New approach to explore and to study cartographical heritage in dialectology: application to the Linguistic Altas of France

Paule-Annick Davoine*, Silvia Gally, Philippe Garat***, Carole Chauvin**, Oton Čopi*, Camille Cavalière***

* Univ. de Grenoble-Alpes, LIG, CNRS-UMR 5217, F-38000 Grenoble, France
** Univ. Grenoble Alpes, GIPSA-Lab, CNRS UMR 5216, F-38000 Grenoble, France

*** Univ. Grenoble Alpes, LJK, CNRS-UMR 5224, F-38000 Grenoble, France

Abstract. This paper introduces GIS and spatial analysis methods for processing dialectology data. Dialectology addresses the study of the linguistic features of languages having a strong oral tradition like local dialects. These linguistic features can be of very different natures: phonetic, morphosyntactic, lexical, semantic or prosodic. Whereas the theoretical approaches used for the construction of linguistic atlas are structured, reliable and homogeneous, data analysis and the elaboration of interpretative maps are still based on manual and non-standardized approaches. This paper deals with the methodologies for integrating geolinguistic data transcribed into ancient maps of Linguistic Atlas of France and for mapping automatically isoglosses interpretaed maps.

Keywords: Dialectology, Cartographical heritage, GIS, Isoglosse.

1. Context and objective

Dialectology addresses the study of the linguistic features of languages having a strong oral tradition like local dialects. These linguistic features can be of very different natures: phonetic, morphosyntactic, lexical, semantic or prosodic. They evolve according to space, time and/or social environment.

To study local dialects, corpuses of phonetic data have been transcribed into Linguistic Atlases. These atlases gather a set of maps on which are registered, for given lexical data (words and concepts) and their phonetic transcription collected at several places or localities (figure 1). In a Linguistic atlas, each locality is identified by one number of inquiry (Contini, 2006). Each map represents one concept (or one word): on the map each locality is associated with a dialectal realization of the concept. Every dialectal realization is transcribed in a phonetic alphabet.

The work has been focused on the Linguistic Atlas of France, which has been printed more than 100 years ago (Gilleron J. & Edmont E., 1902-1910) and currently held at the GIPSA-lab (CNRS-UMR 5216). This document of irreplaceable value supplies the first-rate data for the analysis of the lexical variations. It is used within the framework of international projects, such as the elaboration of a second-generation atlas, called interpretative atlas (Figure 2).

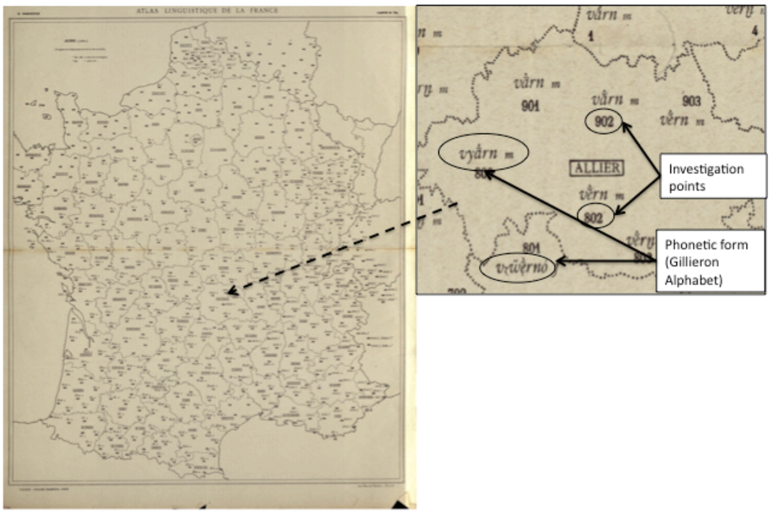

Figure 1. Cartographical heritage in geolinguisitic: Linguistic Atlas of France

Whereas the theoretical approaches used for the construction of linguistic atlas are structured, reliable and homogeneous, data analysis and the elaboration of interpretative maps are still based on manual and non-standardized approaches. One explanation for such a limitation is the fact that software-supported solutions for the exploitation of geolinguistic data have not been yet developed. Most of the work is done manually which is a very time-consuming process. In particular lexical maps remain being hand-drawn as the use of Geographical Information Systems (GIS) and spatial analysis methods is poorly developed for this purpose. Such limitations in the cartographic production capacities do not favour encourage an efficient exploitation of geographical knowledge by researchers in dialectology. It is becoming more and more urgent to propose softwares that will facilitate the extraction, the analysis and the sharing of geolinguistic data.

