



How can *rank distance* help in development of CALL material?

(A comparative study of needs analyses for producing e-material)

Anca Dinu

University of Bucharest
Bucharest

Liviu P. Dinu

University of



Content

- EUROMOBIL project briefing
- Mathematical addenda: Rank distance (RD), Scaled total rank distance (STRD)
- Needs analysis similarity
- Final remarks



EUROMOBIL project briefing

- Student mobility today is supported by numerous EU funded programmes and bilateral contracts between educational institutions.
- The success of a study course abroad strongly depends on the language skills of the students and their knowledge of the culture and social conventions of their host country and host university.



- EUROMOBIL project (www.euro-mobil.org) is a Multimedia Language Learning and Information Programme on CD-ROM with links to the web
- It has been designed with the support of the European Commission
- Its specific purpose is to offer exchange students the opportunity to prepare themselves efficiently for their stay abroad and to raise awareness for intercultural contacts.



- The first part of the project, Euromobil1 (72139-CP-2-2000-1-FI-L2) had as target groups exchange student that were coming to study in Germany, England, Hungary and Finland.
- More programmes to include French, Portuguese, Czech, Polish and Romanian are presently under development as part of the second part of the project, Euromobil2 (225825-CP-1-2005-1-FR-LINGUA-L2).
- The programme can be adapted for further languages.



Needs analysis similarity

- At the beginning of the EUROMOBIL project, a needs analysis survey was conducted in close collaboration with exchange students in order to plan the contents of the programme.
- Exchange students were asked to complete questionnaires in order to reveal their problems in each of the host countries. Thus, project partners could filter out the accidental problems from the current problems the exchange students are confronted with.



Idea

- The idea was: given that most of the needs analyses answers are rankings, can we compare the problems exchange students were confronted with, while studying in a foreign country, and to draw a conclusion upon the similarities?
- What formal method shall we use?



- This article compares and interprets the data and results of Euromobil2 needs analysis by using *Rank distance* (Dinu, 2003), a metric between rankings which was applied with good results
 - in bioinformatics (the DNA similarity problem) (Dinu and Sgarro, 2006)
 - in computational linguistics (similarity of Romance languages) (Dinu and Dinu 2005, 2006)



Rank distance

- A ranking is an ordered list of objects
- Let $u=x_1x_2\dots x_n$ and $v=y_1y_2\dots y_m$ be two rankings of lengths n and m , respectively.
- For an element x_i in u we define its order or rank by **$\text{ord}(x_i | u)=n+1-i$** ;
i.e., the rank of x_i is its position in ranking, counted from the **bottom to the top**.



- **Definition 1:** Rank Distance between two rankings u and v is given by:

$$\begin{aligned}\Delta(u, v) &= \sum_{x \in u \cap v} |ord(x|u) - ord(x|v)| + \\ &+ \sum_{x \in u \setminus v} ord(x|u) + \sum_{x \in v \setminus u} ord(x|v)\end{aligned}$$



Example

Place	Name	ATP Points	ord
1	Federer	7035	5
2	Nadal	4545	4
3	Nabladian	3360	3
4	Ljubicic	3055	2
5	Roddick	2380	1

ATP rankings June 12 2006

Place	Name	ATP Points	ord
1	Federer	6725	5
2	Nadal	4765	4
3	Roddick	3085	3
4	Hewitt	2490	2
5	Davydenko	2390	1

ATP rankings Jan. 2, 2006

$$\Delta(\text{June, Jan}) = |5-5| + |4-4| + |3-0| + |2-0| + |1-3| + |0-2| + |0-1| = 10$$



Drawbacks of rank distance

- The main idea of RD is that it penalizes the unmatched objects (i.e. founded just in one of the rankings) more if they are situated in the initial part of the rankings.
- Scaled total rank distance was created to also penalize the matched objects more if they are situated in the initial part of the rankings.



Scaled Total Rank Distance

- Definition 2: *STRD* between two rankings A and B of length n is:

$$S(A, B) = \sum_{i=1}^n \frac{\Delta(A_i, B_i)}{i(i+1)}.$$



Needs analysis similarity

- The analyzed data (most relevant questions) are grouped into three categories: language, studying and socio-cultural similarities.
- We use the following strategy: for each language, the answers are categorized and the categories are then ranked according to their score.



Language problems similarities

- The results of the student's answers at the question B1 "Please evaluate your language proficiency on your arrival in the host country" are given in *table 1*:

Question B1.	CZ	FR	PL	PO	RO
None	68.75%	5.30%	30.77%	40.82%	40.54%
(very) poor	18.75%	21.30%	33.85%	42.86%	24.32%
fair/satisfactory	8.33%	53.30%	15.38%	12.24%	16.22%
Good	0.00%	17.30%	16.92%	2.04%	8.11%
Fluent	2.08%	0.00%	3.08%	2.04%	5.41%

Table 1



- We extracted from *table 1* a ranking of levels of proficiency arranged in descending order of percentage of students' answers, for each country (*Table 2*). For the sake of simplicity we denoted None by N, (very) Poor by P, Fair/Satisfactory by S, Good by G and Fluent by F.

