

REVIEW

Mikel Martínez-Areta

University of the Basque Country

BAUER, Roland (2009) *Dialektometrische Einsichten. Sprachklassifikatorische Oberflächenmuster und Tiefenstrukturen im lombardo-venedischen Dialektraum und in der Rätoromania*, Ladinia monografica 01, Istitut Ladin Micurà de Rü, 419 p.

0. Introduction

On the following pages, I shall make some few comments about Roland Bauer's recently published *Dialektrometrische Einsichten* (2009). Section 1 is a brief description of the structure of the book and the formal aspects of it. Section 2 is a concise discussion of the dialectometric field. Section 3 is a concluding recapitulation. It may be worth pointing out beforehand that I am a Basque historical linguist and philologist acquainted neither with dialectometric studies nor with modern Romance linguistics. The views presented here, then, will be those of someone from a neighbouring discipline. Although this may imply a handicap in several respects — for I can not competently evaluate most technical aspects presented in the work —, it might be also of some interest to listen to opinions and comments from other (but related) fields.

1. Roland Bauer's *Dialektrometrische Einsichten* (2009)

Had I to define the work as a whole with only one adjective, I think I would choose “exhaustive”. From the very beginning, the author aims at a thorough description of the history of the dialectometric field trying not to forget a single contribution, at least as far as Lombardo-Venetian and Reto-Romance linguistics is concerned. On this line, after some introducing remarks in Chapter 1, in Chapter 2 he goes on to make a thorough history of dialectometry as a discipline. Even if the term arises in 1973 with the Southern French dialectologist Jean Séguy, the main stream of

dialectometric studies has been that of the Austrian School, whose central figure is the Romanist Hans Goebel. The bulk of the chapter actually focuses on Goebel's academic life and evolution, as well as on other dialectometric contributions and projects that have been to some degree inspired by him.

With Goebel's work, electronic processing of dialectological data is used systematically in the analysis for the first time. Indeed, it is his early work that inspired the classical definition of dialectometry as 'dialectal geography + numerical taxonomy', explicitly stated as such in an article by Goebel in 1980.

It seems to me that the rise and evolution of the dialectometric discipline has to be understood in the frame of the rise of several other disciplines —or schools inside an existing discipline— that arose and developed in academic Europe and North America during the 2nd half of the 20th century. Very often, these disciplines have run parallel to each other and have tended to increase mutual interaction over time. Dialectometry has gone the same way. It came about as a discipline sister to other disciplines which sought taxonomical means to measure elements and relations, such as econometry, sociometry, and so on. Likewise, in recent years some articles by Goebel have looked for points of convergence with human population genetics (see p. 50) and geo-onomastics (see p. 54), among others.

Chapter 3 (pp. 87-157) is a discussion about the methodology used in dialectometrical research. The general process to follow is sketched in Figure 1 of page 88. In my humble opinion, in this part of the work Bauer prioritises technical accuracy over clarity of exposition. The presentation of the whole chapter in purely abstract terms prevents any reader not familiarised with taxonomic procedures at top level from approaching the text with some chance to understand the basic ideas. In fact, the author warns in the introduction (p. XI) that the book is targeted mainly at a specialised public, in any case one acquainted with taxonomic devices such as algorithms, matrixes, feature vectors and so on. Thence, he cannot be accused of cheating anybody. What I suggest is that it could have been relatively easy to attract some more readers and to make the dissertation more divulgative if he had just specified more clearly and by means of linguistic data what he exactly refers to by *taxandums*, *taxates*, etc.; if he had put e.g. concrete examples of how he elaborates a *similarity matrix* or a *distance matrix*; and in short, if he had to some degree striven to bring the unspecialised reader and the data of linguistic reality closer to each other.

Concrete linguistic data start to show up by Chapter 4 (p. 158), but by then it has been made perfectly clear that the addressee of the book is planned for a reduced selection of initiated. Even though, from page 164 on the unspecialised reading recovers some interest and this increases in several passages of the remaining chapters. Nonetheless, by the end of the book, some of the questions that I make myself when I think how I should make a dialectometric analysis, if I were a dialectometric scholar, remain unanswered. This is the main point of my critic. What remains unclear to me is how the jump from stage A to stage B in Figure 1 (p. 88) is performed. Apart from the general objections that I present in Section 2, I would have plenty of doubts as for how I should classify a great deal of features. To set forth but one,¹ suppose that a German variety has generalised the regularisation of the verb *schaffen* (*schaffen* / *schaffte* / *geschafft*) but another neighbouring one still preserves the irregular formation in some semantic contexts (*schaffen* / *schuf* / *geschaffen*), but that this happens only to this verb. Should these taxats go into a morphological matrix or into a lexical one? I would have some doubt in such a case and many similar ones, and have not found an answer as for how to tackle them. To put it briefly, whereas taxonomic arguments are exhaustive all over the book, the purely linguistic ones leave some important methodological points unanswered.

