

# **Semiometry : the use of words to describe lifestyles and values.**

Ludovic Lebart,  
CNRS – ENST, Paris, France.  
lebart@enst.fr

◆ In most surveys related to marketing research, it is customary to include information about lifestyles and values.

◆ Such information is generally obtained through a set of questions describing the attitudes and opinions towards a list of sentences or statements.

For example (sample of items):

- *I often feel lonely and isolated from others;*
- *I belong to a closely knitted community;*
- *The death penalty should be re-established;*
- *Family is the only place where we feel good; ... etc.*

◆ "Semiometry" is a technique introduced by a writer interested in marketing research, Jean-François Steiner, to tackle this problem in a more general way.

# I. Basic principles and features

The basic idea is to insert in the questionnaire a series of questions consisting uniquely of words

(a list of 210 words is currently used, but some abbreviated lists containing a subset of 80 words could be used as well).

The interviewees must rate these words according to a seven levels scale, the lowest level (mark = 1) concerning a "most disagreeable (or unpleasant) feeling about the word", the highest level (mark = 7) concerning a "most agreeable (or pleasant) feeling" about the word.

|     |               |    |          |          |          |    |    |    |
|-----|---------------|----|----------|----------|----------|----|----|----|
| 122 | La modestie   | -3 | -2       | <b>X</b> | 0        | +1 | +2 | +3 |
| 133 | Mcelleux      | -3 | -2       | -1       | <b>X</b> | +1 | +2 | +3 |
| 124 | La mort       | -3 | <b>X</b> | -1       | 0        | +1 | +2 | +3 |
| 100 | Une muraille  | -3 | -2       | -1       | 0        | +1 | +2 | +3 |
| 085 | Un mystère    | -3 | -2       | -1       | 0        | +1 | +2 | +3 |
| 105 | Nager         | -3 | -2       | -1       | 0        | +1 | +2 | +3 |
| 043 | Une naissance | -3 | -2       | -1       | 0        | +1 | +2 | +3 |
| 025 | Un nid        | -3 | -2       | -1       | 0        | +1 | +2 | +3 |
| 106 | La nudité     | -3 | -2       | -1       | 0        | +1 | +2 | +3 |
| 071 | Obéir         | -3 | -2       | -1       | 0        | +1 | +2 | +3 |
| 173 | L'océan       | -3 | -2       | -1       | 0        | +1 | +2 | +3 |
| 086 | Un orage      | -3 | -2       | -1       | 0        | +1 | +2 | +3 |

**Facsimile of a questionnaire**

The processing of the filled questionnaires (*mainly through Principal Components Analysis*) produces a stable pattern (*up to 8 stable principal axes*).

Very similar patterns are obtained in ten different countries, despite the problems posed by the translation of the list of words.

There is also an "open version" of the semiometry in which the interviewees are asked openly (*without the help of a previously established list*) to mention which words they consider as agreeable or disagreeable. We will deal with this variant later on.

## Questionnaires in 5 languages

### FRENCH

l'absolu

l'acharnement

acheter

admirer

adorer

l'ambition

l'âme

l'amitié

l'angoisse

un animal

un arbre

l'argent

une armure

l'art

### ENGLISH

absolute

persistence

to buy

to admire

to love

ambition

soul

friendship

anguish

animal

tree

silver

armour

art

### GERMAN

absolut

hartnaeckig

kaufen

bewundern

anbeten

der ehrgeiz

die seele

die freundschaft

die angst

ein tier

ein baum

das geld

die ruestung

die kunst

### SPANISH

el absoluto

el empeno

comprar

admirar

adorar

la ambicion

el alma

la amistad

la angustia

un animal

un arbol

el dinero

una armadura

el arte

### ITALIAN

l'assoluto

l'accanimento

comprare

ammirare

adorare

l'ambizione

l'anima

l'amicizia

l'angoscia

un animale

un albero

il denaro

un'armatura

l'arte

## Surveys carried out since 1990

| Country | Year | Sample Size |
|---------|------|-------------|
| France  | 1990 | 2764        |
| France  | 1996 | 2764        |
| Spain   | 1996 | 2983        |
| U.K.    | 1996 | 1849        |
| Italy   | 1996 | 2606        |
| Germany | 1996 | 3065        |
| France  | 1998 | 2764        |
| Greece  | 2000 | 1062        |
| France  | 1999 | 2763        |
| Canada  | 1998 | 1865        |

## **II. Brief presentation of the main results :**

In each country, a correlation matrix can be computed. This square symmetric matrix with 210 rows and 210 columns is in fact a by-product of the Principal Components Analysis of the arrays of measurements (marks or ratings attributed to words).

