

Chapter three: Methodological steps

3.1 Introduction

Discourse analysis, like pragmatics, studies language as it is used in communication (Johnstone, 2002). Pragmatics, as was already explained in the previous chapter, attempts at understanding people's intentions behind their linguistic choices, namely context-dependent meanings. By the same token, discourse analysis is interested in language beyond the lexical level. Although it looks at patterns, as do semantics, syntax, and morphology, discourse analysis is more interested in how language is related to and reflects social and cultural conventions (Paltridge, 2006).

In this respect, the social and cultural settings determine how discourse is used and understood, and consequently how a person's social identities are expressed. We say identities¹ because, more often than not, people belong to more than one discourse community or communities of practice (Gee, 1999 as cited in Sunderland and Litosseliti, 2002). That is to say that a group of people who share a common activity, hold the same beliefs and values, and, more importantly, communicate in ways that are particular to them (Paltridge, 2006). Thus, an individual's identity and the discursive devices s/he uses accordingly result from his or her relationships with others within a community of practice (Sunderland and Litosseliti, 2002), e.g. Facebook is a virtual community of practice.

The term discourse, therefore, is used to refer to language beyond the sentence used in context to communicate something by way of being aware of the culturally appropriate means to do so (Sunderland and Litosseliti, 2002; Paltridge, 2006). Under the social constructionist approach, it is through the use of language, either written or spoken, that certain social and gendered identities are displayed, performed, and created. In other words, what people say contributes to the construction of views of the world, including who they are, and how they want

¹ Sunderland and Litosseliti (2002) mention that a person's identity, like face, is never static, it rather consists of a continuous decision making about oneself and one's lifestyle.

to be seen, i.e. face. Taking this into consideration, discourse analysis looks at the relationship between language and the contexts in which it is used, by describing and analyzing spoken and written interactions, to understand how speakers manage these interactions within particular groups (Johnstone, 2002; Paltridge, 2006; Sunderland and Litosseliti, 2002).

In this context, discourse analysis, together with corpus linguistics, will be used to fulfill the objectives of this study, which is thoroughly described in the following sections.

3.2 Overview of the study

The present study aimed at 1) identifying which were the most common politically-related insults used on Facebook before, during, and after the 2012 Mexican presidential election, 2) resolving the morphological processes by which such insults came about, and 3) determining the meaning and the pragmatic implications of such insults. The latter was done by means of two corpora created with the comments made by Facebook users on pictures posted on four Facebook pages that were openly against two of the most controversial candidates of the 2012 Mexican presidential elections: Enrique Peña Nieto and Andrés Manuel López Obrador.

The pictures were collected by accessing each of the four Facebook pages on every Saturday for almost a year. Each picture was labeled and saved depending on the candidate that each Facebook page was intended to attack or criticize. The comments made to each picture were copied and pasted into a word document so that they could be entered into the Text Stat program for the creation of the corpora. The word document was given the same label as the picture from which the comments were taken. The quantitative and qualitative analyses were then carried out by determining the frequency of the insults and the linguistic context in which they appeared as well as the morphological processes that gave way to them in order to determine the pragmatic implications behind their morphological structure. Further and more in depth explanations of the methodological steps taken are provided in the sections that come next.

3.3 Data collection

Four Facebook pages were chosen in July 2012 because they appeared to be some of the most active ones during that year's electoral process. This was noticed by the researcher because many of the politically-related pictures shared by her Facebook friends could be traced back to these four Facebook pages, among others. It is hard to tell how many Facebook pages were created due to the 2012 electoral process. Hence, another important factor in determining why these four pages were chosen from among the immensity of pages there might be was that each had more than 50,000 followers (see table 3 to know the exact number of followers that each page had) and because they appeared to be openly against one of the most controversial candidates in this electoral process either Enrique Peña Nieto or Andrés Manuel López Obrador. As a result, the pages were labeled as AMLO1, AMLO 2, EPN1, and EPN2. The former were entirely against Andrés Manuel López Obrador, while the other two were completely anti-Enrique Peña Nieto, this can be inferred from the pages' names (Table 3).

Table 3

Description of the Facebook pages accessed for data collection purposes

Facebook page name	Political tendency	Label	No. of followers
Eres más Falso que las		EPN1	700,145
Propuestas de Peña Nieto	Right wing		
Razones para No votar por Peña Nieto en el 2012	Against Enrique Peña Nieto	EPN2	96,838
Por Los Que Creemos Q Lopez		AMLO1	70.746
Obrador Es En Pobre Pendejo frustrado	Left-centered wing		
No se apendeje, NO vote por el peje	Against Andrés Manuel López Obrador	AMLO2	306,002

Note: The page's names are spelled out exactly as they appear on Facebook.

The number of followers is indicated right below the page's name on Facebook.

It must be mentioned that in order to access the content posted and shared by the managers and users of any Facebook page, one must click on the LIKE button that can be found on the upper right side of the page. Once this was done, each page was accessed every Saturday² for nearly a year, starting on January 7th and ending on December 1st, to collect all the pictures that had been posted on Saturday of the week before. There were 48 dates in total, which, for accurate labeling purposes, were assigned a number from 1 through 48 following a chronological order, thus December 1st was date number 48. Since it is common that more than one picture be posted on the same day, pictures were also labeled to avoid confusions. During the data collection process, a chart was filled for every Facebook page (see figure 1). On the first column of the chart is the date number, followed by the date itself in the column that follows to the right. The third column was used to enter the times at which each picture was posted so that the activity of the page could be more accurately described in the chapter that follows. Finally, the last column was filled with the pictures' label followed underneath by the picture's hyperlink so that it could be accessed in the future. In case no picture was posted on a given date, this latter was marked in red as in the example shown in figure 1.