CZ	FR	PL	PO	RO
N	S	P	P	N
P	P	N	N	P
S	G	G	S	S
F	N	S	G	G
G	F	F	F	F

Table 2



- We want to find out how similar are these rankings w.r.t. each other. Thus, in order to interpret these data, we need to compute the distance between the 5 rankings. We summarize the computed distances in the *table 3*:

STRD normalized	CZ	FR	PL	PO	RO
CZ	0.00	0.60	0.43	0.36	0.03
FR	0.60	0.00	0.53	0.43	0.53
PL	0.43	0.53	0.00	0.09	0.37
PO	0.36	0.43	0.09	0.00	0.36
RO	0.03	0.53	0.37	0.36	0.00

Table 3



- We can see that every time FR finds itself at the biggest distance from every other language.
- This is due to the fact that the exchange students coming to France are more advanced in studying the language of the host country than all the others.
- This supports the decision of the project partners to design for France material for advanced level and for the other countries for beginners level.



- We proceed in a similar manner of analysis with the rest of the answers. We only list here the questions, the table containing the computed distances and its interpretation.
- The computed distances for the question B2 “To what extent do you have difficulties in the following areas: understanding the speech of teachers and room mates, understanding complex sentences, understanding the content, understanding the connecting thought, participating in a discussion?” are given in *table 4*:



STRD normalized	CZ	FR	PL	PO	RO
CZ	0.00	0.24	0.38	0.10	0.10
FR	0.24	0.00	0.57	0.13	0.20
PL	0.38	0.57	0.00	0.63	0.57
PO	0.10	0.13	0.63	0.00	0.07
RO	0.10	0.20	0.57	0.07	0.00

Table 4

- We can see that except for PL, for the other four languages (CZ, FR, PO and RO) the differences between their rankings are very small, between 0.07 and 0.24. We remember that the distance is normalized (all possible values range from 0 to 1). This shows that exchange students have very similar difficulties w.r.t. the language problems mentioned in B2 question.



Studying problems similarities

- From the answers to the question C3 “Were you used to different forms of studying at your home university?” one can see that most of the exchange students find some different forms of studying in the host university (table 5):

Question B1	CZ	FR	PL	PO	RO
Yes – some are different here	54.17%	74.70%	41.54%	48.98%	36.73%
No – all same as at home	37.50%	14.70%	44.62%	26.53%	20.41%
Yes – all are different here	6.25%	9.30%	9.23%	22.45%	21.62%

Table 5



- In order to better prepare the exchange students for the usual forms of study and examination of the host country, question C2 “What forms of academic activity are you involved in at your host university? Lectures, Seminars, Practical training, Laboratory work, Written examinations, Oral examinations, Language courses” determines exactly which forms of academic activities are used in each host country.



- Because more than one category appears in a position in resulting rankings, we used a distance related to scaled total rank distance, that can cope with such situations.
- *Table 6* contains the computed distances between rankings for question C2:

Normalized Distance	CZ	FR	PL	PO	RO
CZ	0.00	0.36	0.21	0.61	0.17
FR	0.36	0.00	0.45	0.4	0.54
PL	0.21	0.45	0.00	0.67	0.25
PO	0.61	0.4	0.67	0.00	0.79
RO	0.17	0.54	0.25	0.79	0.00

Table 6



- We can distinguish two clusters of countries w.r.t. the similarities of the forms of academic activities: CZ, PL and RO on the one hand (with values of the distance between them of just 0.17, 0.21 and 0.25) and FR and PO on the other (the distance from east European countries to FR and PO varies from 0.54 to 0.79, the maximum possible value being again 1).
- One possible explanation for these results could be the regional pattern of teaching: more lectures for the first group and more student activities for the second.
- This means that the project developers should include in the e-material information about the system of teaching and examination *specific* for each group of country (or even for each country).



Socio-cultural and everyday situations similarities

- From question D7 “Have you ever found yourself in a situation in the host country that was funny or embarrassing because you broke a cultural convention?” we can see that the percentage of students who broke some cultural convention is relatively small for all countries (*table 7*).
- It is remarkable that one in four exchange students studying in France found themselves in a funny or embarrassing situation because of that.