These sets of correlation coefficients allow us to define a distance between every pair of words.



## Part of a Kohonen map describing the correlations between words

|   |   |  |   |   |  |   |
|---|---|--|---|---|--|---|
| WORK<br>TO_ORDER<br>TO_OBEY<br>SOLDIER<br>SCHOOL<br>RULE<br>LAW<br>HOUSE<br>ECONOMISE<br>DISCIPLINE | SKIN<br>QUESTION<br>MATERIAL<br>INDUSTRY<br>BUILDER | TRADE<br>REASON<br>PRACTICAL<br>EFFORT                 | TO_PRODUCE<br>TO_BUILD<br>ORIGINAL<br>FLEXIBILITY<br>FIRMNESS<br>EFFECTIVE<br>CERTAINTY<br>ACCURACY | TO_MASTER<br>PERFECTION<br>CHALLENGE<br>BRAVE | VICTORY<br>TO_CONQUER<br>POWER<br>HERO<br>GLORY<br>ELITE<br>AMBITION | WEALTH<br>TO_INHERIT<br>REWARD<br>GOLD                              |
| TRADITION<br>MORALS<br>MODESTY<br>MODERATION<br>HOMELAND<br>CEREMONY                                | ABSOLUTE  | SOLID<br>SCIENCE<br>RESEARCHER<br>PERSISTENCE<br>LOGIC |   |   |  | TO_BUY<br>SILVER<br>PROPERTY<br>JEWEL<br>FASHION                    |
| SOUL<br>SACRED<br>PURITY<br>PRIEST<br>GOD<br>FAITH<br>CREATOR<br>BELIEF                             | REFINED<br>PRUDENCE<br>NOBLE                        | SUBLIME<br>ROBUST<br>PEAK<br>INVENTOR                  | VIRILE<br>INFINITY<br>DYNAMIC   | TO_FERTILISE                                  | PRESENT<br>ELEGANCE  | PERFUME   |
| MONARCH<br>MEDITATION<br>HUMBLE   |   | METALLIC<br>HUGE<br>ESCAPE                             | VOLUPTUOUS<br>SPEED   | SENSUAL<br>INTIMATE<br>DESIRE                 | SOFT<br>FEMININE<br>ETERNAL<br>EMOTION<br>CARESS                     | WEDDING<br>SOFTNESS<br>MATERNAL<br>CHILDHOOD<br>BIRTH<br>ATTACHMENT |

# Part of a Kohonen map describing the correlations between words

(Continuation)

|   |   |                                 |  |   |  |   |
|---|---|---------------------------------|--|---|--|---|
| SACRIFICE   | WALL<br>SECRET<br>RIGID<br>KNOT<br>DETACHMENT<br>BORDER<br>AUDACITY<br>ARMOUR | THINNESS<br>CARNAL              | TO_SEDUCE<br>NUDITY<br>MYSTERY                               | TO_DREAM  | TO_ADMIRE<br>PRECIOUS  | TO_LOVE<br>TOGETHER<br>TENDERNESS<br>FAMILY<br>CHARITABLE   |
| TO_INTERROGA<br>TO_FORBID<br>TO_CRITICISE<br>TO_BREAK<br>TO_AGE<br>IMMOBILE   | SHOUT<br>MASK<br>IRONY<br>HUNT<br>CRAFTINESS                                  | UNKNOWN<br>MAZE                 | WANDER<br>MAGIC<br>ISLAND<br>DIFFERENT<br>BLUE<br>ADVENTURER | TO_SWIM<br>GAME<br>ANIMAL   | FREE   | TO_PROTECT<br>TO_LAUGH<br>TO_CARE<br>HUMOUR<br>FRIENDSHIP<br>FORGIVENESS<br>COMFORT   |
| WAR<br>TO_PUNISH<br>TO_BETRAY<br>TO_ATTACK<br>MISTRUST<br>GUN<br>FAULT<br>DOUBT<br>DISORDER<br>DEATH<br>DANGER<br>ANGUISH | STRANGER<br>STORM<br>REBELLION<br>FIRE<br>CUNNING<br>BLACK                    | WILD<br>RED<br>DESERT<br>CHANGE | TO_CLIMB<br>SPACE<br>OCEAN<br>MOUNTAIN<br>MOON               | TREE<br>TO_WRITE<br>THEATRE<br>RIVER<br>POETRY<br>NEST<br>LIGHTNESS<br>GREEN<br>BOOK<br>ART | WATER<br>VOLUNTEER<br>TO_THINK<br>MUSIC<br>FLOWER<br>COUNTRYSIDE | TO_TEACH<br>TO_RECOVER<br>TO_CONSOLE<br>RESPECT<br>POLITENESS<br>PEACE<br>PATIENCE<br>LIVELINESS<br>JUSTICE<br>HONOUR<br>HONEST<br>CONFIDENCE |