2. Dialectometry and modern linguistics

I go on now to briefly discuss dialectometry from a more general point of view. The idea of performing taxonomic measurings of linguistic elements and relations that help us to predict branching chronologies among languages of a linguistic family,² grades of dialectological similarity, and such things, is somehow reminiscent of the controversy about the lexico-statistical method (also called glottochronology). Indeed, in my opinion dialectometric analysis shares at least some problems with the lexico-statistical method. In particular, neither of them has indisputably shown that there is a single and universally accepted mechanism to process data. In lexico-statistical analysis,

¹ I take my example from German, for I do not feel confident with the Romance data used by Bauer.

² This has been the traditional task of the lexico-statistical method, but diachronic studies have also been undertaken by scholars of the dialectometric discipline, like that of Dees (see Bauer 2009: pp. 57 ff.).

the results of practical exercises have been most often coincident with the analyser's previous position towards the method. If this was positive, the result of the exercise was more or less reasonable according to traditional analyses.³ If negative, the application of the lexico-statistical method was a complete failure far away from linguistic reality.⁴

Since I am not acquainted with Lombardo-Venetian and Reto-Romance dialectology, I shall resort to Basque in order to show several points of my criticism. In standard Basque, and in all central dialects, there is a morpheme *-te* which pluralises the 3rd person agent marker of transitive finite verbs (*du* 's/he has' / *du-te* 'they have', *ematen di-zu* 's/he gives you' / *ematen di-zu-te* 'they give you'). This appears as *-e* in a number of western and eastern areas (Biscayan *dau* / *dab-e* < *dau-e*, Souletin *dü* / *di-e* < *dü-e*). We can pose, then, a triple division *-e* / *-te* / *-e*. A formally and diachronically related morpheme is the pluraliser of the 2nd person agent marker, which is *-e* in most dialects (*du-zu* 'you (sg.) have' / *du-zu-e* 'you (pl.) have'), but *-te* in some few central ones (*du-zu-te* 'you (pl.) have'). The dialectal division here is also *-e* / *-te* / *-e*, although here the central variant is reduced to a very small area. A third related pluraliser is the one of the 2nd person absolutive marker of intransitive finite forms (in standard Basque *zara* 'you (sg.) are' / *zare-te* 'you (pl.) are'), which shows up as *-e* in Biscay (*zari-e* 'you (pl.) are', < **zara-e*), and as *-de* in Souletin (*zira-de* 'you (pl.) are'). Here we have: *-e* / *-te* / *-de*, the central one being predominant.

Now, do we have one, two or three morphemes here?⁵ In other words, what is the unit to be considered as a 'feature'? Does a feature correspond to only one category or to several amalgamated ones? That is, should we consider that here we have just a finite verb pluraliser in general (one morpheme), or that any combination of categories implies a feature, so that we have a 2nd p. abs. marker, a 2nd p. erg. marker, and a 3rd p. erg. marker (three morphemes)?⁶ Who decides whether the feature unit is number, number/person, or number/person/diathesis?⁷

³ This is the case with Gray & Atkinson (2003), whose branching chronology of the Indo-European groups according to (their application of) the lexico-statistical method is rather classical, at least according to some schools.

⁴ A well-known case is Coseriu (1991 [1977]: 175-185), whose application of the lexico-statistical method would bring us to assume, for instance, that Italian Romance branched off Vulgar Latin in the 16th century.

⁵ In fact, the problem is more complex, but this is not the right place to discuss it in detail.

⁶ If we opt for this option, we might get into serious trouble in the analytic process. There is another morpheme *-it-*, which pluralises 1st, 2nd and 3rd person absolutive morphemes, although it is redundant in the 1st and the 2nd: *ga-it-u* 's/he has us', *za-it-u* 's/he has you (sg., but historically pl.)', *za-it-u-z-te* 's/he has you (pl.)', *d-it-u* 's/he has them'. The simplest option would be to consider this as a single

Even if we came to an agreement about what should be considered as a feature, do all features have the same level of representation in the data matrix of dialectometric analyses, as I assume is the ultimate implication of the ‘adansonian’ or ‘isocratic principle’ (p. 92)? Should this mean that a feature that is uttered only once every two days by a standard speaker of a particular variety counts the same as another one which is uttered twenty times per day in the taxonomy of dialectal differentiation? Should the functional load or productivity of particular elements (whether morphemes in morphological matrixes, words in lexical matrixes, and so on) not be pondered in dialectometric analysis?⁸ If it should, who decides how this functional load should be taxonomised?⁹ I think I would prefer to admit the impossibility of taxonomising all these data, even under the risk of being accused of remaining stagnant in old ideas and methodologies, rather than make arbitrary decisions.

