MARIE CURIE ACTIONS

11th International LAB Meeting - Spring Session 2008

European Ph.D. on
Social Representations and Communication
At the Multimedia LAB & Research Center, Rome-Italy

Social Representations in Action and Construction
in Media and Society

"Identity and Social Representations:
Cultural and Mythical Dimensions"

From 26th April - 4th May 2008
http://www.europhd.eu/html/_onda02/07/13.00.00.00.shtml

Participants Presentation

EUROPh D

European Ph.D

on Social Representations and Communication

International Lab Meeting Series 2005-2008

COMMUNICATIVE CONSTRUCTION OF THE SARDINIAN IDENTITY:
Brief insight by using Lexical Analysis of Texts collected from Sardinian Print Media

GIOVEDÌ MARZO 2005   SARDEGNA

università di Aston
remia il sito sardo
www.comitau.org

sito elaborato completamente in sardo, dove la home page diventa "domu". Un archivio telematico d'informazioni sulla lingua isolana, dalla grammatica alla grafia, fino ai neologismi, per parlare e soprattutto scrivere correttamente in limba. Questo è, in sintesi, lo spirito e la funzione didattica di "Comitau.org" (www.comitau.org), ovvero l'Arcivu Sardu de Bilinguismu. Il supplemento della testata giornalistica www.isardi.net, lanciato sulla rete quattro anni fa, ha conseguito un importante riconoscimento scientifico: l'Università britannica di Aston, con il contributo della sezione lingue dell' Unesco, in seguito ad una ricerca sui siti internet dedicati ad alcune lingue minoritarie, ha apprezzato www.co-mitau.org - Arcivu Sardu de Bilinguismu quale miglior sito di lingua sarda. L'equipe internazionale, coordinata dalla professoressa Sue Wright, che dal 5 al 7 aprile sarà ospite dell'Università di Pescara per illustrare i risultati della ricerca, ha setacciato la Rete per l'individuazione e il monitoraggio del più alto numero di siti scritti del tutto o in parte in lingua sarda. Ne sono stati censiti circa duecento, ma il più delle volte si trattava giusto di qualche pagina. I veri e propri spazi web completamente in sardo, come Comitau.org, sono meno di una decina. Il gruppo redazionale della rivista on line è formato da una trentina di collaboratori. «La redazione di Comitau.org è nata in ambito universitario», spiega Amos Cardia, coordinatore editoriale del sito. «Il nostro lavoro è la naturale prosecuzione dell'attività culturale svolta dal "Comitato di studenti per lingua sarda", l'associazione studentesca che nel 2001 organizzò all'Ersu il primo corso di alfabetizzazione alla lingua sarda». L'obiettivo primario di Comitau.org è sensibilizzare la comunità sarda alla tematica del bilinguismo, «che, a nostro parere, deve essere uno strumento indispensabile di sviluppo della società isolana».

Maria Carrozza

sa revista digitale de su   Sòtziu Limba Sarda

Vaduz   Roma
Svizzera
Bern   Monaco   (Sardegna)   Tunis
(Parigi)
Francia
Andorra   El-Jazair   (Algeri)
Andorra la Vella

Shards of Theoretical Background

- According to Tajfel no social group consists of individuals who will all react in the same way to conditions in which they live; and non social group is an island in the same sense in which no man is an island (The Social Psychology of Minorities, 1978:6)

- Marková argued that: "The sense of the past is, both for individuals and for communities, a fundamental precondition in their sense of identity. The past concerns something real and tangible. Building on the past, re-cycling and preserving something one already knows provides a sense of security. To destroy the individual's past involves altering his/her perception of, or hope for, the future." (1997: 13)

- M.Billig (1991) argues that themes such as 'us', 'our homeland', 'our nations' 'their nations' 'honour', 'morality', 'national duty' are themes widely diffused as common sense (Banal Nationalism, 1995:4). In defining a homeland Billig writes that each homeland is to be imagined both in its totality and its particularity.....each homeland must be considered a special space, separated physically and metaphorically from other homelands. In the eighteenth century, it was common for the British, and especially the English, to imagine their island as being God's chosen.