◆ In fact, the semantic similarity is not a transitive relationship.

◆ See the following counterexample:

“ **fact, feature, aspect, appearance, illusion**”

.

◆ But large distances between words (i.e.: significant negative correlation) do exist .

◆ They are better taken into account by the classical Principal Component Analysis method.

◆ The most remote group of points appearing then to occupy opposite locations on the first principal axes.

(Visualization axes and principal planes plans)

Axis 1 is a purely methodological axis, it opposes the extreme notes (1,2, and 6, 7) to the medium notes (3, 4, 5).

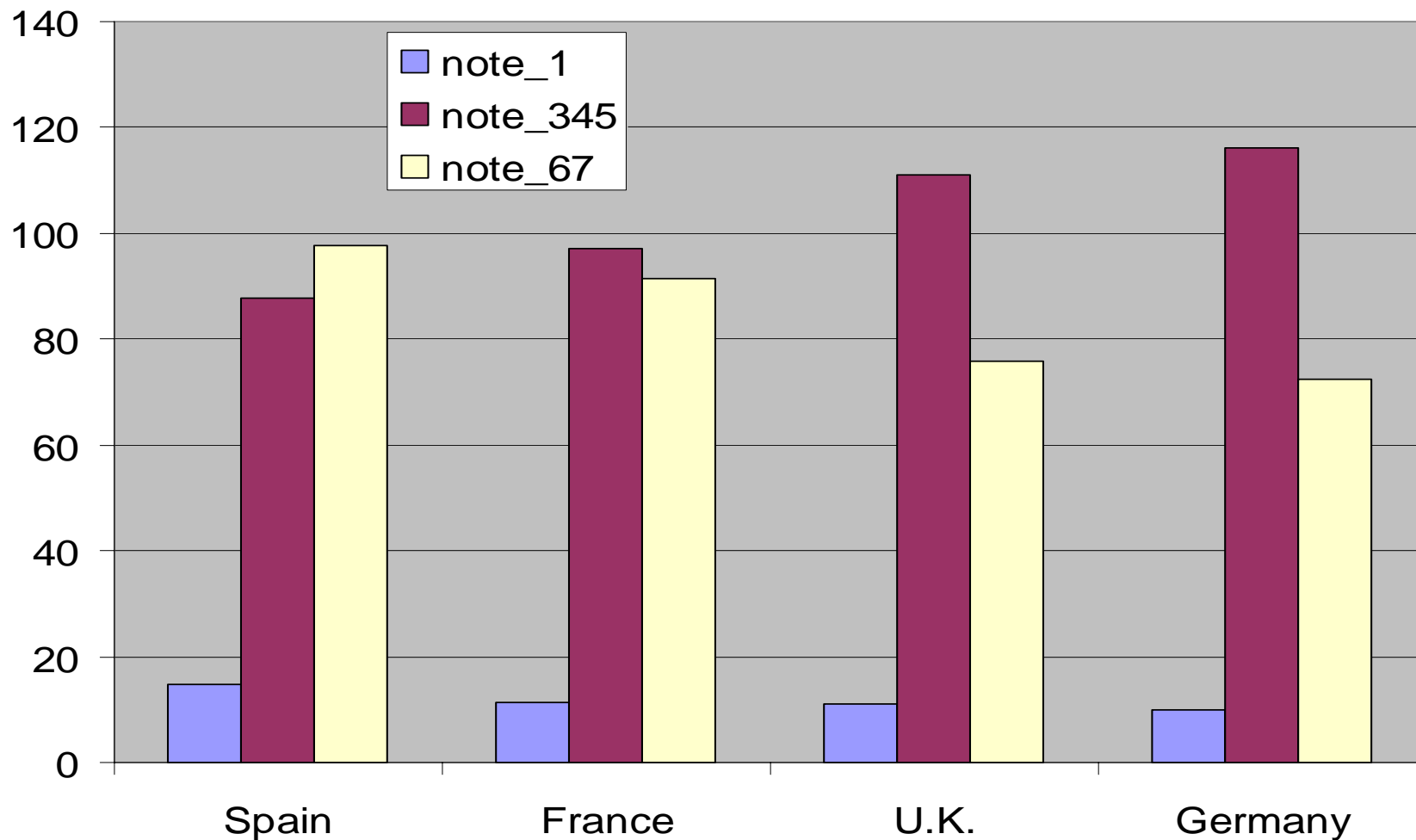
Although it appears in every country, its importance in terms of variance is greater in the north of Europe:

(UK, Germany, where the centre of the scale is largely used)

than in the south of Europe:

(Italy, France, Spain, more inclined to use extreme notes) .

## Differences in profiles of notation in Europe



**Axis 2 can be defined as a *duty-pleasure* axis.**

It is related to the age of the respondent, young people being on the side of the pleasure and elder respondents on the side of the duty, characterised by words such that *law, discipline, God, family*, etc..

However, we encounter a surprising result, this axis also exists if the analysis is performed within a relatively narrow age category, such as the category “35-40”).

**Axis 3 can be defined as a *attachment – detachment* axis.**

It is markedly related to the gender (women being more characterised by the attachment side of the axes, and men by the detachment side, i.e.: words such that: *danger, to\_attack, war, etc.*).

However, another surprising and all the more interesting result is that the same axis appears if the analysis is performed on a sub-sample consisting of one single gender, be it male or female.

However, the order of these axes is not always the same : in particular, there is a permutation of the axes two and three when the analysis concerns countries belonging to the north or to the south of Europe.

The axis *attachment-detachment* comes as a second axis in northern Europe and as a third axis in Spain, Italy and France, countries in which the axis *duty-pleasure* is the second one.



### III Examples of application

These examples aim at showing how Semiometry allows one to characterise a specific subset of the population :

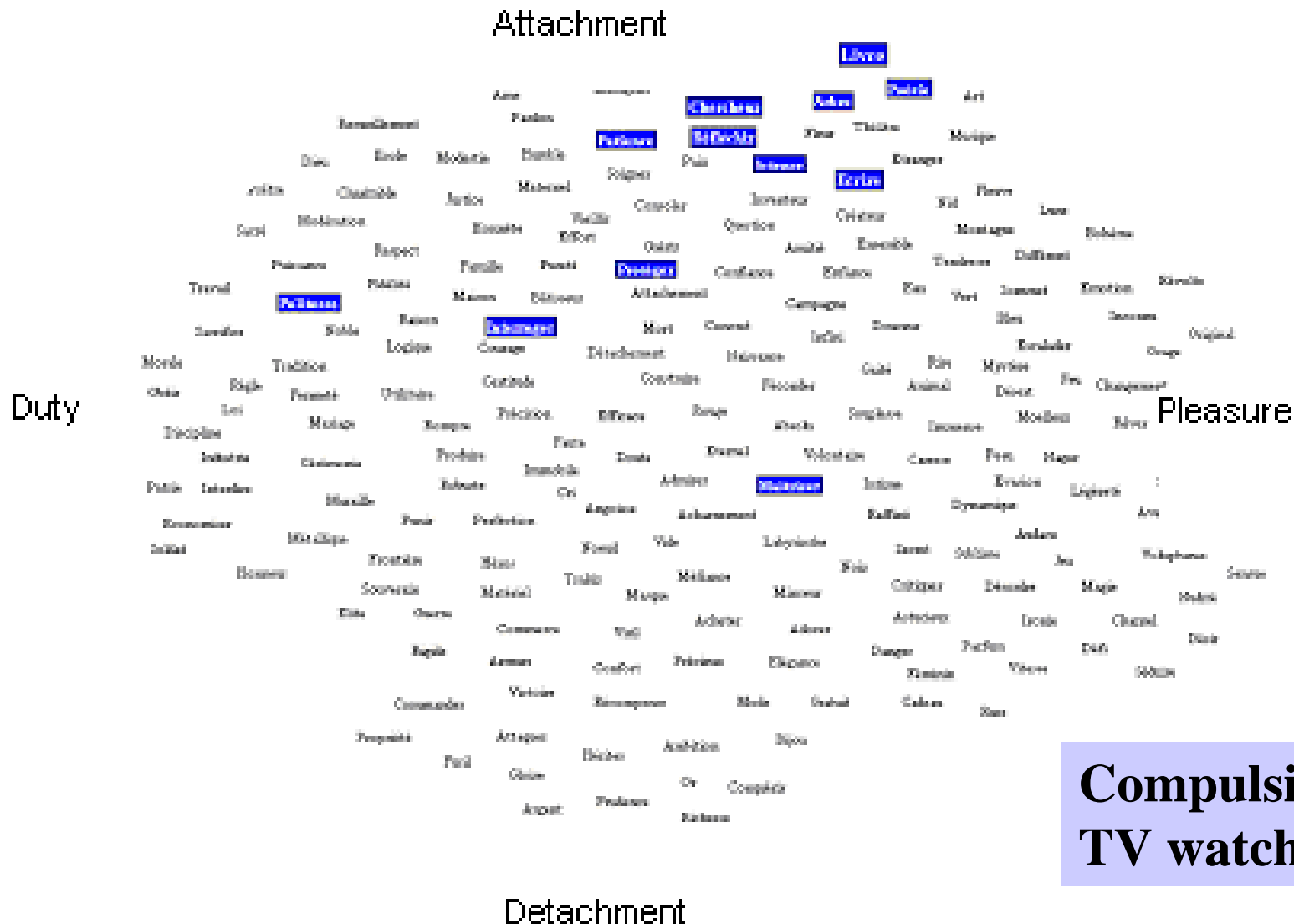
In the first example, the individual having a high (last quartile) TV audience. The second example concerns the evolution of values in France in a relatively short period (from 1996 to 1998).

