CULTURAL ENTREPRENEURSHIP FROM LOCAL. AN EXPLORATORY VIEW FROM STATISTICAL TEXTUAL ANALYSIS APPLIED TO THE REPUBLIC OF COLOMBIA

XII International Conference on Arts & Cultural Management (AIMAC 2013)
Bogotá, Colombia 26 – 29 June, 2013

ÁLVARO VIÑA-VIZCAÍNO[1]
Associated Professor
Entrepreneurship and Innovation Unit Director
Universidad Nacional de Colombia
Bogotá, Colombia
e-mail: auvinav@unal.edu.co

OLGA LUCÍA PATIÑO
Cultural PlanningProfessor
Universidad del Rosario
Cultural Management Advisor
Entrepreneurship and Innovation Unit
Universidad Nacional de Colombia
Bogotá, Colombia
e-mail:freya.politologa@gmail.com

ABSTRACT
This work shows how textual statistics allows an analysis in a systematic way through multivariable methods, and having in mind words and contexts, free responses to open question in surveys or interviews. Due so, perceptions from 376 cultural Colombian entrepreneurs surveyed, they expressed regarding their corresponding internal factors: strengths and weaknesses, and external environment: opportunities and threats, formulated under their business plans development, an explorationin 26 Colombian cities.At the same time, the potentialities this method has in the research are illustrated applied to big amounts of textual information analysis.
Keywords: Statistical textual analysis, Lexicometry, Cultural Entrepreneurship, Colombia

INTRODUCTION
Among the concepts about culture given by social sciences, García-Canclini states an operative definition beyond proposing cultural products provided with signs and symbols, he exhibits those as social processes able enough to be transformed through social uses and re-appropriations, because according to him, *"culture holds the set of social production processes, circulation and consumption of signification in social life"*[2]. (García-Canclini, 2004, p. 34).

The interest for the social uses and re-appropriation of cultural products in Colombia have promoted, since 2005, its economic assessment through the elaboration of a Culture Sattelite Account (National Administration department of Statistics –DANE-[3]), which methodology is correspondent integrally with the one proposed by United Nations System of National Accounts.

From that account recent evidence points that in Colombia the goods and services production coming from creation and culture is very dynamic, and it represents 3,2% of gross income, same as the activity generated by Colombian electricity, gas and water sectors; as well as its consumption represents 2,6% of household expenses. (UNESCO, 2011).

1 Corresponding author
2 Free translation from the authors.
3 No date available.

In a society characterized by informality, inequity and exclusion, where 90% of inhabitants hold only 55% of income, results above reflects the strength and dynamism of a relation between production patterns and cultural consumptions by Colombians, and they illustrate the skill of the so called cultural and creative industries to make an impact in the national economy.

It's so like this, that to face troubles culture and economy consider today: social equity, development sustainability, massive media monopolization -which impose certain values and patterns of culture and consumption-, and also, cultural diversity protection, two topics end up being crucial to expand freedom enjoyment by individuals from any society to obtain what's considered valuable: entrepreneurship spirit, from one hand; and the ability people have to innovate, produce and assimilate changes, on the other hand. It means to propitiate social and economic development locally and regionally; where creativity and human innovations are the only and true factors of production and development(Sen, 2000).

According to the stated above, since 2011 the Ministry of Culture of Colombia is developing and integrated strategy in cultural entrepreneurship along its different provinces towards to strengthen the organizational and managerial processes from cultural and community organizations, formalized or not, linked to activities such as music, dance, visual arts, theater, literature, tangible and intangible heritage, cinematography and communications sector. Developing that, assistance and accompaniment to cultural entrepreneurs in touch with organizations just mentioned for the structuration and formulation of their business plans, were in charge of Universidad Nacional de Colombia, the institution linked to the Ministry's initiative through the Agreement 1622 of 2012.

From the cultural entrepreneurship initiatives object of intervention by the pointed Agreement, located in 26 cities of Colombia, under the direction of one these document authors (Viña-Vizcaíno), a qualitative research was performed towards the characterization, from local, the cultural entrepreneurship. So, using a SWOT analysis to 376 formulated business plans, this document explores the conditioners to which cultural entrepreneurs are exposed starting from the textual statistics and the lexicometric analysis.

It's expected from this document to have important academic, pragmatic and public policy implications. From the academic point of view, results will enrich literature about the studied topic. For professionals, the study could be used as a springboard to promote the cultural entrepreneurship spirit. And for those responsible of public policy formulation, results could expose the benefits of perform and some difficulties to overcome.

Document is structured the following way. First, basic concepts about the foundations of this work are exposed: cultural entrepreneurship, lexicometry and textual statistics. Second, foundations and objectives of the research performed are highlighted, as well as the methodology approach and data treatment. Finally it's concluded by showing results obtained and pointing contributions, limitations and the addressing for future research.
BASIC CONCEPTS

2.1 Cultural entrepreneurship

Achieving understanding on any object of study demands clarity of definition. But this hasn't been the case in which cultural entrepreneurship refers; among other circumstances by the wide variety of creative and economic activities these take place, of several sizes, generated employment and other issues. Although, along history, the cultural entrepreneurs list is long, particularly if entertainment industry is included, the entrepreneurship spirit in arts and culture is a quite new concept. (Towse, 2010, p. 107).