- Anderson thus addressed the qualities which must be an essential feature to each homeland: "the special quality, which marks the homeland as 'ours' continues without diluition right up to the borders: and there it stops, to be separated from the different foreign essence which is marking out the territory on the other side" (1983:26).

- Anderson defined a physical entity such as a land to a shared identity by pointing out that a homeland or nation is imagined as a community because despite the cultural and/or economic differences of its members "is always conceived as a deep, horizontal comradeship." (1991: 7). What is now a concept or imagined identity still needs to be limited with finite although elastic borders.

- According to Hutchinson (2005) Newspapers, e-journals and magazines together with commercial capitalism are two important factors in mirroring the policies of identification and representations with nation states or territories or historic groups within the nation states. Back in 1991 Anderson wrote book and newspapers' publishers were among the first capitalists and being mostly concerned on making profits they were likely to encourage publishing in the widespread vernaculars languages (1991:38).

- On the other hand scholars such as Billig believe in the role of active élites in the context of argumentation and for this reason the need for logic or " art of witcraft" is seen by Billig as deeply intermingled not only with the form of language but also with its contents (1987:133). In addition the contents of common-sense thinking are not "a unitary store of folk wisdom, but instead it may provde us with our dilemmas for deliberation and our controversies for argument."(1987:222)

- This small presentation reports on some previous results of an analysis conducted on a sample of 717 newspapers articles focusing on 4 different narratives of Sardinian identity which were collected between 2003 and 2007 in Britain and in Italy. The newspapers involved in my study were the three major Sardinian newspapers: The "L'Unione Sarda", "La Nuova Sardegna" and "Il Sardegna" (printed from 2004 on).

DATA COLLECTION & TREATMENT

- The overall objective of this preliminary study is to obtain data which can be used together with other sections of the thesis work to explore the communicative construction of Sardinian identity and its Social Representations. Today I will only focus on a description of the work I am doing on four paths of Sardinian identity in the three main regional print media in Sardinia. In particular I will focus on one single newspaper "L'Unione Sarda". It comes as no surprise that if we singled out the study alone it will give us only a shortsighted view on the vast amount of data.

- The whole corpus is made of 717 articles collected by all the major Sardinian daily newspapers between 2003 and 2007. By using a boolean approach I have sorted out 4 macronarratives related to various aspects of Sardinian identity building that and as far as L'Unione Sarda (58% of the whole corpus) is concerned articles collected are 416 which are divided as follows:

  - Ethnic and cultural identity;(Unione Sarda=248 articles= 59.61% of UnioneSardaCorpus)
  - Autonomy(Unione Sarda=84 articles=20.19%);
  - Language as vehicle of identity(Unione Sarda=47 articles=11.30%);
  - New Immigrants (Unione Sarda=37=8,90%)

- As far as this preliminary stage is concerned the data which I will show are grounded only on text units analysis rather than on context units. Context will be analysed at a later stage together with the data stemming from the 4 web forums. Furthermore, the main author who developed the software I am gonna use, (Bolasco 2002) suggests to divide the text into more sub-subject areas or corpora. For this reason in the future I will analyse each macronarrative separately in order to reduce the lenght of the whole corpus (the ideal one being no much than 100 written pages).

- For the moment I decided to leave year 2003 apart given the fact only two of the three newspapers were issued in 2003. The third one, "Il Sardegna", was first issued on February 2004 so the articles being primarily analysed were 687. Once the corpus was created and the variables were chosen (narrative; month; year of publication; newspaper), we set out to run the first stage of data mining which is Grammatical Tagging.

TaLTaC2.5

 We chose TaLTaC2 which is the acronym of "Lexical and Textual Automatic Processing for Corpus and Content Analysis". TaLTaC2 is an application for the analysis of a series of texts (corpus), aimed at describing and interpreting the corpus' content and/or some of its characteristics. It is based on what is known as "lexicometric approach", allowing for the direct study of any type of data expressed in natural language: documents, interviews, press reviews, messages.

- The automatic analysis of texts hinges on the logics of both Text Analysis (TA) and Text Mining (TM). Such an analysis allows to define a quantitative representation of the phenomenon under study, both at the level of text-units (words) and context-units (words). Consequently, both the language and the contents of the text can be examined.