Of course, most application in marketing research will concern the consumer of a particular product or brand or the user of a particular service... but the corresponding data are not available, being confidential.

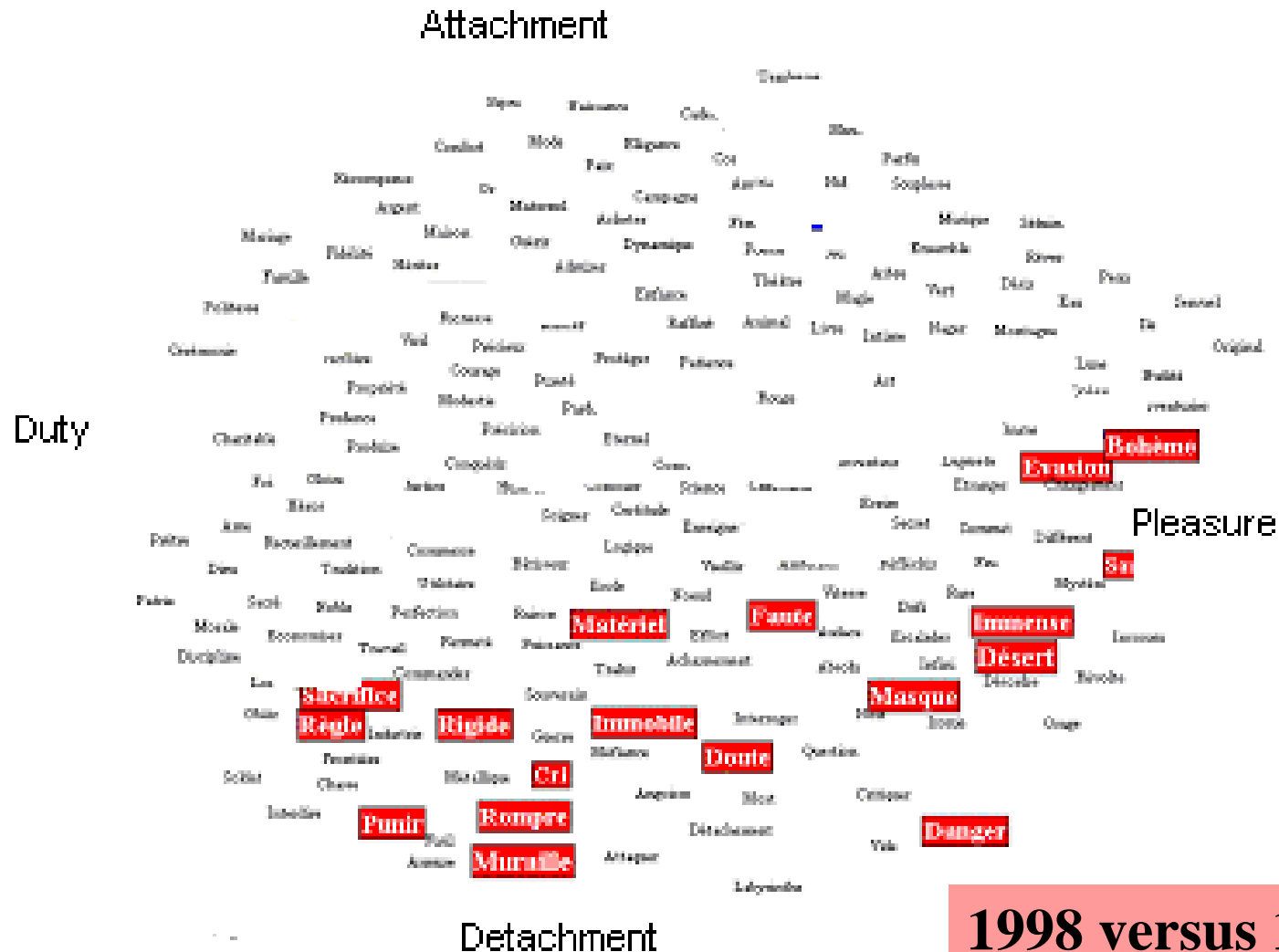
A way of characterising a particular category of respondents is to compare, for each word of the semiometric questionnaire, the average note of this category to the global average of the notes given to the same word.

