

[0810] [Besprek] Wilbert Heeringa, *Measuring Dialect Pronunciation Differences using Levenshtein Distance. Groningen (2004), Groningen Dissertations in Linguistics 46, 315 pp., ISSN 0928-0030.*

This thesis is a dialectometric study the aim of which is to find a tool to measure “[...] linguistic distances between any pair of dialects [...] in an objective way” (p. 5). The need for this tool arises from the awareness on the part of any dialect speaker “[...] that borders exist in the dialect landscape” (p.1). Not only dialect speakers but also dialectologists feel a need to classify dialects according to their observed, supposed or felt kinship.

Dialectometry tries to provide an objective and verifiable basis for this ongoing inclination of dialectologists to draw maps with language and dialect divisions. This is a very useful effort. Dialectometry applies to linguistics methods used in the natural sciences, like mathematics. Dialectal divisions are usually based upon the intuition and knowledge of dialectologists who have been studying a specific dialectal region for a long time. If a formula may be found that enables the user, basing his findings purely on linguistic data, to come to a similar conclusion as that arrived at by dialectological specialists, the dialectometrist will be able to make reliable dialect divisions of any dialectal area. Furthermore, there will be no need to study in detail all linguistic aspects of the dialects involved.

The key to the use of the found tool is its validation. Validation is a new item introduced into dialectometric research by Heeringa, one that represents a definite methodological improvement. So far, calculated distributions have been arrived at intuitively, by comparison with existent traditional dialect divisions. To be able to validate the new method, three questions must be answered:

- what is the reference dialect division?
- how reliable is your reference dataset?
- how big must the resemblance be between the reference dialect division and the computed division to yield a ‘validated’ result?

It must be stated here that the dialectometric part of the thesis looks very impressive (Chapters 3 - 6). The author has tested an enormous number of methods and tried a whole range of ways to convert linguistic data into numerical data. I must admit that I took this - rather technical - part of the thesis for what it was and I am convinced that the author did a thorough job here. What I am concerned with is the dialectological translation and application of the dialectometric tools. Therefore the two examples of application of Heeringa’s method, a dialect division of Norwegian and Dutch dialects, were very interesting to me. So I will go through the three questions posed previously, looking at the dialect divisions presented by Heeringa in Chapters 7, 8 and 9.

A. Reference dialect division

To be able to validate computed findings one has to be able to compare them with a

dialect division that is supposed to be correct. On p. 178 Heeringa mentions his golden standard - a divisional basis devised on the consensus of several dialectologists - which he elaborated in previous publications. He rejects this method here for several reasons, the main one being that traditional dialect divisions are discrete: they only mark differences and fail to say anything about “proximity of linguistic relationships” (p. 178).

Heeringa therefore introduces another validation criterion, the “perceptual distance”: “does one dialect differ strongly from another one?” Reading Heeringa’s text, it becomes clear that he considers the perceptual differences observed by dialect speakers to be the absolute truth about dialect differences.

The choice of this criterion is reasonable - for example, the arrow-method uses it - but in my opinion it is only one criterion, albeit a very important one. Another important criterion is, for example, the genetic relationship between dialects. The West, East and North Frisian dialects will never be grouped together by a method based on perceptual distance, although no linguist doubts their genetic kinship. Other dialectologists prefer the social point of view: upon what standard language does a dialect depend? And one might probably find even more criteria. So Heeringa’s method is useful when one is interested only in testing perceptual distances between dialects. Another point is that, as the title of the thesis makes clear, Heeringa is measuring pronunciation differences, including lexical differences (these are interpreted as very different pronunciations) but language comprises more than just these aspects.

B. Reliability of the reference dataset

In Chapter 7 Heeringa describes his validation method. He applies a perception experiment, executed by Charlotte Gooskens in Norway, involving 15 dialects scattered over a wide area. There are some remarks that I should like to make here.

In each of the 15 towns, a test group of 16 to 27 listeners were asked to judge the perceptual distance from their own speech on a scale of 1 to 10, 1 being identical, 10 most remote. The test group was also asked to judge a recording made in their own town. For some towns the score for an own dialect was 3.44, where a score of 1 was to be expected! In only 2 out of 15 places did all the listeners recognise their own dialect. This result does not inspire a sense of those judgments being very reliable. Leaving out the data for the places themselves - as Heeringa does - may be a methodological escape from this problem; it does not, however, solve practical questions as to the reliability of the experiment. Assuming that we *might* be dealing with dialect differences within one village - as Heeringa does - is an assumption that in my opinion needs verification (par. 7.4.2).

Another problem, in my eyes, is the number of samples and the geographical density. To make my point clear, I will give an extreme example. If you compare only English, German and Chinese, you don’t have to be a very smart dialectometrist because practically every method you apply will give you a correct ‘dialect’ division. The addition of more means that the method has to become more

critical but as long as the 'dialects' under comparison remain quite distinct any method will probably do very well.

The more closely dialects are related, the more sophisticated a method is required to be able to arrive at an outcome that fits with your linguistic or intuitive division. For example, a method suitable for classifying 100 different languages from all over the world may be not sensitive enough to classify 100 related dialects within one country. Heeringa tests the minimum number of data in the data sample needed to come to a reliable classification (p. 176) but we are left with the point of the geographical density and the linguistic vicinity of the data samples. So his validation of the dialectometric methods is valid only for data samples with a geographical density and linguistic variance compatible with the Norwegian data sample. In fact, he refers to this problem (p. 194) but fails to draw any methodological conclusion from it. From Fig. 7.7 we can see that the majority of differences between the Norwegian dialects range between 15% and 30%. Heeringa's rough division of Dutch dialects into 13 main groups lies at the same level of diversity, or even higher. At this level his method indeed achieves reliable outcomes. But at the lower level of details we are confronted with several disputable subdivisions.