So, for the purposes of this research, cultural entrepreneurship is defined as *"the appropriation of symbolic and intangible values from a society in order to create several ways of representation reflected in cultural goods and services, through economic processes based in risk, creativity and innovation, which all together should allow the consolidation of a business idea, company or organization".*(Observatorio de Cultura y Economía, 2011).All this framed, as it's expected, with standards and practices from intellectual property and in particular, the copyrights and other legal rights.

2.2 Discourse Analysis

If cultural entrepreneurships must be at a certain point understandable, those must have their foundations in a common source of ideas, beliefs and stakeholders' interests who perform them, particularly from entrepreneurs. Without this condition, beyond creativeness, culture and the economics, their study could barely be done.

In virtue of this, in this work the entrepreneur's exploration as a social actor, is performed on the classic French tradition from *Discourse Analysis* (Foucault, 1970; Bourdieu, 1999). Derived from applied linguistics and pragmatics, the Discourse Analysis tries to relate the statements to their social, cultural, economic and historical production conditions. As a theoretical practice it takes a discursive praxis located in certain contexts and particular social conditions (in this case the SWOT analysis from the cultural business plans entrepreneurships analyzed); placing it in a necessary interdisciplinary field. In other terms, the characteristics of discursive practices related with participant entrepreneurs in this work are going to reflect not just the vicissitudes of cultural initiatives linked, but also certain characteristics from the socio-cultural, economic and historical environment where they are produced.

2.3 Lexicometry and statistical analysis of textual data

The study of the use of lexical quantification, from counting the units that integrate it, gives the lexicometry its origin. This one is the introduction for the treatment of discursive texts, without its production conditions to be lost: the cultural entrepreneurship of the entrepreneur, and his individual and collective experience herein considered. And so, at the moment to choose for a determined vocabulary, the words as the textual data they are start to configure a repetitive organization of discourse, full of meaning.

From its side, the textual statistics have an origin in the inductive methods application and lineal algebra to mathematic linguistics, derived from developments for multivariable treat of big linguistic data charts done by mathematician J.P. Benzécri in the 60's, to which he nominated *Correspondences Analysis* (Benzécri, 1982); in contrast to deductive approach developed by the linguist Noam Chomski (1956). From the above and the result of his followers, among others, M. Bécue (1991), and L. Lebart and A. Salem (1994), the *French School of Data Analysis* emerges. It's remarkable to point out that these methods haven't had the proper divulgation and application in Colombia.

TECHNIQUES FOR STATISTICAL ANALYSIS OF TEXTUAL DATA
From identification and counting of *graphic shapes* (or words), as basic verbal units constituent of a text *vocabulary,* and *segments* (sequences of two or more words), the *statistical analysis of textual data,* or ADT[4] from now on, allows to explore the similarities among individuals and their shapes; it means, in which way shapes or words used make cultural entrepreneurs look like.

These similarities give origin to a vector which elements are the most used frequencies, and constitute the *lexical profile* of one or more individuals. For example, two entrepreneurs indicate they know and research about traditional music from the Oriental Plains of Colombia, and teach it to new generations; they look alike due to their similarities in their lexical profiles.

Once identified and counted the graphic shapes and the repeated segments, and the text vocabulary created, an*order number* is assigned to make calculations. To do so, it's essential to consider the context of use of graphic shapes; due to their social and abstract character those only have significance in the cultural environment where the communicational interaction takes place. The "cultural manager" segment, for example, in Colombia it could make reference to a cultural entrepreneur, but also to the functionary from private and public sectors who promotes and manages cultural initiatives.

From the shapes vocabulary and the corresponding order number, two contingency charts are constructed and named *lexical chart* and *aggregated lexical chart*. The lexical profile of each individual constitutes the lexical chart. It's a contingency chart (*individual X words*), where intersection of row *i* and column *j*, gathers the frequency which the shape *j* has been used by individual *i*.

From the other hand, aggregated lexical chart is composed by lexical profiles from individuals, when the text has

4 Análisis Estadístico de Datos Textuales – ADT-.

been partitioned in subgroups according to several criteria; for example, subgroups of answers to open questions in surveys, or phrases and paragraphs in prose. It's a contingency chart (*words X texts groups*), where intersection of row *i* and column *j* contents the frequency which shape *i* is used in the text *j*.

The charts above are analyzed through multivariable description techniques; particularly the *simple correspondence analysis* and in a simultaneous and complementary way, the *classification analysis.* The ADT is improved with the lexicometry application, it means, the frequencies in which words are used and the contexts where those are applied (Lebart y Salem, 1994).

