Discourse analysis in contemporary history of Mexico: from quantitative to qualitative approach

Carillo, Elsa

Mitglied der

Leibniz-Gemeinschaft

Discourse Analysis in Contemporary History of Mexico: From Quantitative to Qualitative Approach

*Elsa   Carrillo**

Historiographically, our task is to make a comparison between the two last periods of the Mexican political life: the first which corresponds with the Porfirian dictatorship, dating from 1876 to 1910 and the second dating from 1910 onwards. For the present purposes, this second period can actually be said to have ended in 1976 when the current system gave its first concrete signs of weakness. This means one hundred years of political history of Mexico, a period which traditionally has been divided in two by the revolutionary movement of 1910.

This comparison will be based on the political discourses of each of these systems: We will try to find the image that they themselves wished to portray through their own discourse to those who are not in power: what some people now call the »Civil Society«. This should enable us to evaluate the importance of the break caused by the revolutionary movement, even if only at a representation level.

The origins of this study may be found in analysis which we have already done, of discourses concerning only the revolutionary period. These were a series of debates which we actually analyse, Public Administration reports which every president makes every year during his six years of presidency. For our work, we took the first and the last one of each period. These reports are the only Mexican »official discourses« which present enough homogeneity, necessary for this kind of comparison. They have been pronounced in the same institutional situation, with the same political objective, religiously every year, since the Republican system was founded.

Referring to our first work, we shall say that the Aguasalientes Convention sought the unification of the various factions in the revolutionary movement. It is not useful here to present the results of this work. But we must point out that the utilization of very simple computerised lexicometrical methods was the base for analysing, in a more exact and syste-

---

* Address all communications to Eisa Carrillo, Centres de Recherches d'Histoire sur l'Amérique Latine et le Monde Ibérique, 72 Bd. Port Royal, F-75005 Paris, France.

matic way, the different mental universes which were confronting each
other at this convention.

Although the object of this convention was to set up the bases of the
post-revolutionary government and one of the revolution's objections was
to seek democracy, this word does not appear, with its derivations inclu-
ded, more than nine times in the convention speeches. More frequent
words such as the following appear: Here (117), people (93), us (77), re-
volution (77), »patrie« - fatherland - (65), flag (55), Republic (55), power
(19).

Even if these concepts can be categorized as political, it is not always the
case; for instance, »Republic« very often has a geographic connotation.
Besides, the definition of words such as »people« or »patrie«, or even »flag«
or »revolution«, very often differs from one speaker to the next. Confron-
ted with this divergency of conceptualizations, how could one find a mi-
nimal agreement for a government, if there was not even a consensus
about what was and who formed the country - la patrie -, or about the
causes of revolution? However, this was the kind of discourses which oc-
curred in Mexico, once the revolution was over. It is interesting for us,
therefore, to sec to what degree, at what moment and what intention this
divergency of »criterion«, this lack of coherent political definitions and,
finally, this noncommunication was conserved by the party born from the
revolutionary movement. At the same time, it is interesting to know how
much this kind of discourse is only a revolutionary product or if some of
its elements already existed in the »Ancient Regime«.

Besides the general idea presented above, methodologically we propose
a path which tries to pull out all schemes pre-established for the analysis.
We want the discourses to tell us where to go, in which direction we must
search, rather than telling them what they should say at this or that mo-
ment (1). And, in fact, a reliably and rigorous way of doing so, is with the
help of some computerized statistical methods of Lexicometry.

After the computer has given us the directions for research, it is up to
the historian to continue the research in those directions and, thus, give a
background for the interpretation of the computer results.

For the Convention, we worked with a program called ALINE which is
installed in the »Centre Inter Regional de Calcul Electronique« (CIRCE)
at the Orsay University in Paris.

This program had a lot of possibilities, but is becoming useless because
it never had using manual and people who had worked on it, are now gone
from the university scene. Nevertheless, with help from an informatizian
at the University of Paris I., we were able to get some interesting work
instruments.

First of all, ALINE gives an exhaustive list of words in order to descen-
ding frequency of appearance. Of course, all words which linguists call

»tool-words« - articles, conjunctions, etc. - appear as well; but it is from this list that we set our parameters of the analysis, by looking at the words that are the most frequent - except for »tools-words« - or, on the contrary, looking for words that we expected to appear several times and which are actually quite infrequent.

The program also gives an alphabetical list, where each word is accompanied by the references which allow us to find a word in a text and, thus, study the contexts in which that particular word is used. These references were previously put in alongside the text.

This reference is also useful for applying the program in different ways. For our work, we have made two kinds of listings: firstly, we ask ALINE to apply the program of the intervention; secondly, we ask for a listing by all the speakers, to extract the vocabulary of the group of conventionalists as a whole unit. In that way we were able to make comparisons between the different speakers themselves and to compare each of these results with the general vocabulary employed at the convention.

This reference contains not only a code for the name of each speaker, but also the date of the intervention, the volume where it can be found, the page in this volume and the line on the page where the word is printed. This means that we can analyse the intervention of X or Z, or X and Z, pronounced this or that day, in order to put them in relation with events and, thus, to observe the evolution of their argumentations.

In our President's work, we will apply the program by speaker as well, but we could apply it to all entrance discourses separately from all the latest discourses, in order to compare the enthusiastic investiture from the moderated departure of the presidents; we could also include the subject and the Ministry they are talking about in the reference and, thus, make a study on the evolution of a lonely Ministry, applying the program only to the sections of all the discourses which concern that Ministry.