- The software was created at the Universities of Salerno and at Rome "La Sapienza" during the 1990s under the supervision of Sergio Bolasco, it is the result of the cooperation between several Italian and French universities.

- Taltac employs both statistical and linguistic resources. These are highly integrated with each other and can be customized by the user. This allows both at the lexical and textual level for text analysis and information recovery and extraction according to the principles of data and text mining.

-

  The automatic processing, based on a lexicometric approach, indeed allows researchers to find a number of recurrences in the text, a sort of DNA of the corpus.

- As far as my research is concerned I find these sort of softwares conjugating linguistic and statistical tools quite important not only because they help analysts to extract information on language use by automatic lemmatization, but it also gives researchers the opportunity to better contextualize the use of language and its rhetorics (in this case rhetorics of identity).

Describing/Interpreting

- At this stage of my work my data underwent only a first stage of interpretation focusing on grammar and occurrences. I need to set up a second level of interpretation by shaping semantic universes and by reflecting on semantic and pragmatics. As Fairclough points out (1989-2001: 117) phonology, grammar & vocab are the surface of utterance whilst semantics & pragmatics are the meaning of utterance.

- So what I am trying to do by using TalTac in the first step is descriptive interpretation: recognizing and sorting out words, phrases and sentences. This level provides us with an insight on textual & lexical richness.

- The second level of interpretation consists in assigning meaning to those parts of the texts, either they be single words, chunks or entire paragraphs. In this case is the analyst that must make sense of the utterances by contextualising them. As we see in the first page some images are used in different contexts from those in which they originate

-
  Thus even though a software helps us to de-construct and re-construct semantic fields by combining word-meaning and grammatical information and "work out implicit meanings to arrive at meanings for whole propositions" (Fairclough 2001:120) it is clear that the importance of the analyst and the knowledge of its corpus is paramount.

Few words on the data found

- A preliminary analysis of the data found in Sardinian media reflects the growing awareness of and interest in the different issue of Sardinian contemporary society, in terms not only of its changing in terms of occurrences in Sardinian print media (let's mention for example the growing importance of a recent narrative which describes newly landed immigrants in a region where migration took its toll over the decades), but also in terms of micro-regional differences (the newspapers printed in the central areas cope more with issues related to Sardinian identity as built on ethnicity/culture rather than on demographic change or new cultural and social challenges which followed a relatively massive immigration from the second semester of 2006 on.

TreeTagger

Variable: L'UnioneSarda	P value= >3.9 (level of significativity) The text is highly political and rich.	N° of occurrences	N= 443402 V= 40434 (V/N)*100= 9,119 Richness index as far as years 2004-2007 are concerned is higly significative.	N=total occurrences V=active variables That means many occurrences and few words that occur more than once	9,119 means the text is very dense and rich without too repetitions	
	Sardegna	2.072	NPR	lavoro/sviluppo/cultura/città/università	359/351/350/349/158	N
	sardi	886	N	Presidente/Soru	331/332	N/NPR
	sardo	829	N	Italia/Stato	290/170	NPR/N
	sarda	797	ADJ	Italiano	249	ADJ
	autonomia	463	N	Europa/Mediterraneo	158/1745	N
	lingua	751	N	Cagliari	736	NPR
	isola	394	N	Sassari/Nuoro	325/143	NPR
	Isola	264	NPR	territorio	240	N
	terra	751	N	Roma	172	NPR
	identità	384	N	LSU	13	NPR
	Regione/Regionale	333/362	N/ADJ	Clandestini/convivere/algerino	131/4/8	

Crossing TreeTagger with Rep90

By crossing our internal vocab with the external source (Rep 90, namely a vocab based on all "LaRepubblica" issues from 1990 to 2000.) the software gives us a measure of the richness or banality of our texts when compared to it. The higher the differential with Rep90_FG, the better as far as the originality of the corpus is concerned.