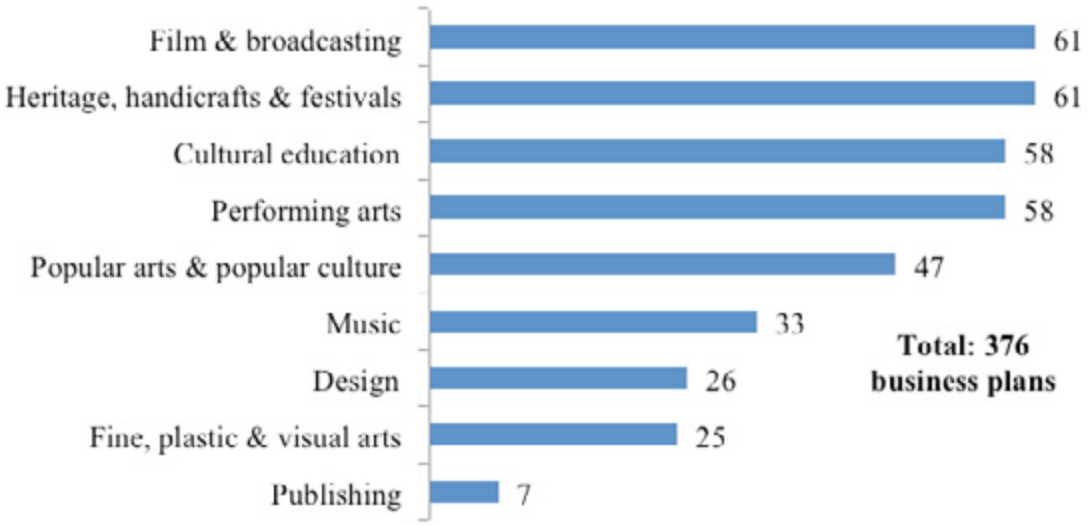
With the *simple correspondence analysis* the different lexical profiles are graphically explored represented as points or vectors in the projection space built by the first factorial axes, taken two by two, due to in a *simple correspondence analysis* a double factorial analysis is performed on the contingency chart: one over the row profiles and another over the columns profile. Its interpretation is done according to *owned values* (quantity of information contained by each factorial axis), and the *inertia percentages* (lexical profiles dispersion measure). This information is contained in an *owned values histogram*, and it's complemented with *coordinates relative coefficients factorial* (quality measure of lexical profiles representation over the axis).

And so, with the *classification analysis* groups are created and characterized so their individual within can be as most homogeneous as possible, and other individuals from other groups can be as most different possible. Two individuals will be similar (and will make part of the same group), if they assume similar values compared to the same variables; it means, when they use the same words. If the contrary, they will be dissimilar ad will make part of different groups.

As a consequence, individuals from a group will enjoy common characteristics *i.e.* the lexicon used differentiates them from other individuals from other groups, it's clear that one single characteristic cannot define a group, in a graphic way or a particular segment. Interpretation is done from any similarity measurement, being the canonical Euclidean distance the most common.

APPLICATION OF ADT IN CULTURAL ENTREPRENEURSHIP FROM LOCAL EXPLORATION. THE COLOMBIAN CASE
For this work, 376 SWOT's –under the traditional focus (internal factors: strengths and weaknesses; and external environment: opportunities and threats)-, were taken for an ADT analysis. These SWOT's were done under the business plans formulation by the same number of cultural entrepreneurs, accompanied by the Universidad Nacional de Colombia in 2012.Figure 1 shows multiple cultural activities in which entrepreneurs are involved with.

Figure 1:Cultural industry where entrepreneurs are involved with.



Computational process was done through the DTM_VIC free software. According to lexicometry, the following results were obtained:

Chart 1: Summary of lexicometric results

Results	SWOT	Strengths	Weaknesses	Opportunities	Threats
Number of responses	1434	347	376	376	335
Number of words (tokens)	24679	5185	7265	7195	5034
Number of distinct words (types)	6506	1380	1812	1852	1462
Percent of distinct words (types)	26.3	26.6	24.9	25.7	29.0

With the aim of reconstruct the lexical significance underlying the SWOT analysis, only those words giving some information were kept, it means, adjectives, nouns and verbs; and the frequency of occurrence was more than 4.The ten first are listed in Chart 2 in descending order of frequency.

Chart 2: Selection of kept words by order of frequency– Frequency threshold = 4

	Strengths	Freq	Weaknesses	Freq	Opportunities	Freq	Threats	Freq
1	Experience (experiencia)	103	Shortages (carencias)	163	Cultural (cultural/ culturales)	93	Shortages (carencias)	48
2	Arts & cultural knowledge (conocimiento especializado)	75	Resources (recursos)	80	Culture (cultura)	49	Market competition (competencia)	46
3	Training &education (formación y entrenamiento)	55	Start-up capital (capital semilla)	24	Marketing (mercadeo)	41	Support (apoyos)	25
4	Years (años)	35	Support (apoyos)	22	Public organizations (entidades públicas)	35	Audience (públicos)	24
5	Entrepreneur team (equipoemprendedor)	32	Working capital (capital de trabajo)	22	Cultural industry (sector cultural)	34	Cultural (cultural/culturales)	22
6	Cultural industry (sector cultural)	29	Equipment (equipamiento)	22	Supply (oferta)	27	Cultural industry (sector cultural)	22
7	Recognition (reconocimiento)	25	Infrastructure (infraestructura)	22	Region (región)	26	Culture (cultura)	17
8	City (ciudad)	22	Marketing (mercadeo)	18	Support (apoyos)	25	Market (mercado)	17
9	Region (región)	22	Informal organizations & administrative process (informalidad)	18	Tendering process (convocatorias)	25	Unfair competition (competencia desleal)	15
10	Quality (calidad de producto o servicio)	20	Lack of administrative knowledge (Falta de conocimiento en gestión)	17	National (nacional)	24	Public organizations (entidades públicas)	15