A simple test of comparison of means allows one to test the significance of the hypothesis of independence (i.e.: the mean in the particular category is equal to the global mean) and then (taking into account the problem of “multiple comparisons”), to select the significant words.

*Words “rejected”: Book, poetry, researcher, to think, science,  
to write, patience, to protect, tree*

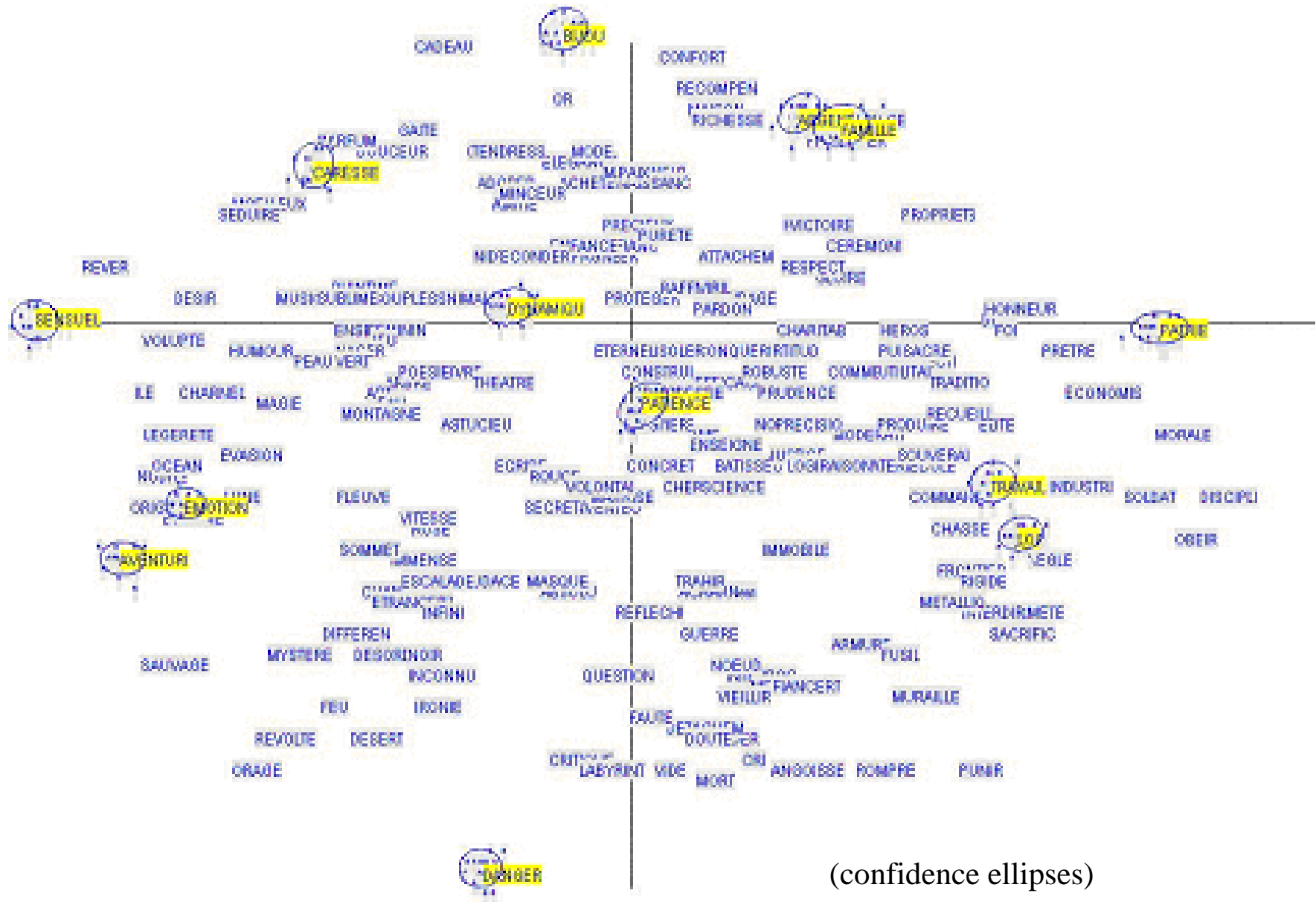


Words “in progress”: desert, punish, mask, rigid, wild, cry, doubt, fault.



