[Paper presented at the International Symposium *The Lower Classes, Scripturality, and the History of Language. An Interdisciplinary Balance*, Christian-Albrechts-Universität, Kiel (Germany) 6 and 7 November 2014]

Under-privileged social actors: their texts, their language, their history[1]

Rita Marquilhas

Universidade de Lisboa, Centro de Linguística

I would like to start thanking the organizers both for the invitation and for the careful preparation of the Symposium. The selection of the principles and questions you listed in order for us to prepare our interventions is a precious guide, indeed. In accordance, I will keep track of them all along my presentation and I will use them both as backbone structure and as pretext for further remarks.

- *What was the actual state of alphabetisation and literacy in the lower classes in the modern era?*
- *What connections are there between the (written) textual forms of the upper classes and the lower classes?*
- *Which language models do the lower classes use for orientation?*
- *Are there independent developments in the lower classes, or are the linguistic forms to be regarded as imperfect acquisitions from the upper class ("Gesunkenes Kulturgut")?*
- *What impact did the development of media have (presence and availability of books, paper, postal services...)?*
- *What sources for the written language use by the lower class can be found?*
- *Which edition techniques are useful for lower class texts (principles of diplomatic transliteration, methods of digital indexing, xml-editions, tagging, etc.)?* Colloquium site (http://www.scripturalite.romanistik.uni-kiel.de/de/english-1)

-
- I will start with the observation about the need for cross-disciplinary work when approaching the subject of the language history and lower classes history. I agree of course with this remark made by the organizers, but having practiced cross-disciplinary studies that match history, sociology and linguistics for some years, I would like to add some remarks of my own about this research behaviour.

When specialists spot the need for cross-disciplinarity in order to get deep into an object of research, there are two things to do next: First) put more than one discipline into practice, either singlehanded or joining forces with colleagues who possess other competences, and Second) never forget why the object needed to be approached by more than one field of knowledge in the first

---

place. We can only prove the benefits of a cross-disciplinary adventure if we manage to show that the complexity of our object gets somehow less mysterious. Because we cannot ignore that there is a lot to lose, also, when abandoning the logic of single disciplinary approaches. Less academics will follow us, for one thing. Very specific terminology has to be partly abandoned because broader terms have to be used. Clear-cut principles have to be embraced less eagerly too, and that can cost us the respect of our peers. (For instance, the principle of the innateness of language has to be combined with the principle of language as a social object). In order to keep the balance between explanatory gains and academic respect, the only way is to keep going back to the research object and thinking again and again why it is so complex.

To get concrete here, let's look straight into the Western lower classes and their language as an historical object of inquiry. Lower classes are a social category, so we need to know, first, the principles governing of social life independently of history and, second, the idiosyncrasies of social 'specimens' in the chronology we are studying. On the other hand, the lower classes' language, like all linguistic behaviour, is neither independent of its speakers nor is it independent of the universal ways language works. This is not being too ambitious, trying to grasp it all about the history of lower classes and their language, because experience proves that we get short of answers if we only know principles of social life, or details of historical societies, or Pragmatic explanations to speakers performances, or generative versus typological explanations of language universals.

That is why, looking closer to the seven questions that the organizers prepared for us, one can start by identifying the several oppositions that are at stake and the different fields of knowledge that are at ease with the successive dichotomies: literate versus illiterate, up in society versus low in society, spoken language versus written language, models for writing versus actual written texts, developed language variants versus acquired variants, presence of media versus absence, existence of written sources versus lack of them, useful editorial techniques for those sources versus less useful ones.

-Literate social strata versus illiterate strata. History of education and cultural history grasp this opposition best BUT the historians have to rely a lot on descriptive statements from social actors of the chronology in question, typically literate actors and typically dominant ones practicing what was called in the Early Modern ages the political (polite) discourse; to decide whether it was accurate or not in what concerns literacy, performance of signatures can be measured by the historian BUT it is always difficult to establish how reliable those measures are.