Chart 2 displays how cultural entrepreneurs' strengths derive basically from their experience and knowledge obtained by training or years of dedication to a cultural activity. It's also important the recognition the entrepreneurship team might have in the province it works. The weaknesses on the other hand, correspond to the lack of the most diverse resources: financial, infrastructure and marketing; it deepens the organizational informality.

Making reference to opportunities perceived by entrepreneurs, these became from supports by the tendering processes for cultural sector offered by national and local public organizations; and the biggest marketing activities efforts it might involve, for example, when it becomes necessary to identify and know better the cultural audiences. Precisely, this topic is pointed among the Colombian cultural sector's threats as one of the biggest annoyances; because of the lack of audience's support or because of the unfair competition from other alternatives they consider being exposed.

Concerning the internal factors, it calls the attention how the cultural entrepreneurs focus the strengths basically in the self (experience, knowledge, training, years), in about 268 mentions. Meanwhile, the weaknesses are focused basically towards the deficiencies from the collective or the entrepreneurship team (shortages of diverse sources), at least 243 mentions. Referring the external environment, they perceive opportunities as those generated by the Colombian State, which joint frequency is close to 54 mentions; while the threats are derived from the market, due to competition or audiences, with 70 mentions. This is displayed in Chart 3 through some responses characteristically from the group; it means, *verbatim* from the ADT made.

Chart 3: Entrepreneurs operational context according to their perception

	English translation	**Selection of characteristic individuals or responses (verbatims)**
Strengths	Certified experience with professional diplomas.	.52 - 2 experiencia certificada con títulos profesionales
	Owner of Teatro Cádiz, location for the enterprise work. 25 years of experience practicing arts.	.46 - 5 propietario del teatro Cádiz espacio de sede para el funcionamiento de la empresa. 25 años de experiencia en el oficio de las artes
Weaknesses	Lack of work capital.	2.06 - 1 falta capital de trabajo
	Low organizational experience, not having a location of our own, low organizational structure.	1.87 - 1 poca experiencia organizacional. no poseer local propio. poca estructura administrativa
Opportunities	Entrepreneurship law. Royalties' resources. Tendering processes participation.	1.52 –2 ley de emprendimiento. recursos de regalías. participación en sus convocatorias
	Development plan for Valle del Cauca towards software development. Cultural industries regulations.	.58 - 4 plan de desarrollo del valle del cauca para desarrollo de software. legislación de industrias culturales
Threats	Society doesn't yet understand the importance carried by arts and culture. Bad communication in the region to achieve it.	.64 - 4 aun la sociedad no comprende la importancia que trae consigo el arte y la cultura. la deficiente comunicación en la región para ello
	Another individual with capital to copy it and started before me.	.47 - 3 otro con capital lo copie e inicie antes que yo

The lexical chart of SWOT analysis crosses 1434 rows of individuals (corresponding answers about it were given by 376 entrepreneurs), with 6506 columns or most frequent words; chart analyzes as illustrative in order to obtain, from one hand, those entrepreneurs who use equal or similar shapes in their perceptions. On the other hand, in order to represent closeness among such words and assess and derive the contexts according to different Colombian provinces. A graphic vision is obtained as it was pointed, through a correspondences factorial analysis of the lexical chart.

Figure 2 shows, for all entrepreneurs from different Colombian provinces (from now on bold highlighted) the perceived strengths represented by 30 words or most frequents graphic shapes, given in their original context. In this figure interesting information can be appreciated, by taking the following "rules" of interpretation:

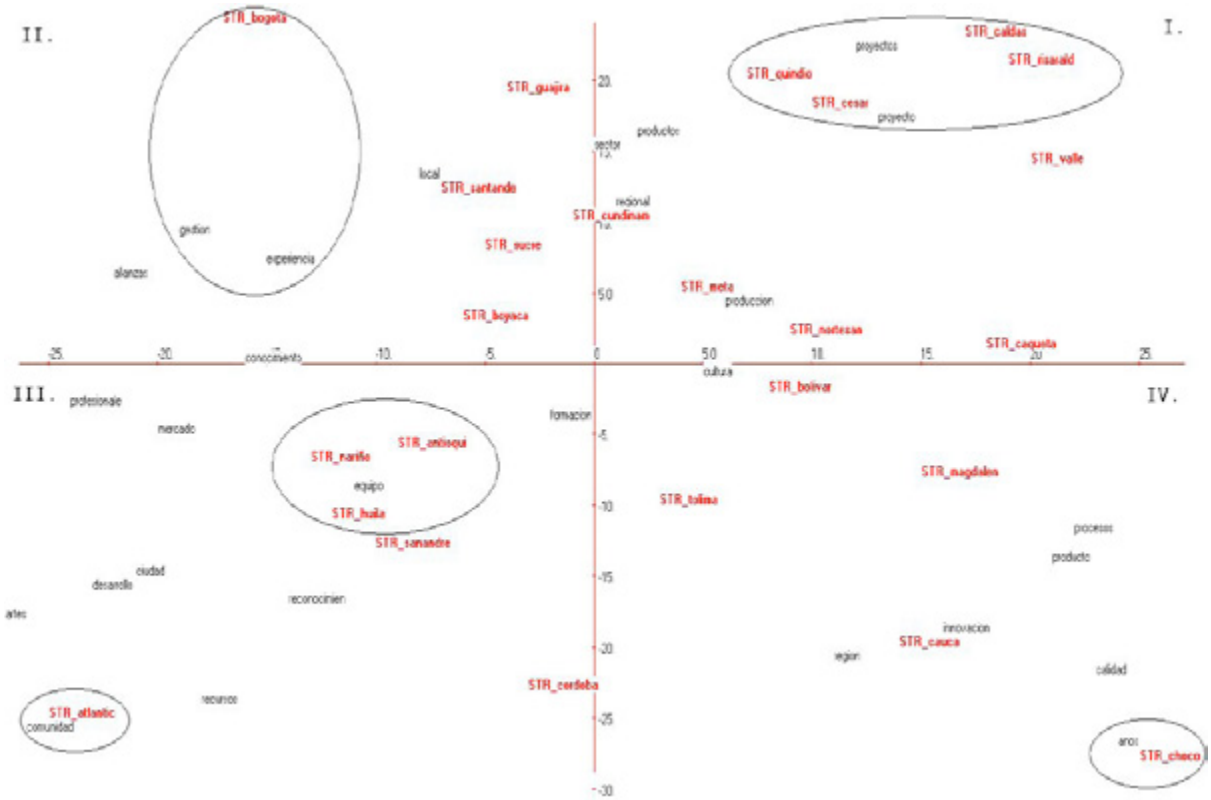
Proximity among the provinces' entrepreneurs: when two or more provinces are close, for example, Caldas and Risaralda, or Quindío and Cesar, in the quadrant I, conclusion can be that these have similar lexical profiles *i.e.* and same graphic shapes were associated, in this case projects (proyectos) and project (proyecto). It means, entrepreneurs from those four provinces, estimate that one of their strengths are their corresponding initiatives or cultural projects.

Proximity between two graphic shapes (words): when two or more words are close, for example management (gestión) and experience (experiencia) present in quadrant II, it can be concluded that associations of individuals to this lexicon are similar and intercorrelated, as it comes relatively with cultural entrepreneurs in Bogotá.

Proximity between provinces' entrepreneurs and words: Combining the two rules above, it can be said that there is a strong association between each word and its corresponding province. It's in such a way that entrepreneurs from Antioquia, Huila and Nariño, represented in quadrant III, estimate that one strength is given by knowledge (conocimiento) bearded by the entrepreneurship team (equipo).

Proximity of words and the provinces' entrepreneurs to origin: When a province is portrayed very close to its origin as in Boyacá province case, it means it behaves as the "average of provinces". For the same reason, a little is differentiated from the rest of provinces. Analogically with the word culture, it has a little descriptive power.

