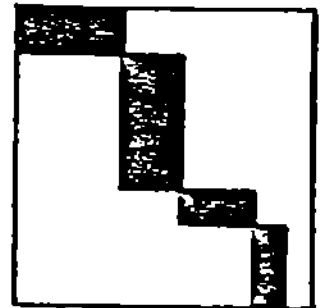
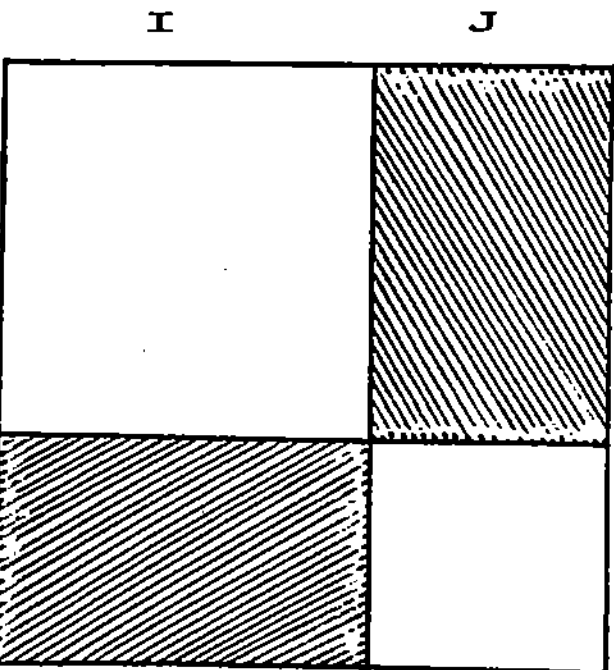
EQUIVALENCE  
RELATION

I



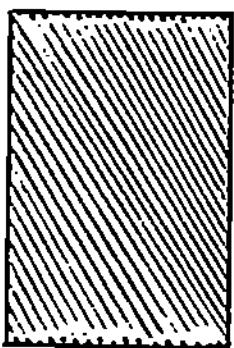
SQUARED &  
NON SYMMETRIC

I



BLOCK  
CORRESPONDENCE

J

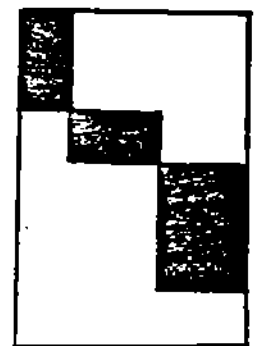


RECTANGULAR

I

J

J



BLOCK  
CORRESPONDENCE

UADRI-DECOMPOSITION

SIMILARITY AGGREGATION

**QUADRIDECOMPOSITION AND DICTIONARIES**

$\delta$  - DICTIONARY OF SYNONYMS

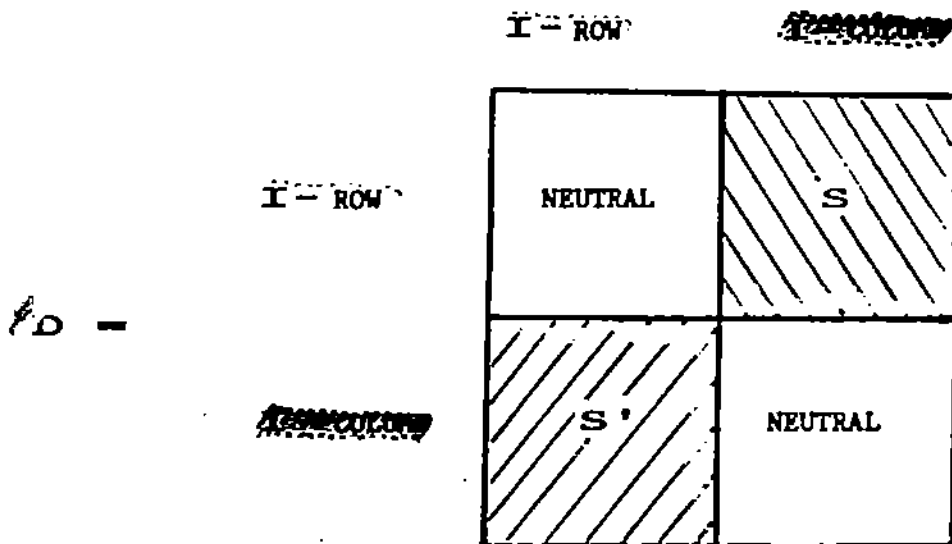
$I$  - { words belongin to the connected compon of an entry from  $\delta$  }

(S) -  $(\delta_{ij})$   $i$  and  $j = 1, \dots, |I|$   
relational matrix of synonymy given by.