The voices that embodied the dominant discourse in Portugal in the 16th and 17th century were very well studied in a series of essays by the cultural historian Diogo Ramada Curto. They played the game of an aristocratic taste for debating language issues and they were constantly concerned with the distance between the language performance of *the noble compared to the plebeian, the well-advised compared to the indiscreet, the vicious compared to the well-educated* (Duarte Nunes de Leão, 1576, apud Curto2007: 69). Dante's De Vulgari Eloquentia had had a great effect, so it was fashionable all across Europe to debate languages' merits and to draw general conclusions from variation among speakers. But alongside the literal meaning of the elite discourse of the time, the

Social Scientist can search there too for the unconscious presence of the notion of cultural capital and distinction, as defined by Pierre Bourdieu.

The Royal court, above all, was the context for embodying cultural capital and become socially distinguished. No accumulation of study could compensate for the lack of this socialisation opportunity.

As for literacy rates in Portuguese history, one can start by the first national population census that included a question on the ability to read. In 1878, almost 80% of the population over 6 years old could not read (Ramos 1988: 1067). This squares right with Jack Goody's concept of restricted literacy in traditional societies (Goody 1968), but a systematic study of the ability of the Portuguese to sign depositions before the court in the 17th century, two hundred years before, gives us a much different picture of literacy in the same country (Marquilhas 2000). The overall Portuguese population, urban and rural, would be as literate as the urban Spanish one, as measured by Marie-Christine Rodríguez and Bartolomé Bennassar (1978). On the other hand, the comparison with British literacy, as measured in Norwich by David Cressy for the same chronology, gives us the rural northern Portugal with a lower literacy rate than the British, but the urban southern Portugal either with a similar rate or a superior one, in Lisbon.

Social-occupational category	Rural north	Center (Lisbon)	Urban south
clergymen, monks and nuns	2%	99,0%	100%
upper nobility, servants to powerful families, clerks, students, professionals, upper administrative staff, upper tradesmen, civil officials of the Inquisition	87,5%	92,3%	93,7%
lower traders, artisans, pilots, sailors, suburban artisans	51,6%	58,9%	51,6%
farmers and landowners living by the rent, local nobility, citizens, upper officials in local administration	48,3%	68,0%	66,0%
common servants, labourers, apprentices, soldiers, fishermen, slaves, beggars	15,7%	30,7%	21,6%
unspecified socio-professional category	34,1%	54,4%	32,4%
Average	60,1%		

David Cressy wrote, in 1980, that 'without the push of a protestant literacy campaign, without the demands of a complex and dynamic economy, it was only natural that the larger part of Mediterranean Europe would remain illiterate". More than an assumption, based on Max Weber's well-known explanation for the rise of capitalism on the roots of Calvinism, this proved to be an oversimplification. Since writing provides an alternate solution for face-to-face interaction, every

dynamic society in terms of population displacement has good reasons to depend on writing as a functional support to maintain constant activity between its social networks' nodes. The explanation for the comparatively high levels of literacy in Portugal in the 17th century can be plural, but let me remark that, at the time, the population was scarce (less than two million people by 1640), the capital – Lisbon – was a large port and economically vibrant, thus a constant destination for rural emigrants. At the same time, the struggle to maintain an overseas Empire required soldiers, sailors and settlers. All these forces working together forced people, especially men in the rural areas, away from their original social groups, and communication by writing became one of the needs in order to survive, at least in emotional terms. Two hundred years later, when the census of 1878 counted almost 80% of illiterate Portuguese, the Asian and the Brazilian empires were gone and the African was still to be explored. The population had more than doubled since the 17th century, from 1.9 million in 1640 to 4.7 million in 1878. And the literacy campaigns led by the government were still about to start fully. One hundred years later, in 1960, only 17% of the population was illiterate.

I took too long to deal with the first opposition, between the literate versus the illiterate in Early Modern Portugal, so I will go faster now.