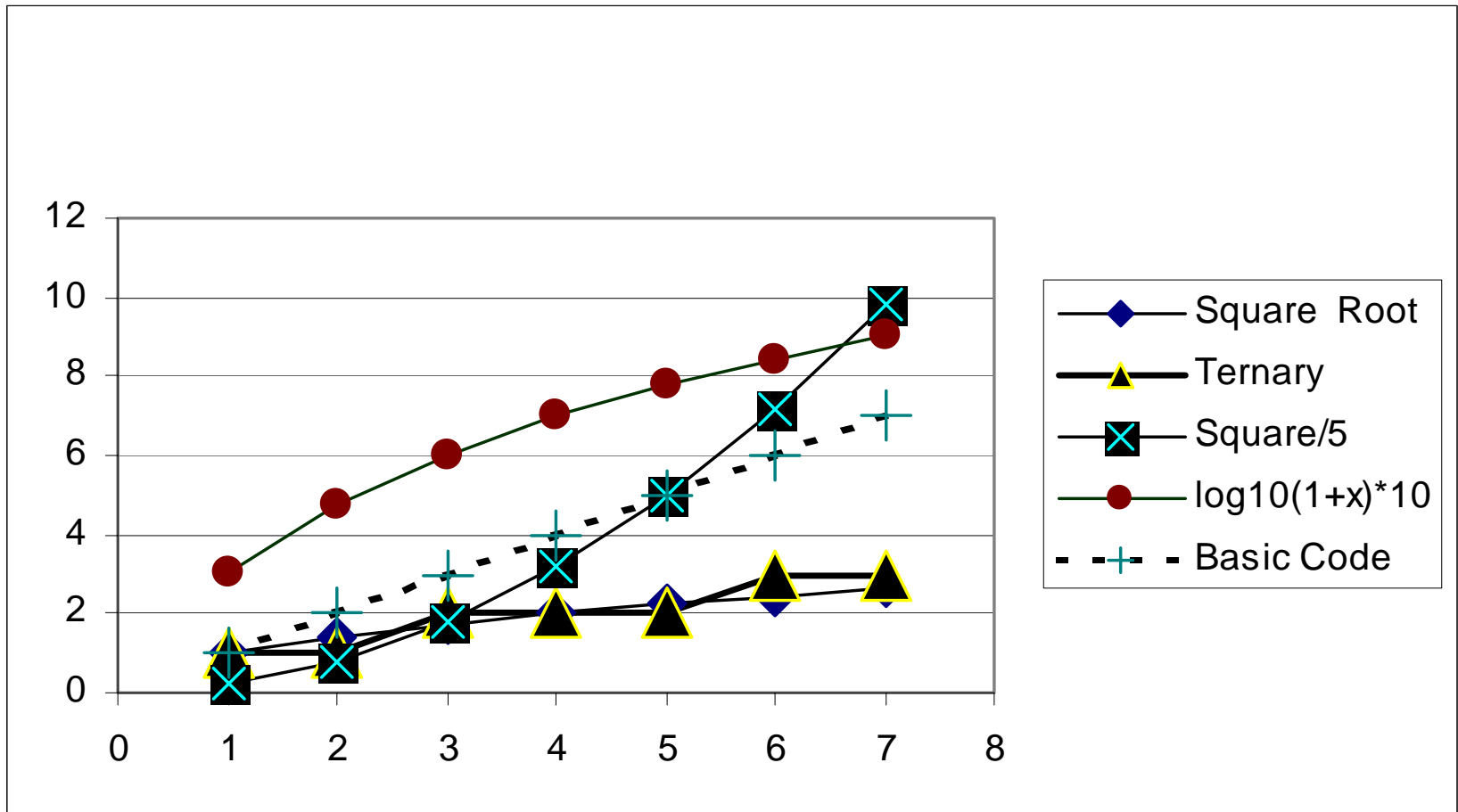
1998 versus 1996

## IV. Stability of results



(confidence ellipses)

## Stability vis-à-vis the initial coding



## V. The “Open Semiology”

Instead of starting from a pre-established list of words, the following experiment has been conducted: it has been asked to a sample of 1191 respondents to spontaneously quote the words they consider as most agreeable, and the words they consider as the most disagreeable (experiment realised in France in 1995).

The responses constitute a text made of 41547 words(or: tokens), 7170 of which being distinct words (or: types).

| <i>French</i> | <i>frequency</i> | <i>English</i> |
|---------------|------------------|----------------|
| amour         | 731              | love           |
| soleil        | 628              | sun            |
| vacances      | 498              | vacation       |
| #maladie      | 475              | #illness       |
| #guerre       | 439              | #war           |
| #mort         | 421              | #death         |
| amitié        | 379              | friendship     |
| enfant        | 366              | child          |
| famille       | 358              | family         |
| fleur         | 328              | flower         |
| bonheur       | 305              | happiness      |
| voyage        | 292              | travel         |
| mer           | 286              | sea            |
| musique       | 277              | music          |
| #chômage      | 265              | #unemployment  |



| Characteristic<br>Spontaneous words | Test-<br>Value | Proba-<br>bility | English<br>translation |
|-------------------------------------|----------------|------------------|------------------------|
| <i>Male, less than 30</i>           |                |                  |                        |
| 1 dormir                            | 3.24           | .001             | to sleep               |
| 2 plaisir                           | 3.03           | .001             | pleasure               |
| 3 manger                            | 2.84           | .002             | to eat                 |