$L(i) = \{j \in I : \delta_{ij} = 1\}$  - synonyms with word  $i$   
"ROW-WORD" (calling word)

~~C(i)~~ -  $\{j \in I : \delta_{ji} = 1\}$  - words being synonyms with  $i$   
→ "COLUMN-WORD" (called word)

\* very often:  $L(i) = C(i)$ . It is very interesting to "double" the words in the quadridecomposition matrix.



\* DURING THE PROCESSING, "ROW-WORDS" AND "COLUMN-WORDS" KEEP SEPARATED

\* SOME WORDS WILL BELONG TO TWO DIFFERENT CLUSTERS IN THE OBTAINED SOLUTION:

$I$ -ROW /  $J$ -COLUMN == "BRIDGE-WORDS"

OBTAINED SOLUTION ON  $I$  -

PARTITION WITH IMBEDDED CLUSTERS



## CONDORCET'S SOLUTION

commune  
municipalité

agglomération  
bourg  
bourgade  
hameau  
localité  
village

bled  
coin  
patelin  
pays  
trou

capitale  
centre  
cité  
métropole  
ville

+ additional informations on  
links between clusters

Connected component: "VILLE"

# QUADRI-DECOMPOSITION TECHNIQUE

VILLE	1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
METROPOLE	0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
CITE	1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
CAPITALE	0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
AGGLOMERATION	1 0 1 0 1 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0
CENTRE	1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
VILLAGE	0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 0 0 0
LOCALITE	0 0 0 0 1 0 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0
BOURGADE	0 0 0 0 0 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0
BOURG	0 0 0 0 0 0 0 1 0 1 1 1 1 0 1 0 0 0 0 0 0
TROU	0 0 0 0 0 0 0 1 0 0 1 1 1 1 0 0 1 0 0 1 0 0
PATELIN	0 0 0 0 0 0 0 1 0 0 1 1 1 1 0 0 1 1 0 1 0
BLED	0 0 0 0 0 0 0 1 0 0 0 1 1 1 0 0 0 0 0 0 0
HAMEAU	0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 0 0 0
COMMUNE	0 0 0 0 1 0 1 0 1 1 0 0 0 0 0 1 0 0 1 1 0 0 1
COIN	0 0 0 0 0 0 0 0 1 0 0 1 1 1 0 0 1 0 0 1 0 0
PAYS	0 0 1 0 0 0 0 0 0 1 1 1 1 1 0 1 0 1 0 1 0
MUNICIPALITE	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1

RAW DATA MATRIX

connected component: VILLE

-Classe 1:

CITE-L  
VILLE-C

-Classe 2:

CAPITALE-L, CENTRE-L METROPOLE-L VILLE-L  
CAPITALE-C, CENTRE-C METROPOLE-C CITE-C AGGLOMERATION-C

-Classe 3:

VILLAGE-L BOURGADE-L BOURG-L COMMUNE-L LOCALITE-L AGGLOMERATION-L  
VILLAGE-C BOURGADE-C BOURG-C

-Classe 4:

BLED-L COIN-L PATELIN-L TROU-L PAYS-L  
BLED-C COIN-C PATELIN-C TROU-C

-Classe 5:

HAMEAU-L  
HAMEAU-C LOCALITE-C

-Classe 6:

MUNICIPALITE-L  
MUNICIPALITE-C COMMUNE-C

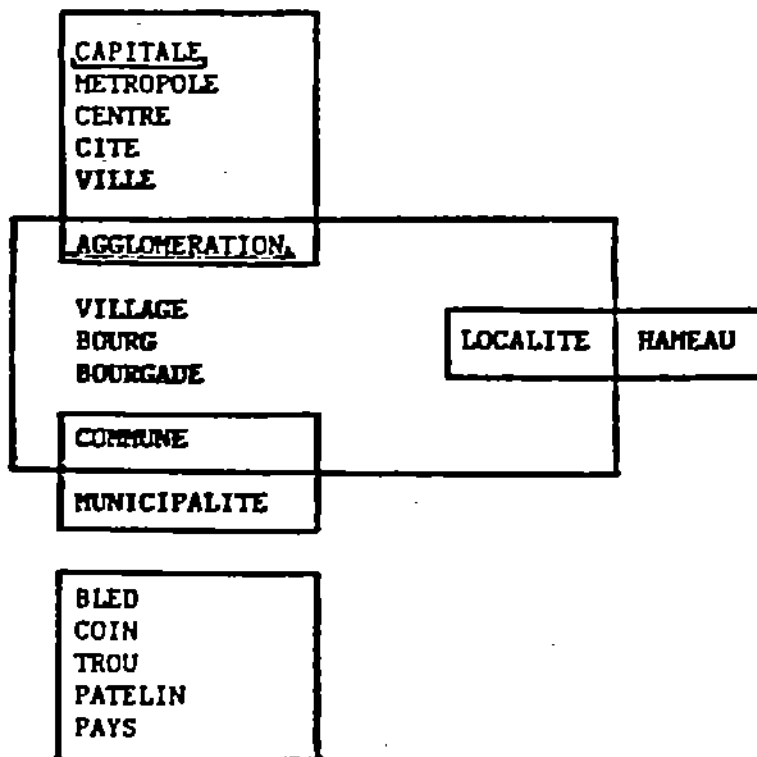
-Classe 7:

PAYS-C

RESULTING PARTITION

	VILLE	AGGLOMERATION	CAPITALE	CENTRE	CITE	METROPOLE	VILLAGE	BOURG	BOURGADE	BLED	COIN	PATELIN	TROU	HAMEAU	LOCALITE	MUNICIPALITE	COMMUNE	PAYS
CITE	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CAPITALE	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
CENTRE	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
METROPOLE	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
VILLE	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
AGGLOMERATION	1	1	0	0	1	0	1	1	1	0	0	0	0	0	1	0	0	0
BOURGADE	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
BOURG	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	0	0
COMMUNE	0	1	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	0
LOCALITE	0	1	0	0	0	0	1	1	1	0	0	1	1	0	1	0	0	0
VILLAGE	0	0	0	0	0	0	1	1	1	1	0	1	1	1	1	1	0	0
BLED	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0
COIN	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	0	0	0
PATELIN	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1
PAYS	0	0	0	0	1	0	1	1	0	1	0	1	1	0	0	1	0	1
TROU	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0
HAMEAU	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0
MUNICIPALITE	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0

DATA PERMUTED ACCORDING TO THE RESULT



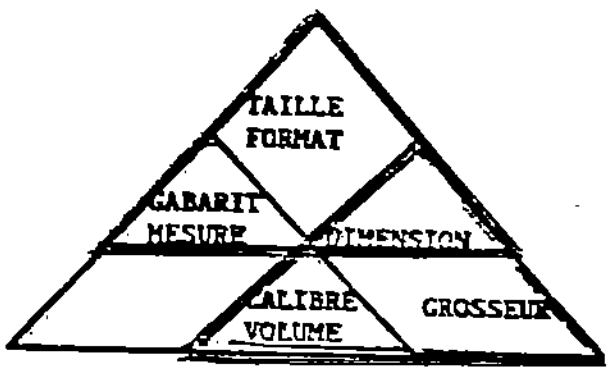
EMBEDDED CLUSTERS PARTITION

AIRE  
TERRAIN

ESPACE

SURFACE      SUPERFICIE  
ETENDUE

GRANDEUR



ZONE  
FIEF  
TERRITOIRE

SECTEUR      PARAGES

COIN  
EMPLACEMENT  
ENDROIT  
LIEU  
PLACE  
POINT

CONTREE

REGION

TERRE  
PAYS

PROVINCE

SITE

SITUATION  
POSITION  
EXPOSITION  
ORIENTATION

SOCLE  
PIEDestal  
SUPPORT

BASE

ASSIETTE  
FONDATION  
SOUBASSEMENT

PIED  
BAS

AXE  
DIRECTION  
LIGNE

FOND

IMBEDDED CLUSTERS PARTITION

CONNECTED COMPONENT: "COIN"