In spite of this criticism, my judgement of dialectometry is in general terms positive. Dialectometry has, at least, an extremely significant advantage over the lexico-statistical method:¹⁰ data are collected by people who are specialists in the languages and dialects they belong to. That implies that the collecting process will be accurate and taxates will be usually rid of those mistakes that are so frequent in lexico-statistical samples.

In fact, whether one agrees with the way dialectometric scholars interpret data or not, the massive collection of data gathered by them is a huge contribution to dialectology in general. If we interpret the data collected as the means and the

morpheme, but if we have decided that the feature unit should correspond not to the number category alone (sg./pl.), but to the number-person complex (1st p. sg. / 2nd p. sg. / 3rd p. sg. / 1st p. pl. / 2nd p. pl. / 3rd p. pl.), then we should be coherent with this decision and to split *-it-* into three (or four) morphemes.

⁷ That is, whether the *-te* (~ *-e* ~ *-de*) of *zare-te* should be counted as a single feature or included into a general group of finite verb pluralisers (along with *du-zu-e* and *du-te*), or as a third option into a group of finite verb 2nd p. finite verb pluralisers (along with *du-zu-e* but not with *du-te*).

⁸ As a matter of fact, the isocratic principle has been used not only by dialectometric scholars. At least in Basque dialectology, it is also typical among scholars working with traditional methods, in order to count convergent and divergent features when presenting the final balance of the position of a particular dialect in a dialectal spectrum.

⁹ I am not aware of having read any explanation for these doubts in the whole book. My main objection, as I said before, is that it does not explain the process of collection of data and its sorting out into different matrixes as accurately as it does the taxonomisation of data.

¹⁰ As a matter of fact, I think that linguistics should not even consider the lexico-statistical question as a settled debate, concluding that the method is absolutely useless and should be definitively given up. Such a clear-minded linguist as Szemerényi (1986 [1982]: 112) e.g. does not, and prefers to give the method a chance under the condition that it be subjected to constant criticism, even by its defenders. It is far away from the scientific spirit to take controversies for definitively closed.

dialectometric taxonomisation as the ends (or aim to be reached), one can hardly think of any realm where one can so indisputably state that “the ends justify the means”. Another question is whether the means justify the ends. To judge from the results of the dialectometrical papers that are most familiar and linguistically close to me —those about Basque dialectology, like e.g. Aurrekoetxea (1992, 2004) and Aurrekoetxea & Videgain (2007)—, the conclusions drawn are not at odds with linguistic reality as examined with traditional methods. In fact, they make plenty of sense as for the similarity level among Basque dialects. Now, since the debate about dialect classification among Basque varieties has been extremely active in recent years, I wonder if the dialectologists who, working with traditional methods, have proposed dialectal classifications different from the classical one and different from Aurrekoetxea’s, would reach the same conclusions as they now do, if they applied dialectometric technics. I suspect that they would, for neither the way of collecting the data nor the way of processing them are universal, as there is always margin to interpret elements in one’s own interest.

If this were so, a comparison more proper than the one between dialectometric and lexico-statistical methods would be that between dialectometry and surveys before political elections. It is usually said that they often fail, but that failure is relative. The range in which they oscillate has certain limits, and our surprise when we see a somewhat unexpected result does not usually reach the level of astonishment. Obviously, the result of the survey is often influenced by the ideological tendency of the institution that makes it, but this should never go too far from objective reality, if it does not want to lose credibility. My suspicion is that dialectometrical analysis can be something similar. The same object of study can offer divergent results, but inside certain limits. It is therefore useful as a taxonomic device, but as a method it has to be exposed to constant criticism.

3. Conclusion

To sum up, Bauer’s *Dialektrmetrische Einsichten* is an exhaustive and updated compilation of everything that is to be said about the dialectometric discipline today, taking Lombardo-Venetian and Reto-Romance dialectology as a testing ground.

Although unfortunately it does not include among its tasks to spread the main ideas and conclusions reached by the dialectometric discipline to other linguistic realms, it is obvious that this work is a milestone in dialectometric studies.