C. The resemblance between the traditional Dutch dialect division and Heeringa's computed one

One very interesting and, for a Frisian dialectologist such as myself, verifiable case study is Heeringa's attempt to calculate dialect relationships and divisions for the languages and dialects spoken in the Netherlands and Flanders. In particular his map showing the dialectal continuum, Map 9.33, is impressive. The general divisions he obtains look more reliable and fit existing linguistic knowledge better than does Hoppenbrouwers FFM method (cf. my discussion in *Us Wurk* 51 (2002), pp. 159 - 162).

Heeringa was lucky to have at his disposal a dialect map of the Netherlands and Flanders, the one by Jo Daan mainly based on the same criteria as he used for validating his own methods, i.e. perceptual distance. So, in fact, his case study of a dialect division between 'Dutch' dialects offers another chance to compare computed dialect divisions and distances with a more traditionally perceived model.

When it comes to the actual dialectal division we are confronted with various severe problems. From a dialectometric division one expects a certain level of recognition of traditional patterns. In this respect the new method indeed reproduces the general image familiar from Jo Daan's map. But the value of dialectometric methods is that they increase the understanding of dialectal divisions such that they introduce new ideas concerning them. Where the computed division diverges from the traditional one, interesting linguistic arguments for this and new insight into dialectal relationships and the nature of the differences might be expected.

One interesting case is the relationship between the dialect of Breskens and the other dialects of Zeeland, which might be the consequence of intensified traffic enabled by the ferry between Vlissingen and Breskens (p. 263). But most of the deviations in Heeringa's division from the traditional dialect division are due either to differing transcription - the material used has various transcribers - or is of the nature of the Zevenbergen case, where a dialect in West-Brabant is linked up with Limburgish dialects - i.e. totally unexplicably linguistically.

With respect to the transcriber differences, we have to conclude that Heeringa's method is so sensitive to phonetic differences that in reality a number of borders owe their existence to transcriber differences (cf. Figures 9.3 to 9.7). This, on the one hand, pleads for the method. On the other, in a real world in which transcriber differences are hard to avoid, it also implies a handicap. Furthermore, we come across a problem that I already also observed in my discussion of the FFM by the Hoppenbrouwers brothers: that linguistically relatively unimportant, highly frequent features may heavily influence dialectal division. Heeringa mentions that one reason why the West Brabantish dialect of Steenberghe is classified as Limburgish may be the simple difference between [r] and [R], which is in the practise of the language totally irrelevant; [R] would be a 'Limburgish' feature. Taking into account also that the [R] in Limburgish is of rather recent origin (100 years) and is emerging at high speed throughout the whole of the Netherlands, the relation between dialectometric criteria and linguistic reality becomes doubtful.

The Stellingwerven case

Another case more close to my own field of experience, Frisian, relates to the situation of the dialects of the Stellingwerven. In any traditional dialect division the Stellingwerven dialects are assigned as Low Saxon. It is also known that of all Low Saxon dialects the Stellingwerven dialects show the strongest Frisian influence. In Heeringa's method the traditional area of Stellingwerven dialects is cut into two segments according to a transcriber difference. One part is assigned as 'Frisian mixed varieties', a subgroup of the Frisian group. The other part is assigned to the Overijsel group of Low Saxon dialects. In a multidimensional scaling plot presented in a lecture, the Stellingwerven dialects appeared more Low Saxon than Frisian. This already fits better with my linguistic intuition.

One phonetic feature that weighs heavily in this, and any other, method based on phonetic features is the fact that most of the Stellingwerven dialects have unvoiced [s] and [f] in anlaut, like Frisian and unlike most of the Low Saxon dialects. This relatively small difference (it will hardly influence mutual intelligibility) may cause a high score in the Levenshtein difference, as it did in the FFM method.

The multidimensional scaling of the data (9.5.1, p. 266 ff.) shows that the second dimension distinguishes Frisian dialects from the other dialects. The word with the strongest correlation with this second dimension is 'father', where Frisian differs lexically from most of the other dialects. And lexical differences cause a high Levenshtein difference. The Frisian loanword for 'father', *heit*, in the Stel-

lingwerven dialects underlines the special relations between them and Frisian, as ‘noted’ by Heeringa’s method.

So, in respect of the Stellingwerven case, the method is correctly sensitive for Frisian influence in some of the Stellingwerven dialects but the way features are counted leads to a wrong classification of the dialects. This classification is not wrong due to any suggestion that no relationship exists between Frisian and Stellingwerven dialects. It is misleading in that the classification as mainly a Frisian mixed variety rather than a Low Saxon dialect with some Frisian influence is wrong from the genetic point of view and, indeed, from my own personal perceptual one.

Conclusion

Heeringa’s work represents a definite improvement on the FFM method by the Hoppenbrouwers brothers. It is able to deliver acceptable overview maps and beautiful continuum maps of the general dialect and language variety of the Netherlands and Flanders. These maps reflect a computerised form of pronunciation differences between dialects. For any other linguistic feature one has to develop other tools.

But also regarding pronunciation differences, we must conclude that where we might expect from Heeringa’s computations some extra understanding of complicated details his method offers no answers or, in some cases, obviously wrong answers. This is partly a consequence of the drawbacks of the RND material. Heeringa also provides us with an outlook of expected better results from the FAND data. As a transcriber of a part of the FAND data, I myself offer him little hope in that respect. So the only solution here would be a completely homogeneous dataset. And even if that was possible, I have my doubts about the validity of the method for detailed problems involving dialects that differ less than 20%. The stress on several highly frequent phonetic features may still distort dialectal oppositions in such cases.

Arjen Versloot
Fryske Akademy