| 4 loisir                            | 2.29           | .011             | leisure                |
| <i>Male, over 55</i>                |                |                  |                        |
| 1 courage                           | 3.59           | .000             | courage                |
| 2 fraternité                        | 2.80           | .003             | brotherhood            |
| 3 propreté                          | 2.68           | .004             | cleanness              |
| 4 santé                             | 2.24           | .012             | health                 |
| <i>Female, less than 30</i>         |                |                  |                        |
| 1 chocolat                          | 3.06           | .001             | chocolate              |
| 2 bébé                              | 3.02           | .001             | baby                   |
| 3 animaux                           | 2.46           | .007             | animal                 |
| 4 maman                             | 2.26           | .012             | mummy                  |
| 5 câlin                             | 2.13           | .016             | caress                 |
| 6 été                               | 2.12           | .017             | summer                 |
| <i>Female, over 55</i>              |                |                  |                        |
| 1 merci                             | 3.00           | .001             | thank you              |
| 2 affection                         | 2.83           | .002             | love                   |
| 3 politesse                         | 2.53           | .006             | politeness             |
| 4 bonjour                           | 2.14           | .016             | Hello                  |

# Conclusion

- Versatile tool, likely to be used in many circumstances.
- International comparisons made possible in domains considered as intra-cultural.
- Evolutions and chronological variations, thanks to the relative permanence of the items.
- Richness and simplicity of the open questionnaire.

Thank You

Domo Arigato Gozaimashita

Merci

Danke

Obrigado

Gracias