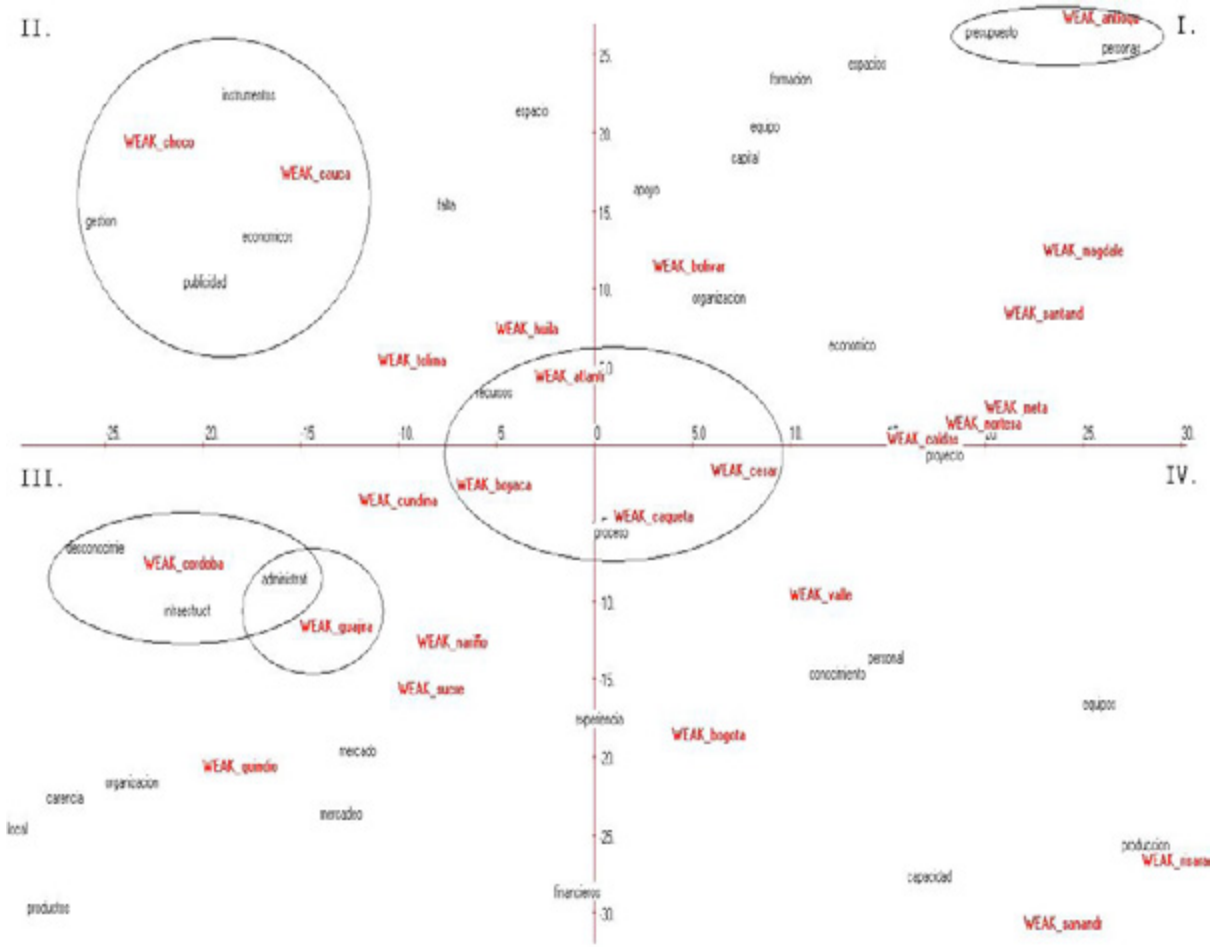
Figure 2 Representation on the two first axes of the strengths present in Colombian provinces



Applying the recently pointed criteria to weaknesses, as it's shown in Figure 3, it can be concluded that cultural entrepreneurs from the provinces of Atlántico, Boyacá, Caquetá and Cesar, represent the average behavior of all provinces: lack of resources (recursos). More specifically, as it's shown in quadrant I, entrepreneurs from Antioquia feel to have more budget difficulties (presupuesto) tan those from the province of Bolívar. From the Cauca and Chocó provinces' side, represented in quadrant II, they express more difficulties concerning to economic resources (económicos), equipment (instrumentos), management (gestión) and advertising (publicidad). In

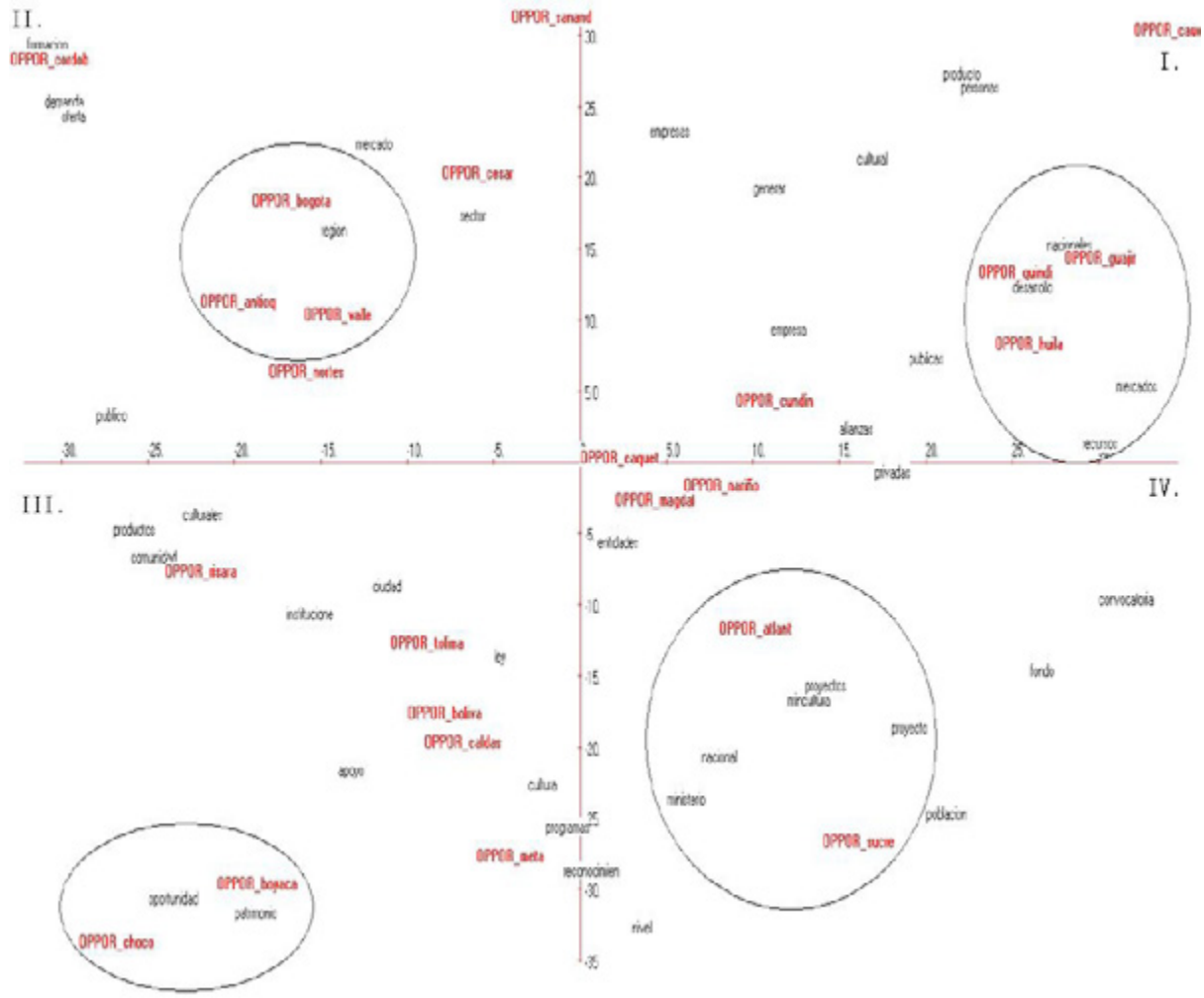
quadrant III is evident how entrepreneurs from Córdoba province state to have difficulties in infrastructure (infraestructura) and management (desconocimiento administrativo). A situation shared with entrepreneurs from La Guajira province.