In the context of GeoDialect and CartoDialect[1] projects, we were interested in how geomatics tools and spatial analysis methods can contribue to improving the processing of geolinguistic data and maps. The case of study is how lexical and dialectological data - originating from historical car-

---

[1] [1] GeoDialect : Labex PERSYVAL (ANR--11-LABX-0025) et CartoDialect : PEPS interdisciplinaire « en réseau » HuMaIn 2013.

tographical documents - could be integrated into a GIS. We also study which cartographical methodologies may be adapted for the treatment of dialectological data.

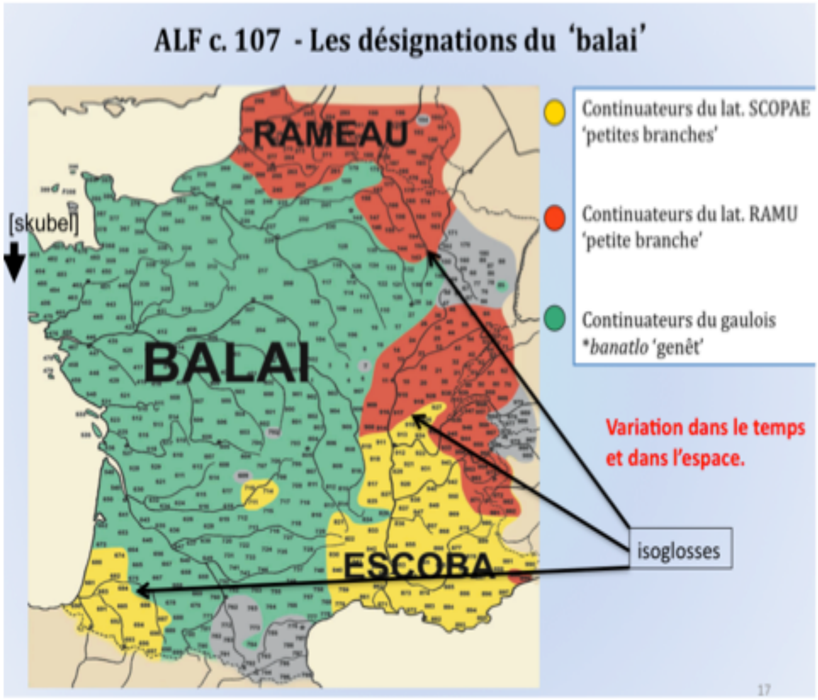

Figure 2. Interpretative map: linguistic (i.e. lexical) aeras are defined by special boundaries, called isoglosses (from Brun-Tringaud & al., 2005: 69).

This paper introduces GIS for processing dialectological data. We firstly present the features of the dialectological data coming from Linguistic Atlases of France. Then, we propose a literature review on GIS and spatial analysis in geolinguistic, before to present our methodology for integrating geolinguistic data of Linguistic Atlas of France into a GIS. Our proposal is based on the design of geolinguistic information layers integrating phonetic, lexical, morphosyntactic, motivational, and geographical dimensions. In the last section, we focus on the geovisualization of dialectological data and more specifically on mapping isoglosses. We propose various algorithms for automated creation of "isoglosses", namely limits that separate different linguistic areas. By taking into account the heterogeneous spatial distribution of geolinguistic points of enquiry, these algorithms allow to create homogeneous linguistic areas based on various dialectological criteria.

2. Cartographical heritage in dialectology: example of Linguistic Atlas of France

2.1. Linguistic Atlas of France

From 1902 to 1910, Jules Gilliéron and Edmond Edmont - pioneers in the geolinguistic data description, handling and analysis - made the first Gallo-Romance dialectal atlas of France, called the *Linguistic Atlas of France o*r *Atlas Linguistique de la France* (ALF). The preliminay plan of the ALF had been reviewed in order to include the Catalan dialects of France (Rousillon region) and the Italo-Romance dialects of Corsica. The field surveys have been held from 1897 to 1901 in 639 localities. The Atlas was published in a printed edition of 12 volumes over a period of eight years, compiled in 35 brochures and about 1920 geolinguistic maps. Furthermore, the geolinguistic maps display a snapshot of the french dialectal situation at the beginning of the 20th century.

The ALF is a first generation atlas made of more than 1.200.000 (1920 maps x 639 localities) raw lexical data, trustworthy transcripted in a homogeneous way by adopting the Rousselot-Gilliéron alphabet: an unique questionnaire have been used stating the localities, the dates and the surveys' conditions. This Atlas represents nowadays a big lexical database and the biggest example of romance lexikon available. Several years later, Gilliéron and Edmont published two Atlas' volumes annexes - *Tables* (1912) and *Suppléments I* (1920) – including an index of the dialectal forms explored and the materials collected outside the lexical questionnaire process. With all these Atlas' materials Gilliéron carried out various analysis and gave birth to the geolinguistics at the beginning of the 20th century. The ALF will serve as a model for many other dialectological and geolinguistic atlas and works undertaken all around the word[2] in the 20th and the 21st centuries.