SYNONYMS AND DEFINITIONS

MAISON. : endroit où vivent les gens

FOYER. : lieu où réside la famille

LOGIS : endroit où on loge, où on habite\*

LOGIS : local où on vit avec confort et intimité

LOGIS : lieu où l'on trouve à se loger

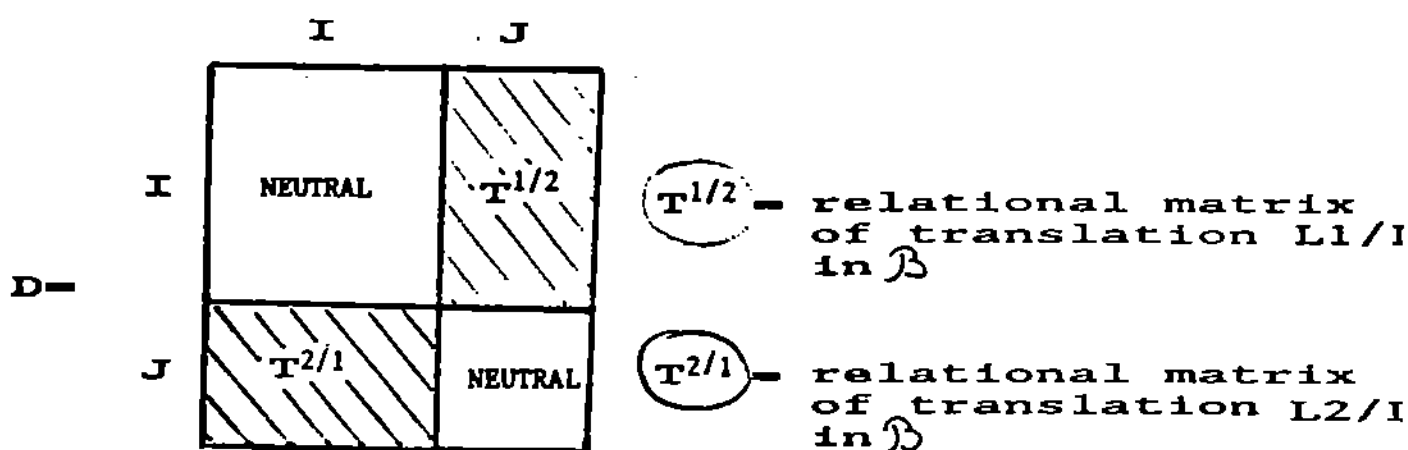
→ ENHANCEMENT OF SYNONYMS CLUSTERS

→ CHECKING THE INTERNAL COHERENCE BY CROSSING  
THE SYNONYMOUS PARADIGMS

→ STUDY OF THE SYNTAGMATIC RELATIONS BETWEEN  
CLUSTERS

B - BILINGUAL DICTIONARY Language 1 (L1) / Language 2 (L2)