Figure 3 Representation on the two first axes of the weaknesses present in Colombian provinces



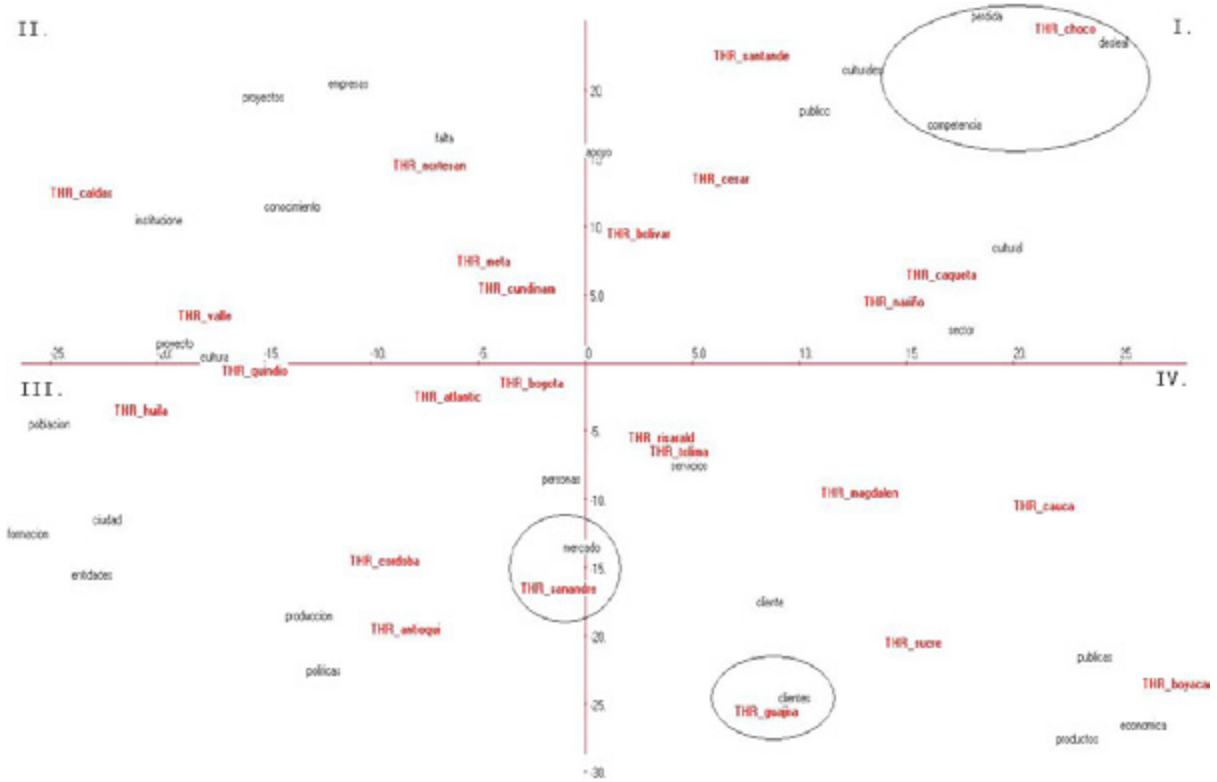
In figure 4 are shown how the corresponding associations to opportunities are opposed between financial resources given to entrepreneurs from tendering processes from national level, basically provided for Guajira, Quindío, Huila, Atlántico and Sucre provinces; meanwhile in the three relatively richest territories such as Bogotá, Antioquia and Valle, these resources come from the cited provinces as shown according to quadrant II. According to quadrant III of this same figure, heritage resources (patrimonio) are perceived as a shared opportunity for Boyacá and Chocó provinces.