Nevertheless, the grid of 639 localities was relatively weak to capture a complex linguistic situation of 36.000 municipalities. That is why around 1950 a huge work of *Regional Linguistic Atlases of France* (Regional ALF), have been undertaken, and prescribed by Dauzat at the end of the 30's. Thus, the CNRS supported and financed a project which essentially provided for the progressive realisation of 25 atlases, programming 1 atlas per 'Region' in order to covert all the french national territory. By this huge geolinguistic work - which has now been completed - 70 atlases' volumes have been published in a printed edition: these atlases considerably enriched the lexical database by including, alongside, phonetic and morpho-

---

[2] The Gilliéron and Edmont works inspired, for example, the creation of the *Linguistic Atlas of Italy and Switzerland* (*AIS*) edited by Karl Jaberg et Jakob Jud in 1928, or of the *ALI Linguistic Atlas of Italy* (*ALI*) first edited by Matteo Bartoli in 1924.

syntactical data, with ethnographic and iconographic documents – a major innovation compared to Gilliéron's work[3].

2.2. Interest of Linguistic Atlas of France (ALF)

Nowadays, the largest linguistic database presented in the ALF and in the Regional ALF is used to develop second-generation atlases, the interpretative ones. On the basis of raw ALF data, geolinguists generate "interpretative maps" published in printed editions for different projects, as the *Atlas Linguarum Europae* (ALE) and as the *Romance Linguistic Atlas* (*Atlas Linguistique Roman - ALiR*)[4]. With the introduction of their innovation inaugurated around the 80's, by interpreting geolinguistic and dialectal data, these atlases (i.e. ALiR and ALE) actually represent a way to experiment new approches like new trends for linguistic analysis in semantics and in lexical reconstructions.

Furthermore, the raw data registered in the first generation atlases (i. e. ALF and Regional ALF) can provide materials to analyse phonetic, phonological and morphosynctactical variations in a given geographical area, in synchrony (analysis of the current state) or in diachrony (linguistic variation analysis over time). These raw data enable to highlight, at both, areal distributions of phonetic phenomena (consonants, vowels, syllabic structures and intonational patterns) and phonological systems or morphological elements for nouns and verbs.

The value of dialectal and geolinguistic researches have been illustrated many times over: however it should also be pointed out that it goes beyond the discipline framework. The study of the local names of plants and animals is relevant for the botanists' and specialists' of the animal kingdom; both dialectologists and ethnologists are interested in the relation between words and objects; considerations about diatopical, diachronical and diastratical variation of dialects lead towards collaboration with geographers, sociologists, historians and geneticists.

---

3 Similar works on *Regional Linguistic atlases* have been undertaken in other Romance and european countries like Spain, Italy, Portugal and Romania, but also in the Southern America.

4 The *Romance Linguistic Atlas* (ALiR) was carried out in 1987 by the Grenoble Dialectology Center (France), which joined the *Voice, Linguistic Systems and Dialectology* team, of the *Speech and Cognition Departement*, at the GIPSA-lab (UMR 5216). The aim of ALiR is to sudy lexical, phonetic, phonological and morphosyntactical aspects of all the european romance varieties. Each volume of the ALiR contains one 'atlas' with interpretative maps and one book with the maps' comments and the linguistic interpretations.

3. GIS and dialectology: state of the art

Dialectoly and geolinguistic researches have not been taking the advantages of GIS tools and methodologies (Hoch & Hayes, 2010). The recent technological advances is mainly used for the development of electronic linguistic atlases for the scanning of maps of old linguistic atlases, as the *Linguistic and ethnographic Atlas of Italy and Southern Switzerland* (AIS, http://www3.pd.istc.cnr.it/navigais/), sometimes compined with field booknotes, as the *Linguistic Atlas of Iberian Peninsula* (ALPI, http://westernlinguistics.ca/alpi/more_info.php.). Other projects concern the development of multimedia database (audio recording and atlases' transcripted data) in order to disseminate the results of geolinguistic investigations, as the Thesaurus Occitan (THESOC; Dalbera & al., 2012). On this database and its website (http://bdlc.univ-corse.fr/) phonetics and morphosyntaxics module are integrated and it is possible to download and visualize dialectological data for a selected location. However, these projects don't integrate GIS functionnalities and don't use spatial anlaysis modules neither for exploring nor for map production or spatial analyses. Lee and Kretzschmar (1993) have already mentioned the interest of the overlay and spatial relationship functionalities proposed by GIS to identify linguistic structurations - by combining differents linguistic maps - or to establish relationship between linguistic data and environmental or socio-demographical data.