Because I want to show some of the methods we are using in Lisbon in order to find, publish and study the writings of the everyday man in the Early Modern age, I will only recover one more opposition from all the seven I listed earlier. This is the opposition between written and spoken language, a dichotomy that does not match, regardless of appearances, the opposition between upper class and lower class varieties of language. Sociolinguistics goes to a great length in trying to prove how artificial, static and non natural written language is. With not so many results as it would be desirable. The appearance of natural language projected by written texts blinds even linguists, who should be more equipped to watch an mirroring phenomenon. There is a recent title on that subject, published this year, so allow me to cite the authors on a rather long quotation. I bring it here because I find this is an essential point to be taken before studying the apparent natural language kept in historical (and synchronical, for that matter) popular writings:

One of the main points of modern linguistic research is that spontaneous spoken languages are not a defective or imperfect implementation of written standard languages. These latter languages are based on cultural elaborations of the former and they are to a lesser or greater extent artificial. This explains why non-educated speakers make mistakes when trying to speak a written standard language: they tend to naturalize that artificial language in order to make it adequate for normal everyday linguistic interaction, introducing a degree of variation and the under-determination typical of natural languages, but which is not permitted in cultivated languages. Cultivated languages can only be usefully described and studied when the natural languages on which they are based are taken as a reference. On the contrary, natural languages can be studied in themselves, without reference to their elaborated versions. Of course, in communities with a standard language tradition, some influence of the standard language on the spoken natural language can be expected; but this influence is superficial [...] and does not affect

any essential features of the corresponding natural language (Moreno & Mendívil Giró 2014: 153).

This is a creed of much importance, because it shows how relevant to language history are the writings of lower class speakers of the past, the ones that for lack of training with the tool of written language, are more immune than the fully literate to the artificiality of the practice and more capable of being faithful in the recording of spoken language feature. Nevertheless, we should always be aware of the fact that it is their culture above all that we will be able to find there. Grammar features should be hypothesized beforehand, and only then tested against the data of the corpus, not the other way around.

As for the collection of Portuguese writings by the Lower Classes in history, we have a good idea of what a representative sample of Portuguese informants from the Early Modern period would be like because we know the changing volumes of Portuguese population between the 16th and the early 19th century:

Years	Mainland Portugal	Islands	Total
1530	1,120,000		1,120,000
1580	1,200,000		00,000
1620	1,200,000		1,200,000
1640	1,900,000		1,900,000
1700	2,050,000	125,?97	2,175,897
1732	2,143,?8	159,?	2,303,289
1768	2,?98		2,409,698
1801	2,931,9?	?3,400	3,215,330
1821	026,450		3,026,450
1835	3,?684		3,061,684

Excerpt from a chronologically broader chart published in Rodrigues (2009: 519)

The historian of language in society may combine these types of statistics with the criterion, suggested by Labov, of collecting linguistic samples that correspond to at least 0.025% of the population of speakers under investigation (Labov 1966; Hernández-Campoy and Schilling 2012: 67). The numbers required to obtain a sufficiently representative sample of linguistic informants, divided into reasonably narrow periods of 50 years, are by no means unattainable, as we can see in the next table:

Time interval	Number of necessary informants
1551-1600	300
1601-1650	470
1651-1700	500
1701-1750	530
1751-1800	600
1801-1834	780
Total	3,180

Ideal sample in terms of validity to represent European Portuguese spoken in the early Modern period

That is to say, in order to achieve this ideal sample, it would be necessary to track down around 3,200 reliable informants for whom there are informal records that are mostly conversational in nature but have been documented in writing. These do in fact exist. As was researched between 2008 and 2010 by the *CARDS, Unknown Letters* project (Marquilhas 2012), and now with the *P.S. Post Scriptum* project that has succeeded it (Vaamonde et al. 2013), there are unpublished but contextualized Portuguese documents conserved in sufficient number to enable the constitution of that historical sociolinguistic sample.

The *P.S. Post Scriptum* project locates, publishes and studies documents of private interaction, namely private letters in Portuguese and Spanish. The documents are extraordinarily rich sources for the history of the language as spoken on an everyday basis from the 16th to early 19th centuries. Most were addressed to relatives, friends, sweethearts or mere acquaintances, and were archived by the civil and ecclesiastic courts, together with documentation that contextualizes them. This documentation also sheds light on the sociological profiles of the writers or recipients, as what interested the courts was the incriminatory value of the letters. The discourse used in them comes from very diverse social strata: they were written by and to men, women and even children, the powerful and the disempowered, from all parts of Portugal, and its overseas territories. In some cases, we know the life stories of these people, despite their modest origins, through the testimony of defendants or witnesses.