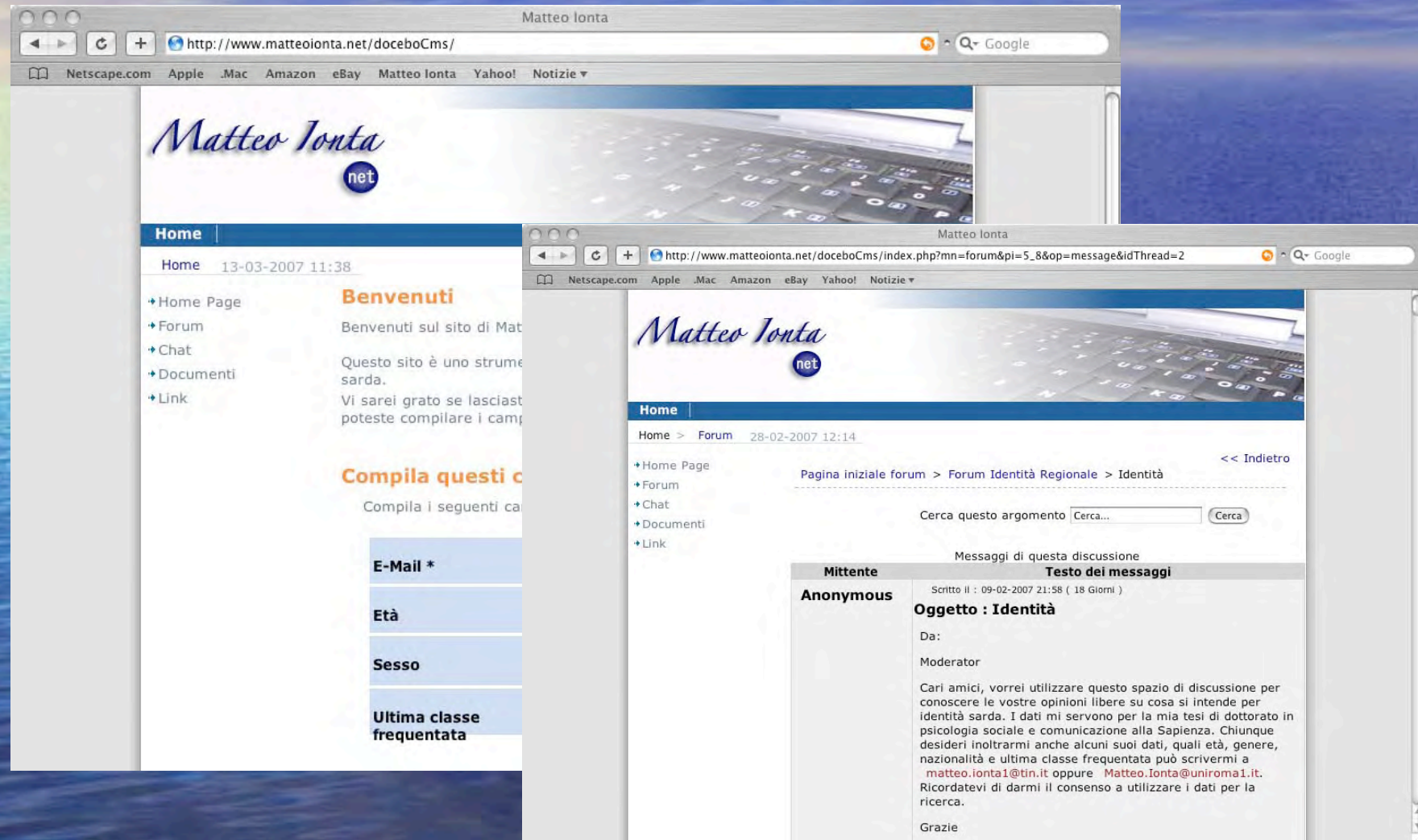| Differential	Word	Occurrences
300,3770836 53776	sarda	548
56,20921112 24379	Mediterraneo	1745
25,23743218 5651	isola/Isola	1797
17,35569236 35944	territorio	756
7,498536974 22481	sardo	1.157
27,07662877 38715	Regione	987
7,514360865 38156	Statuto	998

Rep90 is TalTac's external source which analyses what is called: peculiarities, whilst the specific characteristics of the corpus I have created must be analysed separately with a proceed that analyses all the frequencies and what are the "positive recurrences". From then on the research becomes quite personal as the researcher will create the semantic universes by himself.