Alternatively, geolinguistic studies based on dialectometry use quantitative methods for linguistic spatial analysis in order to produce interpretatives maps automatically (Goebl,2008; Aurrekoetxea & Videgain, 2009). Indeed, dialectometry focuses on matrix calculations of linguistic correlation of lexical or phonetic data.

Hoch S. and Hayes J. (2010),have showed that the spatial analysis methods can improve researches in dialectology. Grieve and al. (2011) have studied the regional linguistic variations and identified spatial patterns for specific linguistic variables by using three spatial analysis methods: spatial autocorrelation, factor analysis and cluster analysis.

More recently, many researchers deal with the question of cartographic visualization (Silber & al., 2012; Luebbering & al, 2013; Dell'Aquila & Iannàccaro, 2013), even if Ormeling (1992) proposed in 1992 to use the typology of main maps and symbologies for the geolinguistic cartography.

Concerning geolinguistic data visualization, the most significant advances are reflected in the use of interactive maps in geolinguistic databases or interface websites (see projet ALAVAL[5]). By selecting query on the map, the

[5] Projet ALAVAL (Atlas Linguistique audiovisuel du francoprovençal valaisan) http://www5.unine.ch/dialectologie/Atlas001/001AtlasDemo.htm

user can access both, to semantic contents of the linguistic atlas and, also, to raw data used for creating maps such as video and audio records or notes gathered by geolinguists during the surveys (Diémoz & Kristil, 2012). In those cases the map is used as an intermediary for the database querying. Démioz and Kristil (2012) encouraged the development of interactive maps inspite of difficulties encountered in production of efficient and legible interpretative maps.

However, a global reflexion about the application of semiotic principles to geolinguistic data cartography has been made recently (Ormeling, 1992 ; Silber & al, 2012 ; Dell'Aquila & Iannàccaro, 2013 ; Luebbering & al, 2013). If the constraints of mapping geolonguistic data was put forward in recent researches, none have proposed a general reflexion about methodological solutions and overall techniques about mapping linguistic isoglosses automatically (Hoch & Hayes, 2010; Luebbering & al, 2013)

4. Methodological approach for integrating ALF dialectological data in GIS

In spite of the interest for the Linguistic Atlas of France (ALF) data, this cartographical heritage has never been integrated into a GIS and the survey points have never been georeferenced.

4.1. Methodology for georeferencing survey points of the Linguistic Atlas of France

The georeferencing of data contained in cartographical heritage consists to assign georeferenced control points on scanned maps. This method shall extract some control points (of a known origin) from an old map and match them to a recent map, for example. However, ALF maps doesn't contain reference points aligned with control points. The geolocations of place names on the ALF map don't correspond exactly to the geographical locations of investigated places. These are listed in the appendice in the ALF. From its description it is possible to create a digital version of the investigated places and run the geocoding process to extract the names. Goecoding process consists to the assignment of geographical coordinates to place names by using georefenced names from a gazetteer or geographical databases.

The geocoding accuracy depends on the quality of baseline referential geographical information. However, in recent geographical databases some of the place names have changed or disappeared since 1902. In some cases the spelling may be slightly modified or the name could be written in different ways. In order to take into account ambigous place names or homonyms, two geocoding parameters were taken into account: place name and county code. We have used the National Geographic Institut dabatabases: the BD

Nymes© containing toponyms for villages and hamlets and BD GeoFla© for county names.

The geocoding was realised by a semi-automatic process divided in two stages: 1) by automatical « matching » of ALF place names and reference geographical information layer (BD GeoFla©); 2) when the matching could not be etablished, the place name was inserted into a list on the basis of the expertise of dialectologist researcher and by using the ALF appendice. After the first matching, 553 ALF points out of 621[6] have been geocoded. A second step of geocoding based on the use of BD Nymes© database allowed to georeference the remaining points (figure 3).

Once the ALF point's georeferenced, it is possible to map them (Figure 4). The comparison between a first created map and the ancient - original map of ALF has shown that some sites and some points could not be geocoded. In that case, we used Wikipedia and Google Map to find and to assign the geographical coordinates to these place names.