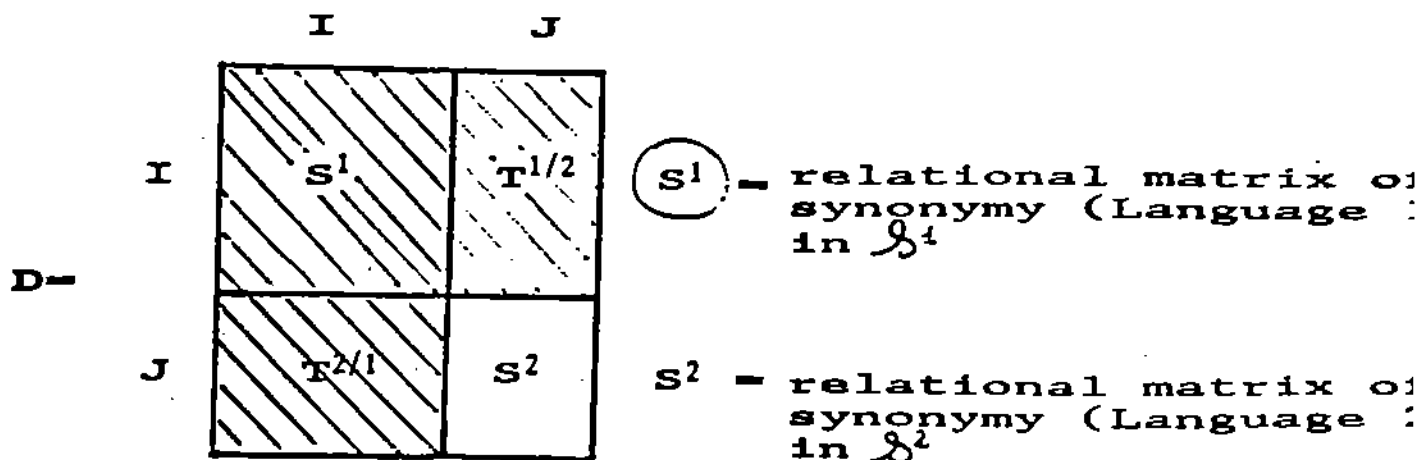
- I - {words in L1 from the connected component of an entry in L1}
- J - {words in L2 from the connected component of an entry in L2}



$T^{1/2}$  &  $T^{2/1}$  are not trivially derived the one from each other

\* \* \*

$\mathcal{B}$ .  $\mathcal{S}^1$  - synonyms (L1) ,  $\mathcal{S}^2$  - synonyms (L2)



calibre CALIBER  
gabarit FORMER

grandeur BULK  
grosueur SIZE  
volume VOLUME

taille EXTENT

mesure MEASURE  
MEASUREMENT

étendue EXPANSE  
dimension DIMENSION

format FORMAT

espace SPACE

emplacement LOCATION  
place SITE  
situation SITUATION

orientation ORIENTATION

aire AREA  
superficie SURFACE  
surface

direction DIRECTION  
WAY

axe LINE  
ligne

coin PLACE  
endroit POINT  
lieu SPOT  
point

parages PART  
QUARTER

fief FIEF

province PROVINCE

région REGION  
territoire TERRITORY  
DISTRICT

contrée COUNTRY  
pays LAND  
terrain  
terre

zone  
secteur ZONE

pedestal PEDESTAL  
socle PLINTH  
SOCLE

fond FOUNDATION  
fondation

bas BASE  
base BOTTOM  
fondement FOOT  
ped

assiette SEATING  
assise

support SUPPORT

## QUADRI-DECOMPOSITION

### BILINGUAL DICTIONARY:

**A** = French / English

**B** = English / French

dimension	
format	
grandeur	FORMAT
mesure	

calibre	
gabarit	
grosueur	FORMER
taille	
volume	

étendue	BULK
	CALIBRE
	DIMENSION
	EXPANSE
	EXTENT
	MEASURE
	MEASUREMENT
	SIZE
SPACE	
VOLUME	

aire	
espace	
superficie	SURFACE
surface	
zone	

coin	
parages	PART
secteur	

terrain	
pays	LAND
terre	

fief	FIEF
------	------

↑ territoire région contrée ↓	AREA
	COUNTRY
	DISTRICT
	PROVINCE
	QUARTER
	REGION
	TERRITORY
	ZONE

assiette	
assise	
fondation	SEATING
fondement	
soubassement	

pedestal	PEDESTAL
socle	SOCLE
support	

bas	BASE
base	BOTTOM
fond	FOOT
fond	FOUNDATION
ped	FLINTH
support	SUPPORT

axe	
direction	DIRECTION
ligne	LINE
orientation	

exposition	EXPOSITION
point	POINT

emplacement	LOCATION
endroit	PLACE
lieu	POSITION
place	SITE
position	SITUATION
site	SPOT
situation	

### QUADRI-DECOMPOSITION

BILINGUAL DICTIONARY &

DICTIONARY OF SYNONYMS

[N] - French synonyms

[A] - translation French / English

M - English synonyms

[B] - translation English / French