Figure 4 Representation on the two first axes of opportunities present in Colombian provinces.



Finally, regarding threats, Figure 5 shows difficulties due to loss of audiences because of unfair competition (competencia desleal) as it's the case of Chocó province in quadrant I; or perception of difficulties due to market threats (mercado) as it's in San Andrés province, represented in quadrant III. And in the same sense, absence of audiences (clients) in La Guajira province, as it's illustrated in quadrant IV.

Figure 5 Representation on the first two axes of threats present in Colombian provinces



When lexicometric analysis results are contrasted listed in Chart 2 with specificities at local level that the cultural entrepreneurs show from different Colombian provinces, just as it was illustrated in Figures 2 to 5, corroborates in the last interpretations given in the chart.

Conclusion

This work was able to reconstruct the context in which the explored cultural entrepreneurs operate, evoking their perceptions from a lexicometric analysis and textual statistics applied to SWOT analysis by those performed in their corresponding business plans. The stated above, allows us to conclude that the contexts where cultural entrepreneurs operate alternates between the individual's strengths, weaknesses of the collective it belongs, opportunities given by the Colombian State and the market threats.

Analytical exploration of data charts as those herein considered wouldn't be possible if the classical focus of "*expost* codification" would have been taken, due to the size of the chart, particularly because of the number of variables or graphic shapes considered. As a consequence, a second conclusion of this work is that textual statistics allows analyzing systematically through multivariable methods *i.e.* correspondences analysis, and having in mind words and contexts, the free response to open questions in surveys and interviews. This way becomes possible the analysis of big amounts of textual information.

However, is necessary to point out that when interpretations depend of the analytical researcher, more than just a weakness of textual statistics, it is considered strength because those leave the chance to make other kind of analysis according to different significations. From this last perspective, the interdisciplinary analysis is considered fundamental.

Independently of the stated above, it's expected from this writing to have important academic implications,

pragmatic and of public policy. From the academic point of view, results will enrich literature about the studied topic. For professionals, the research could be used as a launch to promote the cultural entrepreneur spirit. And for those responsible to make public policies, results could show the benefits of performing and difficulties to overcome.

Colombian cultural entrepreneurs are a target of research with many topics yet to be discovered for social sciences and their opinions go beyond the regular entrepreneurs from other industries, they have ironically management troubles issues like it was stated above, and for their fortune, the Ministry is giving tendering calls and opportunities to improve their business development. But yet they feel like they need more other topics of attention and that's why an interdisciplinary analysis should take place in the future.

ACKNOWLEDGES

Authors wish to express acknowledge first, to the Ministry of Culture of the Republic of Colombia, for the provided support in the frame for the Agreement 1622 of 2012 done with Universidad Nacional de Colombia; this allowed the access to primary information source here analyzed. Second the participant cultural entrepreneurs in the cited Agreement, and the reason for this exploration. And third, the team of interdisciplinary professionals from the Entrepreneurship and Innovation Unit of the Universidad Nacional de Colombia.

REFERENCES

Bécue, M. (1991). *Análisis de Datos Textuales. Métodos Estadísticos y Algoritmos.* Paris: Cisia.
Benzécri, J. (1982). *Historie et Préhistorie de l´Analyse des Données.* Paris: Dunod.
Bourdieu, P. (1999). *¿Qué significa hablar? Economía de los intercambios lingüísticos.*Madrid: Akal.
Chomsky, N. (1956). *Syntactic Structures.* La Haye: Mouton.
Departamento Administrativo Nacional de Estadística, DANE. (s.f.). *Metodologia de la cuenta satelite de cultura.* Bogotá: DANE, Documento No.43.
DTM - VIC. Data and text mining, visualization, interface, clasification. Freely downloadable from the website: www.dtm-vic.com.
Foucault, M. (1970). *El Orden del Discurso.* Barcelona: Tusquets.
García Canclini, N. (2004). *Diferentes, desiguales y desconectados. Mapas de la interculturalidad.* Barcelona: Gedisa Editorial.
Lebart, L., & Salem, A. (1994). *Statistique Textualle.* Paris: Dunod.
Observatorio de cultura y economía. (2011, Noviembre 19). *Emprendimiento cultural. El Concepto.*Retrieved Octubre 11, 2012, from http://culturayeconomia.org/foros/emprendimiento-cultural-el-concepto
Sen, A. (2000). *Libertad y desarrollo.* Buenos Aires: Planeta.
Towse, R. (2010). *A Textbook of Cultural Economics.* New York: Cambridge University Press.
UNESCO. (2011). *Batería de indicadores en cultura para el desarrollo.*Bogotá: UNESCO - Ministerio de Cultura.