ID *	Shape *	POINT	LIEU	DEPART	X	Y
8	Point	935	Surjoux	01	5,799903	46,02994
9	Point	913	Villars-les-Dombes	01	5,045984	45,99985
17	Point	776	Lézignan-Corbières	11	2,765101	43,19833
18	Point	784	Rivel	11	2,005488	42,919142
19	Point	785	Ladern-sur-Lauquet	11	2,392941	43,102249
20	Point	786	Tuchan	11	2,736532	42,911309
21	Point	787	Sigean	11	2,979367	43,03428
22	Point	793	Axat	11	2,242055	42,796594
23	Point	773	Fanjeaux	11	2,035164	43,192137
24	Point	718	Laguiole	12	2,866889	44,695653
25	Point	724	Rieupeyroux	12	2,235421	44,303566
26	Point	727	Espalion	12	2,762465	44,519856
27	Point	728	Séverac-le-Château	12	3,078274	44,300124
28	Point	735	Calmont	12	2,525277	44,265039
29	Point	737	Saint-Rome-de-Tarn	12	2,892253	44,03667
30	Point	746	Belmont-sur-Rance	12	2,751691	43,824983
31	Point	748	Nant	12	3,289054	44,02061
32	Point	716	Conques	12	2,41668	44,591572
33	Point	872	Martigues	13	5,047361	43,384848

Figure 3. Georeferenced survey points (localities) of Linguistic Atlas of France (extract)

[6] Les points linguistiques compris dans le territoire français.

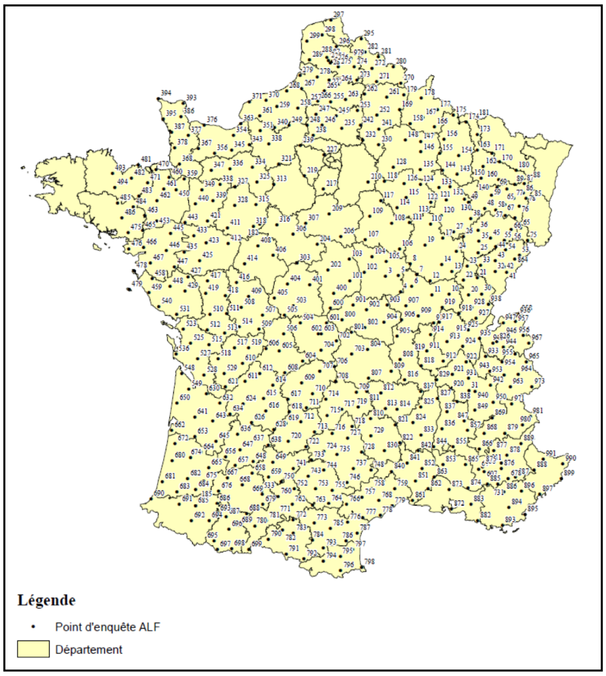

Figure 4. Map of surveyed localities of Linguistic Atlas of France**.**

4.2. Design of dialectological geographic information layers

In geolinguistics different types of data can be distinguished: raw data contained in atlases, as the ALF; interpretative data resulting from lexical, phonetic, morphosyntactical and semantic (motivational) analysis of raw data. Regarding to our project we made lexical researches based on the ALF data, by selecting common features used in lexical analysis as, the phonetic form of a local realisation, the lemma related to a local realisation, its etymology (looking both for suffix and prefix), gender of occurrence ("m" for masculine and "f" for feminine)[7] and, whenever possible, a semantic interpretation based on the motivational approach. According to geolinguistic analysis, other parameters can be taken into account, which depend on the type of analysis (phonetic, prosodic, morphosyntactical, etc…). As regard to the semantic interpretations built, we took account to the degree of reliability of our semantic (*i.e.* motivational) interpretation by taking it a thematic data.

---

[7] Only in those case where a gender information of a local realisation is transcribedon on the ALF map.

Globally lexical data are organized in spreadsheets but to integrate it into GIS they had to be structured as a geographical information matrix. An attribute joint between georeferenced ALF points and thematix data allows creating a geographical information layer for each linguistic notion. The figure 5 presents data structuration for the lexical analysis of "mushroom" on the ALF ("champignon" map), and the figure 6 presents the lexical interpretative map for the "spin top" on the ALF ("toupie" map) elaborated with ArcGIS software (Gally & al, 2013).