TalTac also gives the possibility to transfer or insert data from other softwares such as Spss/Lexico3/Spad_T.

MACRONARRATIVES	N°of Articles Found Over 5 years
Ethnicity/Culture 58,71%	n°= 421
Autonomy 21,48%	n°= 154
Language 11,16%	n°= 80
New Immigrants 8,65%	n°=62

Macro	Ethnicity/Culture	Autonomy %	Language	New Immigrants
2003	20 (66.7%)	3 (10%)	6 (20%)	1 (3,3%)
2004	34 (72,34)	7 (14,90)	6 (12,76)	0 (0)
2005	102 (51,51)	56 (28,29)	37 (18,69)	3 (1,51)
2006	194 (67, 83)	57 (19,93)	22 (7,69)	13 (4,55)
2007	71 (45,51)	31 (19,87)	9 (5,77)	45 (28,85)

As we can see in the tab, articles on EthnicCultural identity kept steady over the years together with those articles related to Sardinian autonomy and identity. In a way Ethnic/Culture and Autonomy have in common the revisitation of the past, the memory of colonization, the belief in an alleged group affinity bot cultural and ethnic. The necessity to shape and re-shape commonsense through the in-group has been widely investigated by British scholars, yet, the majority of those scholars did not refer to SR in order to describe the processes of identity construction and date-back ingroup origin through rhetoric and metaphoric discourse. Noticeable language was given less space after the Region of Sardinian issued a common standard for the language in 2006. Standardisation process has to do with many factors of inclusion and exclusion in a number of social economic spheres that are gaining some importance in Sardinia: employment in Public Office, Digital Tv, Publishing. However as we may infer from the scarce number of articles dedicated to language once Sardinian standard was achieved, this issue is not a "hot" one. It losts its appeal and its presence in the media dropped despite language is an issue still at stake in the political agenda, academic and activist agendas. On the other hand, new immigration became a consistent narrative from September 2006 on especially when a consistent wave of young people from Algeria started to land over the island. The wave was depicted as a dangerous flood by the media.

Further research was conducted by creating 4 forums of discussion one for each of the narratives found on the Print Media.

SOME REFERENCES

Anderson, B., (1983, 1991) *Imagined Communities*, Verso, London

Bart, F. (1996) Ethnic Groups and Boundaries, in J.Hutchinson, A.D. Smith (ed) *Ethnicity*

Billig M., (1987) *Arguing and Thinking. A rhetorical approach to social psychology*, Cambridge University Press
Billig M., (1995) *Banal Nationalism*, Sage Publications, London

Bolasco S. (2002) Integrazione statistico-linguistica nell'analisi del contenuto, in B. Mazzara (ed.) *Metodi qualitativi in Psicologia Sociale. Prospettive teoriche e strumenti operativi*, Carocci Ed. Rome, pp. 329-342.

Bolasco S., Giovannini D. (2002). Il Trattamento Automatico Lessico-Testuale per l'Analisi del Contenuto (TALTAC): un'applicazione allo studio delle immagini della formazione professionale trentina, in B. Mazzara (ed.) *Metodi qualitativi in Psicologia Sociale. Prospettive teoriche e strumenti operativi*, Carocci Ed., Rome, pp. 343-361.

Fairclough N. (2001). *Language and Power*, second edition, Pearson Education, UK

J.Hutchinson, A.D. Smith (1994) *Nationalism*, Oxford University Press
J.Hutchinson, A.D. Smith (1996) *Ethnicity*, Oxford University Press

Jaffe A. (1999). *Ideologies in Action. Language Politics on Corsica*, M. Heller, J. R.Watts (ed.), Mouton De Gruyter, Berlin-New York

Marková, I., (1997) The Individual and the Community: A Post-Communist Perspective, in Orford & Stephenson (ed.) *Journal of Community & Applied Social Psychology*, 7, pp.3-17
Spence D. (1982) *Narrative Truth & Historical Truth*, New York, Norton

THANKS FOR YOUR ATTENTION!!