Question D7	CZ	FR	PL	PO	RO
yes	16.67%	25.30%	10.77%	2.04%	8.11%
no	77.08%	64.00%	87.69%	97.96%	83.78%

Table 6

- The computed distances for the question B6 “Have you had any difficulties in dealing with practical matters or getting to know the university as: Registration, Housing, Banking services, Finding your way around the campus, Using the library, Computing center services, Health care services, Meals?” are given in *table 7*:



Normalized distance	CZ	FR	PL	PO	RO
CZ	0.00	0.50	0.41	0.53	0.25
FR	0.50	0.00	0.53	0.35	0.47
PL	0.41	0.53	0.00	0.78	0.79
PO	0.53	0.35	0.78	0.00	0.48
RO	0.25	0.47	0.79	0.48	0.00

Table 7

- We can only remark that the lowest distance between the difficulties when dealing with practical matters is reached between CZ and RO. Practically, except banking services, which are more problematic for exchange students in RO than in CZ, the rankings for the two countries are quasi the same.

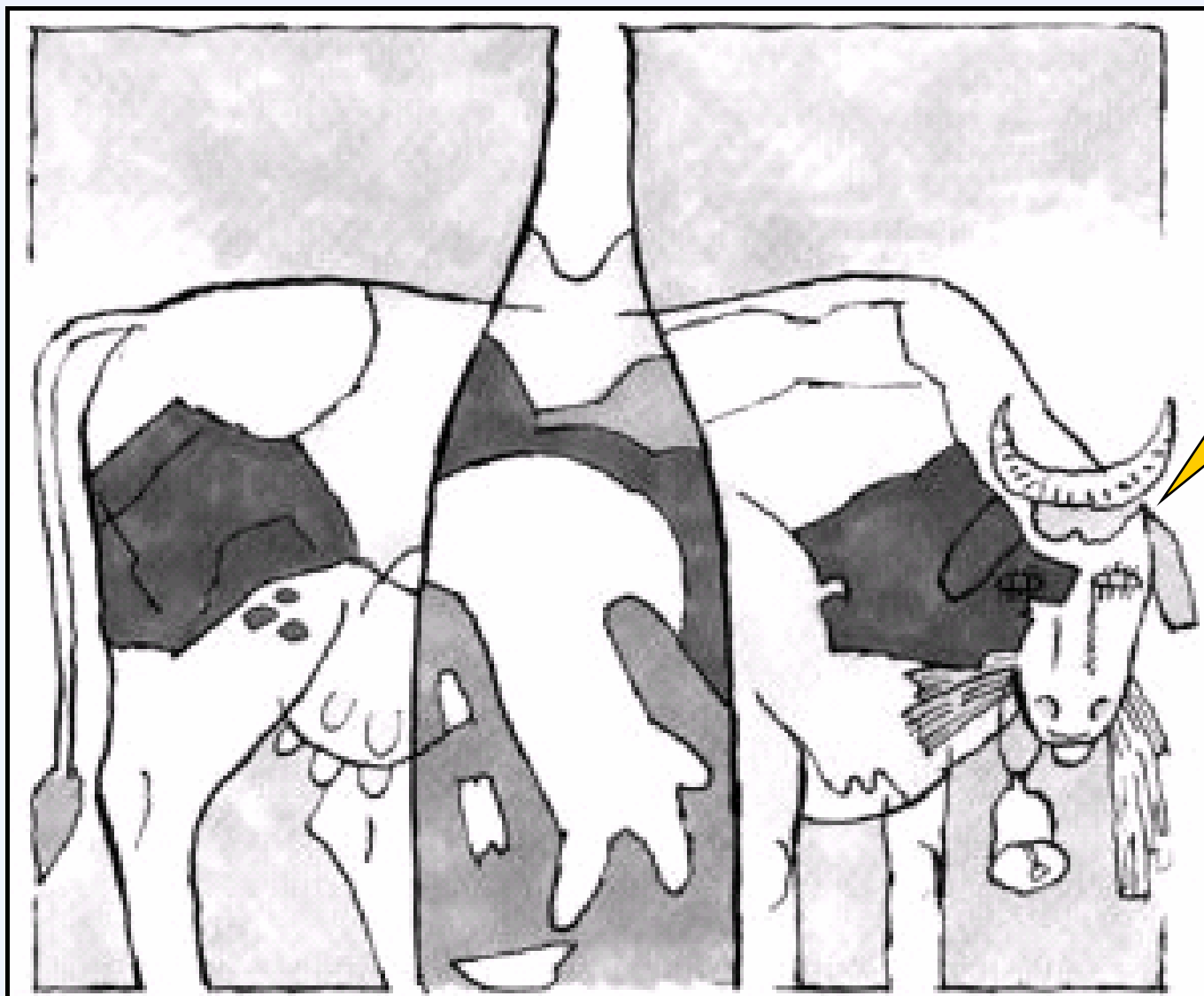


- For the rest of the countries, the distances are quite high, so the developers of the e-materials have to include information about dealing with practical matters or getting to know the university, independent of each other, according to the specific priorities for each country, as they emerge from needs analysis.



Final remarks

- Our research not only confirms and/or details some natural intuitions, but offers some insights, otherwise inaccessible through direct observations.
- The results of this comparative study are expected to aid in the development of efficient e-material, specific for each target group, during Euromobil2 project.



**Thank
you!**