lemme	traduction	points	IPA	origine linguistique	prefixe	etymon racine	suffixe 1	suffixe 2	signification	fiabilite
champignon	champignon	1	[ʃˈãpĭɲɔ̃]	latin		CAMPĀNIUS	ŌNE(M)		végétal qui pousse dans les champs	1
champignon	champignon	6	[ʃˈãpĭɲɔ̃]	latin		CAMPĀNIUS	ŌNE(M)		végétal qui pousse dans les champs	1
champignon	champignon	12	[ʃˈãpĭɲɔ̃]	latin		CAMPĀNIUS	ŌNE(M)		végétal qui pousse dans les champs	1
champignon	champignon	13	[ʃˈãpĭɲɔ̃]	latin		CAMPĀNIUS	ŌNE(M)		végétal qui pousse dans les champs	1
champignon	champignon	16	[ʃˈãpĭɲɔ̃]	latin		CAMPĀNIUS	ŌNE(M)		végétal qui pousse dans les champs	1
champignon	champignon	17	[ʃˈãpĭɲɔ̃]	latin		CAMPĀNIUS	ŌNE(M)		végétal qui pousse dans les champs	1
champignon	champignon	19	[ʃˈãpĭɲɔ̃]	latin		CAMPĀNIUS	ŌNE(M)		végétal qui pousse dans les champs	1
champignon	champignon	24	[ʃˈãpĭɲɔ̃]	latin		CAMPĀNIUS	ŌNE(M)		végétal qui pousse dans les champs	1

Figure 5. Interpeted and raw geolinguistic data: example of lexical designation of "*Champignon*" (From ALF).

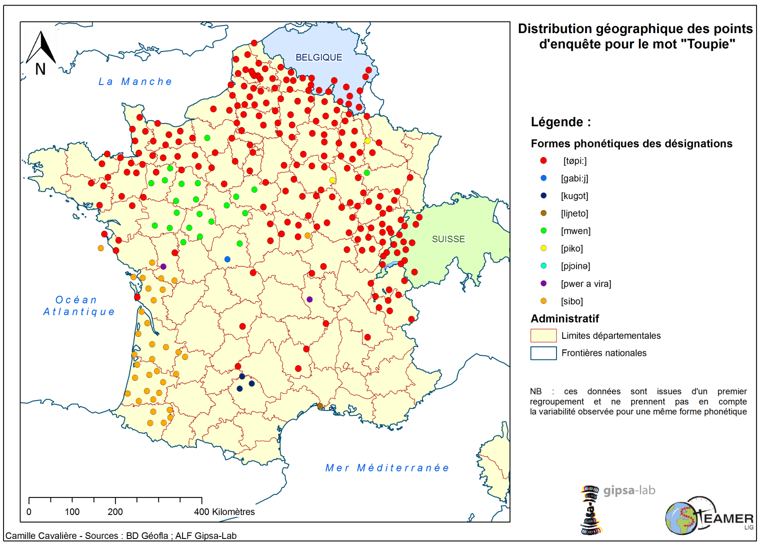

Figure 6. Cartographic representation of linguisitic interpreted data: example of lexical designation of "*Toupie*" (from ALF).

5. Isoglosses mapping

The issue of isogloss mapping is the priority of our project. Geo-linguistic data are by nature non-quantitative but qualitative data and, therefore, most of the usual interpolation techniques, such as kriging, kernel smoothing or weighted regression, cannot be applied to them. The automation of isogloss-lines construction needs specific qualitative-oriented processing techniques. In a first attempt, we have implemented the Thiessen polygons method, which consists in attributing to any locus the nearest observed value. This technique typically generates polygonal-shaped isoglosses when displaying the interpolation results as a map. Thus, aiming at diminishing the *angular* effects of the Thiessen polygons, we have implemented a variant method, derived from an already existing technique of spatial statistics, namely, the geographically weighted regression (GWR).

The principle of the GWR interpolation, when applied on qualitative data, is the following: let Z be a target geo-linguistic variable, and let $\{Z_1, Z_2, \ldots, Z_p\}$ be the set of all its possible values on a given territory. The cardinality of this set is p. Let us denote z_i ($i = 1,\ldots,n$) the values of Z observed at n spatial sites s_i located at geographical coordinates (x_i,y_i) $i=1,\ldots,n$. Let s_0 be an arbitrary point located at some coordinates (x_0,y_0) and with unknown z_o value ; we shall call $\hat{z}_0$ the predicted value of Z at this point. The GWR approach consists in calculating $\hat{z}_0$ as being the value Z_k which maximizes a pseudo likelihood function valuated at locus (x_0,y_0) : without any prior knowledge on the spatial structure of Z, the likelihood of a given value Z_k at locus (x_0,y_0) is seen to be proportional to the sum of all occurencies of Z_k among the sample locations s_i in the vicinity of s_0 :

$$\sum_{i=n}^{n} w_i(\mathrm{x}_0, \mathrm{y}_0)\, I(z_i == Z_k)$$

where $I(\,)$ is the indicator function, equalling 1 for all sites s_i that exhibit the value Z_k. , null otherwise. In addition, the term $w_i(x_0,y_0)$ designates the weight of the site s_i relatively to s_o. The larger the geographic distance between s_i and s_o , the smaller the weight. Most usual weighing scheme make use of *kernels*, such as the Gaussian kernel :