Having identified the documents in question the *P.S. Post Scriptum* team makes use of all possible innovations in philology, the digital humanities and corpus linguistics in order to grasp their cultural, social and linguistic significance. Consequently, the treatment of these original manuscripts has to begin with diplomatic transcription, which gives rise to a genetic critical edition that reconstructs the writing process. It uses the most universal format in terms of digital

publishing (XML-TEI) and feeds a biographical database (of authors and addressees) whose information may be compared with the textual contents of the letters that they wrote or received, and with the communicative situation in which they participated. Using automatic tools, the word strings of the diplomatic transcription are broken up into graphic units – tokens – aligned one by one with modern orthographic words, their lemmas and more abstract categories, both morphosyntactic and lexicosemantic. Thus the scholarly digital edition becomes a historical corpus.

This is, in short, the programme of the *P.S. Post Scriptum* project, which involves linguists, historians and systems engineers, and which will be completed in 2017. From the perspective of historical sociolinguistics, the project is interesting because of what it might yield about the internal and external factors of variation and change over four centuries of Portuguese and Castillian.

References

Cressy, David, 1980, *Literacy and the Social Order. Reading and Writing in Tudor and Stuart England.* Cambridge, Cambridge University Press.

Curto, Diogo Ramada. 2007. *Cultura escrita (séculos XV a XVIII).* Lisboa: Imprensa de Ciências Sociais.

Curto, Diogo Ramada. 1988. *O discurso político em Portugal (1600-1650).* Temas de cultura portuguesa 12. Lisboa: Projecto Univ. Aberta. Centro de Estudos de História e Cultura Portuguesa.

Curto, Diogo Ramada. 201_. *A cultura política no tempo dos Filipes (1580-1640).* Lugar da história 78. Lisboa: Edições 70.

Leão, Duarte Nunes de. 1576. *Orthographia da lingoa portuguesa.* Em Lisboa: per Ioão de Barreira.

Marquilhas, Rita. 2000. *A Faculdade Das Letras: Leitura E Escrita Em Portugal No Século XVII.* Lisboa: Imprensa Nacional-Casa da Moeda.

Marquilhas, Rita. 2012. "A Historical Digital Archive of Portuguese Letters." In *Letter Writing in Late Modern Europe*, edited by Marina Dossena and Gabriella Del Lungo Camiciotti, 31–43. Amsterdam: John Benjamins.

Moreno, Juan-Carlos, and José-Luis Mendívil-Giró. 2014. *On Biology, History and Culture in Human Language: A Critical Overview.* Sheffield & Bristol: Equinox.

Ramos, Rui. 1988. "Culturas De Alfabetização e Culturas Do Analfabetismo Em Portugal: Uma Introdução à História Da Alfabetização No Portugal Contemporâneo." *Análise Social* 24 (103-104): 1067–1145.

Rodríguez, Marie-Christine e Bennassar, Bartolomé, 1978, «Signatures et niveau culturels des témoins et accusés dans les procès d'inquisition du ressort du Tribunal de Tolède (1525-

1817) et du ressort du Tribunal de Cordoue (1595-1632)», *Caravelle. Cahiers du Monde Hispanique et Luso-Brésilien*, 31. Toulouse, Université de Toulouse-Le Mirail, pp. 19-46.

Rodrigues, Teresa Ferreira, ed. 2009. *História Da População Portuguesa: Das Longas Permanências À Conquista Da Modernidade*. Porto: CEPESE & Edições Afrontamento

Vaamonde, Gael, Rita Marquilhas, Ana Luísa Costa, and Clara Pinto. 2014. "Post Scriptum: Archivo Digital de Escritura Cotidiana." In *Humanidades Digitales: Desafíos, Logros Y Perspectivas de Futuro*, edited by Sagrario López Poza and Nieves Pena Sueiro, 473–82. Janus [en línea], Anexo 1.