² Saturdays were chosen for data collection because many events that are relevant in the electoral process, e.g. presidential debate, took place on Saturday. In addition, since it is not a working or school day, it is more likely that Facebook users will spend more time sharing all sorts of content, especially during electoral times.

RAZONES PARA NO VOTAR POR PENANIETO – EPN2
<https://www.facebook.com/pages/Razones-para-NO-votar-por-Pe%C3%B1a-Nieto-en-el-2012/233518030001539?fref=ts>

	Fecha	Hora	Link y etiqueta
1	7 de enero de 2012		1.1.EPN2 1.2.EPN2
2	14 de enero de 2012		
3	21 de enero de 2012	4:30pm 9:55pm 9:56pm	3.1.EPN2 https://www.facebook.com/photo.php?fbid=342125605807447&set=pb.233518030001539-2207520000.1363406701&type=3&theater 3.2.EPN2 https://www.facebook.com/photo.php?fbid=342268405793167&set=pb.233518030001539-2207520000.1363406701&type=3&theater 3.3.EPN2 https://www.facebook.com/photo.php?fbid=342269005793107&set=pb.233518030001539-2207520000.1363406701&type=3&theater
4	28 de enero de 2012	12:44pm 12:48pm	4.1.EPN2 https://www.facebook.com/photo.php?fbid=347071298646211&set=pb.233518030001539-2207520000.1363406701&type=3&theater 4.2.EPN2

Figure 1
Example of data collection chart

All the pictures posted on each date were saved and labeled in the following way: **X.Y.FBPageLabel.Z**, where X stands for the number assigned to each date; since in some cases more than one picture was posted on the same date, Y stands for the position in which each particular picture was posted in relation to other posted on the same date. X and Y are followed by the label of the Facebook page from which the picture was taken, which can be AMLO1 or EPN2. Accordingly, Z stands for the last digit of the label assigned to each Facebook page, so Z can be either 1 or 2 depending on the case. All pictures downloaded and saved in a JPG format.

Once all pictures had been downloaded and properly labeled, saved, and registered in the data collection chart (figure 1), an online random number generator (Stattek.com) was used to determine which pictures out of all those that were collected would actually be used for the next phase in the data collection process: the creation of the corpora. As can be noted, the pictures' label starts with two digits (X, Y), which can go from 1.1 as in the example in figure to 48.100. These were used as a match with the random numbers generated, which were a total of

50 in order to get at least ten matches per Facebook page and thus 40 pictures. Nonetheless, little matches occurred between the pages labeled as AMLO1 and EPN1.

In the first case (AMLO1) this was not very surprising because it was the Facebook page where the least number of pictures (only 19) were posted throughout the entire year. Only three matches coincided between the random numbers generated and the pictures belonging to the AMLO1 and EPN1 pages. Thus, although it was initially planned that a total of 10 pictures per page would be used, the researcher decided to use the three matches for these two pages plus fifteen matches that occurred with the other two pages respectively. Therefore, out of 620 pictures, the comments made to 36 of them were copied and pasted into a word document which was saved under the label of the picture to which the comments belonged so that it could be then entered into the TextStat program for the creation of the corpora.

Two corpora were created. One was labeled as AMLO corpus labeled and the other as EPN corpus. This has a type- token ratio of 8363: 50,059, whilst the former had a type-token ratio of 22,650:223,072. After the documents were uploaded and added to create each corpus, a frequency and collocation list was obtained by means of the TextStat program itself. Then, in order to continue with the quantitative and later the qualitative analyses, the researcher looked for some of the insults that she had previously seen and from which she came out with the idea developing the present project and sorted them in table 4, which can be seen in chapter four.

3.4 Data analysis

3.4.1 Quantitative analysis. Once both the pictures and the comments made on them were properly saved as was already indicated, an online random number generator (stattrek.com) was used in order to determine which pictures and their respective comments would be used for the analysis in an unbiased way. Since a table of random numbers requires that items be labeled with the same number of digits, the X and Y values were used in order to

select the pictures that would be used³. Throughout almost a year, a total of 620 pictures were collected from all four Facebook pages, however, only 36 pictures were actually used. There was no previous expectation regarding the total number of pictures that would be gathered. Nonetheless, it was decided that ten pictures per Facebook page, hence a total of 40 pictures, would be a representative sample. However, there were only 36 coincidences with the random numbers generated. Thus, only 36 pictures were actually used for the present study. It must be noted that the Facebook pages that posted the largest number of pictures had 15 coincidences, while those that posted less pictures only had 3 coincidences with the random numbers generated. After the pictures were selected, the word documents that corresponded to each of them were entered into the Text Stat program to find out the frequency of the insults and their collocations, and subsequently carry out the morphological and pragmatic analyses, i.e. discourse analysis, which are further explained in the following section of this chapter.

3.4.2 Qualitative analysis. Since the vocabularies of languages are thought to reflect the knowledge, values, and interests of a speech community, giving way to the theory of scripts (Chilton and Schäffner, 1997), the qualitative analysis is intended at determining the political knowledge, values and interests of those against Peña Nieto and those against Andrés Manuel López Obrador that are encapsulated within the morphological units that make up each insult by first identifying the word formation processes behind them and later inferring their pragmatic implications by means of analyzing the linguistic context in which they occur.

3.5 Final comments

The results obtained from both quantitative and qualitative analyses are included in the following chapter, while the actual explanation of the morphological and pragmatic analysis is provided in the last chapter. The pictures used for the study can be found in the appendices section, at the very end of the document.

³ The period in between them was not taken into account at this point of the data collection process.