$$w_i(\mathrm{x}_0, \mathrm{y}_0) = \exp(-d_i^2(x_0, y_0)/b^2)$$

with $d_i\,(x_o\ ,\ y_o)$ being the geographical distance between s_i and s_o. The bandwidth parametre b is a scale factor which has to be correctly trimmed using a cross validation technique, for instance. Many kernel types, other than Gaussian, can be encountered in interpolating applications. The Exponential kernel, Uniform, Inverse Distance, etc. are of common use in environmental sciences. Depending on the choice of the kernel and the extend of the bandwidth parameter, the resulting interpolation map can be quite or smooth.

The following interpretative maps (figure 7) illustrate the GWR approach described thereabove, and compare the use of four different kernels. In particular, we have used the *mushroom* lexical data of the ALF and clustered in 28 *lemma* (lexical labels) over the French territory (excepting the Britanny area). Maps (*a*), (*b*), (*c*) and (*d*) show the GWR interpolations run respectively with the Exponential, Gaussian, Uniform and Inverse Distance kernels. Then, map (*e*) shows the Thiessen polygons' method. It can be seen that the GWR interpolation with the Inverse Distance kernel in map (*d*) gives very similar results to the Thiessen polygons' in map (*e*), but with a better level of smoothness. These last two techniques (d) and (e) are said *exact* interpolations because they restitute exactly the sample values, without any loss of information. Whereas the interpolations (a) to (c) are said *non exact*, due to a certain amount of spatial smoothing, which has erased the isolated 'spots', considering that they have low statistical significance or credit due to their rarety.

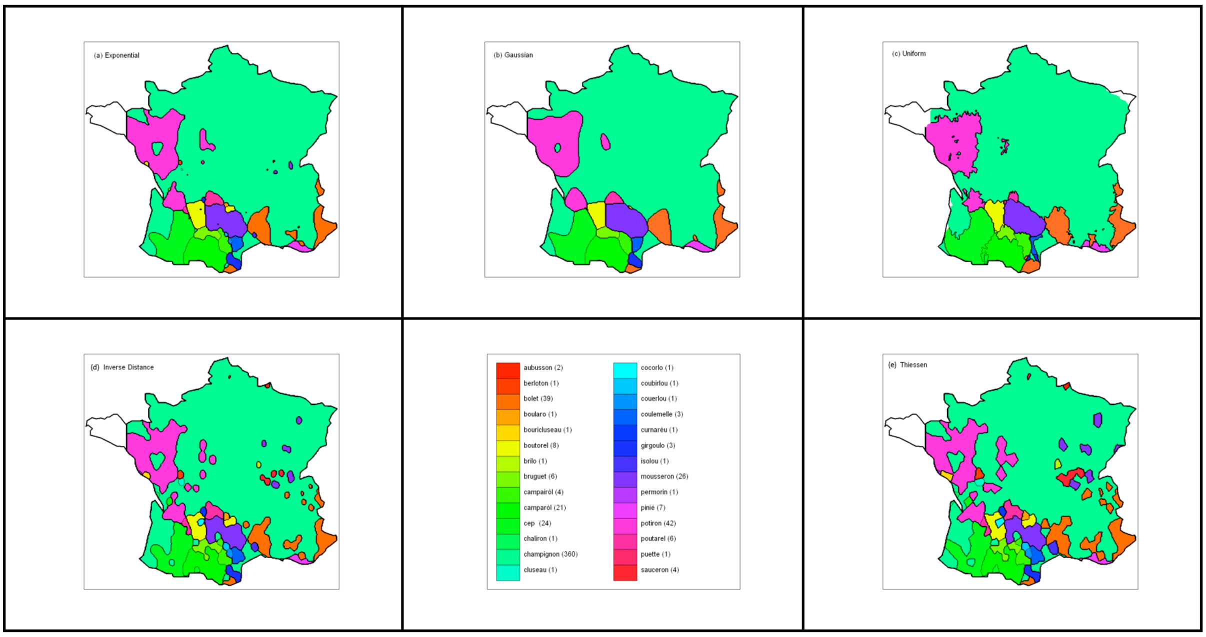

Figure 7. Issoglosses mapping according different appraoches. Example of linguistic designation of "*Champignon*"

6. Conclusion

If geomatics tools and Geographical Information Systems (GIS) are widely recognized for their capacity to manage spatial and cartographical information, few researches regarding the integration of geolinguistic data from cartographical heritage (i.e. lexical data from linguistic atlases) into GIS have been conducted. Nowadays mapping geolinguistic data by suggesting automatic construction of linguistic isoglosses is a real improvement. In present time methodological and technological proposals are rare. In fact, the issues concerning geolinguistic cartographical representation of isoglosses and graphical semiology related had rarely been handled (Luebbering & al, 2013). Predominantly, geolinguistic researchers have focused on specific methods and approaches for raw data collection (as linguistic atlases) and on interpretative analysis deriving from raw data, rather than on making automatic cartographic methods for interpretative maps.