In fact, these references allow us - besides their first function of localization - to proceed in different figure combinations for different kinds of studies, our discourse data bank, thus, becoming a little mine of research subjects.

Finally, ALINE gives us the number of different words that are utilised in the discourse. This function is called »number of forms« and it is very useful for providing an idea of richness of vocabulary of each speaker (2).

Some other information which is given by ALINE is:

- the number of lines which were treated
- the maximum length of forms
- the average length of these forms
- the maximum frequency
- the average frequency

6

- the total number of words
- the number of separators (,. : ; etc.)
- and the size in octets of each of these information.

All these renseignements are very useful in deep linguistic analysis. For our study, the utilization of the frequency and localization lists and the number of forms and words were quite enough.

Once, all these information have been re-organised and analysed, we arrive at the interpretative stage. After we have taken out the concepts with their different definitions and have compared them, we replace the text itself in the historical context in which it has been elaborated. In this way, we can compare this reality with the image that speakers had of it. In other words, we try by this method, bring out their own definitions and attitudes in front of a given reality, in a way which, we think, prevents as much as possible, the interference of our own views. Thus, quantitative may allow us to do a more detached qualitative analysis of, by itself am-biguous discourse, as are the political's.

But ALINE is not the only program suitable for this kind of work; after doing the Convention vocabulary study, we have learnt of similar pro-grams which are very good. Here we will give a general view of »Lexicloud« programs functions which, for the moment, work only on mainframe and about another one which actually works on microcomputer as well as on mainframe.

The first one is a series of programs which have been developed by a linguistic research team at the »lnstitut National de Langue Francaise«, a group which forms the research unit of »Lexicomctry and Political Text«. For our kind of work, this group has the advantage of working for over 15 years especially on Political Text. Their studies arc extremely performed and their analysis goes very deep into the text. Some of their programs have been adapted to Russian and to Spanish language.

The second one is the microcomputer version of the Oxford Con-cordance Program (OCP), which was developed at the Oxford University Computing Service, by Jeremy Martin. This program has been elaborated »after a widespread consultation with potential users«, so it appears as a very simple utilization program which is easily applicable to all kinds of texts: from poetical text to political text and in several languages.

With reference to facility, we shall expose OCP's results first. For OCP, once the text is converted into ASCII on the microcomputer, this program gives us a list of words by descending frequency; then an alphabetical list with the frequency of apparition and the line references where a word can be found. With this program, other references can be used - speaker, scene, etc. - so, in this alphabetical list, the number of times that a word appears within each speech - or scene, etc. and the line on which it appears, will be listed beside the total frequency.

Thirdly, it gives a list of words with their total frequency and the context lists for each of their apparition (occurrence), accompanied by the number of line and reference which indicate where a word appears in the text.

This context (concordance) is modifiable and can be viewed in several formats. This is, mainly, what OCP does. There are, of course, much more possibilities (as defining our own alphabet set or asking for some simple statistical results) which would need a more lengthy explanation than we can give here.

Lecicometrical and Political Text group, work by »items« which can be phrases or »segments« of text and which are defined by punctuation signs. All their works are based on deep statistical calculus into the text. Thus, they do not only work on words, but on groups of words (segments, co-occurrences, couples, etc.), looking as well for the subtilization of words which form these items, by their disambiguation (with homonyms, for instance) or their »lemmatization« - categorization - (grouping singular and plural of nouns, feminine and masculine of adjectives, etc.).

These groups of words or items can be found in several forms:
-   two or three words which appear always together, one before the other (COUPLE of words);
-   a phrase (words between strong punctuation signs) which is frequently repeated (SEGMENTS);
-   two words which appear frequently near one another, separated by some other words (COOCCURRENCES).

Most of these items are imbricated one into the other in the text and the analysis of Lexicloud programs tries to get the different kinds of imbrications a text may contain. In fact, this group tries, and is nearly succeeding, in getting a sort of radiography of the text in order to have as much information, as quickly as possible, about the construction of a text. In other words, they show how an idea is introduced to the public.

So, the results of these programs include not only the typical alphabetical or descending frequency lists but also an enormous amount of information which must be gradually asked by the user, depending on his own questions about the text.

Perhaps, and they themselves always set it up in their articles, they have not yet arrived at the level »radiography« as they would like. Some technical and methodological questions remain to be resolved. But as users we think they have done a lot and their actual results are quite enough when doing deep studies of political texts.

For users the problem of capture still remains, although steps towards the perfecting of scanners have been made. For historians, it will be a great feat, when scanners can recognize manuscripts. This leaves us wondering as we know that historians work not only with manuscripts, but even with

old impressions whose characters are not as well formatted as they are with newest computerised printing.

Faced with this situation, there are a lot of historians who still resort to using the computer as a simple typewriter, rather than an information bank or a work-tool of research. But the time will come, we hope, when all this nightmare of compatibility, information, and above all, the long and deserting capture will be merely a question of three months preparation research work.

Notes

1. Because we think that in doing so, we risk to misinterpret a text. Traditionally, we analyse a text from a certain point of view, knowing in advance what we want to pick up from the text. Sometimes - as in the convention case - we do not find that what we think is important for our subject of research, and so, we fail to notice other characteristics of the text which can be more important than those we are looking for, or which give an indirect explanation to our research objectives. So, we think it is best to start with as much an »open mind« as possible.
2. In putting it in relation with the total number of words.