We have studied how the lexical data coming from ancient dialect investigation and mapped on paper document can be integrated into Geographical information System. We have studied also what kind of spatial analysis algoritms may be adapted to treatment of dialectological data with cartophical methodologies. Our proposals define the first specifications of a future Geolinguistic Information Systme for extracting and analysing lexical data from ancient atlas used in geo-linguistics. The objective of this work is twofold: firstly, to save a whole documentary heritage in geolinguistics (the French Linguistic Atlas) and, secondly, to automatize lexical analysis - manually performed until now – with experts in geolinguistics, in order to reveal and exploit a cultural and linguistic heritage that that might be ofinterest to other reaserchers in Linguistics or in Ethnography and in Anthropology.

Acknowledgement

This study was co-financed by the Persyval Labex (ANR--11-LABX-0025) and CNRS PEPS HuMain 2013-2014 (National Center for Scientific Research).

References

Aurrekoetxea, G., Videgain, C. (2009) « Le projet Bourciez : traitement géolinguistique d'un corpus dialectal de 1885 », Dialectologia, n°2, p. 81-111. Disponible sur http://www.publicacions.ub.edu/revistes/dialectologia2/

Brun-Trigaud G., Le Berre Y., Le Dû J. (2005), *Lectures de l'ALF de Gilliéron et Edmont. Du temps dans l'espace*, CTHS

Contini, M. (2006), Quel avenir pour la dialectologie ?, in *I encontro de Estudios Dialectologicos – Actas, Açores 6-7 novembro 2003*. Ponta Delgada, Istituto Cultural de Ponta Delgada : 17-46.

Dalbera J.-Ph., Brun-Trigaud G., Oliviéri M., Ranucci J.-C. (2012), La base de données linguistique occitane Thesoc. Trésor patrimonial et instrument de recherche scientifique, in *Estudis Romànics* 34 : 367-387.

Diémoz F., Kristol A. (2012), L'Atlas Linguistique audioviusel du francoprovençal valaisan et les défis du polymorphisme, in Kattenbusch, D.et Tosques, F. (eds), *20 ans de géolinguistique numérique, Actes du colloque de Berlin (Humboldt-Universität)2-3 novembre 2012. Berlin, Humboldt-Universität zu Berlin* : 160-180.

Gally S., Chauvin C., Davoine P.-A., Demolin D., Contini M. (2013), GéoDialect : Exploration des outils géomatiques pour le traitement et l'analyse des données géolinguistiques, in *Géolinguistique n°14*, Grenoble, ELLUG.

Gilliéron J., Edmont E. (1902-1910), *Atlas Linguistique de la France*, 35 fascicules, Paris : Champion.

Goebl H. (2008), La dialettometrizzazione integrale dell'AIS. Presentazione dei primi risultati, in *Révue de Linguistique Romane* , n°72, 25-113.

Dell'Aquila V., Iannàccaro G. (2013), Quelques considérations sur la cartographie des données linguistiques, in *La Bretagne linguistique n°17/2013* : 235-286.

Hoch S., Hayes J. (2010), Geolinguisitcs : The Incorporation of Geographical Information System and Science, in *The Geographical Bulletin*, 51 : 23-26.

Lee, J., Kretzschmar W. (1993), Spatial analysis of linguistic data with GIS functions, in *International Journal Geographical Information Systems,,* vol 7/n°6 : 541-560.

Lowe D. (2004), Distinctive image features from scale-invariant keypoints, in *International Journal of Computer Vision*, 60/2 (2004), pp. 91-110.

Luebbering C., Korine K., Prisley S. (2013), The Lay of the Language : Surveying the Cartographic Characteristics of Language Maps. Cartography and Geographic Information Science, Nov 1, 2013

Ormeling, F. (1992), Methods and Possibilities for Mapping by Onomasticians, Discussion Papers, in *Geolinguistics* 19-21 : 34-46.

Silber P., Weibel R., Glaser E., Bart G. (2012), Cartographic Visualization in Support Dialectology, in *Proceeding AutoCarto 2012*, Columbus, Ohio, USA.