

# Measuring Norwegian dialect distances using acoustic features

Wilbert Heeringa<sup>a,\*</sup>, Keith Johnson<sup>b</sup>, Charlotte Gooskens<sup>c</sup>

<sup>a</sup> *Meertens Institute, Variationist Linguistics, P.O. Box 94264, 1090 GG Amsterdam, The Netherlands*

<sup>b</sup> *University of California at Berkely, Department of Linguistics, 1203 Dwinelle Hall, Berkeley, CA 94720-2650, USA*

<sup>c</sup> *University of Groningen, Scandinavian Department, P.O. Box 716, 9700 AS Groningen, The Netherlands*

Received 12 May 2006; received in revised form 23 July 2008; accepted 23 July 2008

## Abstract

Levenshtein distance has become a popular tool for measuring linguistic dialect distances, and has been applied to Irish Gaelic, Dutch, German and other dialect groups. The method, in the current state of the art, depends upon phonetic transcriptions, even when acoustic differences are used the number of segments in the transcriptions is used for speech rate normalization.

The goal of this paper is to find a *fully* acoustic measure which approximates the quality of semi-acoustic measures that rely on tagged speech. We use a set of 15 Norwegian dialect recordings and test the hypothesis that the use of the acoustic signal only, without transcriptions, is sufficient for obtaining results which largely agree with both traditional Norwegian dialectology and the perception of the speakers themselves.

We use formant trajectories and consider both the Hertz and the Bark scale. We experiment with an approach in which *z*-scores per frame are used instead of the original frequency values. Besides formant tracks, we also consider zero crossing rates: the number of times per interval that the amplitude waveform crosses the zero line. The zero crossing rate is sensitive to the difference between voiced and unvoiced speech sections.

When using the fully acoustic measure on the basis of the combined representation with normalized frequency values, we obtained results comparable with the results obtained with the semi-acoustic measure. We applied cluster analysis and multidimensional scaling to distances obtained with this method and found results which largely agree with both the results of traditional Norwegian dialectology and with the perception of the speakers. When scaling to three dimensions, we found the first dimension responsible for gender differences. However, when leaving out this dimension, dialect specific information is lost as well.

© 2008 Elsevier B.V. All rights reserved.

*Keywords:* Acoustics; Acoustic features; Dialectology; Dialectometry; Dialect; Phonetics; Phonology

## 1. Introduction

Computational dialectometry has been proven to be useful for finding dialect relationships and identifying dialect areas. The first to develop a method of measuring dialect distances was Jean Séguy, assisted and inspired by Henri Guiter (Chambers and Trudgill, 1998). Strongly related to the methodology of Séguy is the work of Goe-

bl, although the basis of Goebel's work was developed mainly independently of Séguy (Goebel, 1982, 1993). In the methodology of both Séguy and Goebel two items to be compared (lexically, phonetically, syntactically or at other levels) are the same or different. Distinctions are binary.

Kessler (1995) used the Levenshtein distance for finding linguistic distances between Irish Gaelic dialect varieties, and the same algorithm was applied to Dutch dialect varieties by Nerbonne et al. (1996). The Levenshtein distance is a sensitive measure with which distances between strings (in this case transcriptions of word pronunciations) are calculated. This means that distinctions between

\* Corresponding author. Tel.: +31 50 363 5970; fax: +31 50 363 6855.  
E-mail addresses: [wilbert.heeringa@meertens.knaw.nl](mailto:wilbert.heeringa@meertens.knaw.nl) (W. Heeringa), [keithjohnson@berkeley.edu](mailto:keithjohnson@berkeley.edu) (K. Johnson), [c.s.gooskens@rug.nl](mailto:c.s.gooskens@rug.nl) (C. Gooskens).

pronunciations of a particular word are gradual rather than just binary. Gooskens and Heeringa (2004) showed that linguistic distances between 15 Norwegian varieties measured with Levenshtein distance correlate significantly with perceptual distances measured between the same 15 Norwegian varieties ( $r = 0.67$ ).

Although the introduction of Levenshtein distance in the field of dialectometry was a significant improvement for comparing dialects phonetically, the results still depend on the quality of the phonetic transcriptions, which may vary greatly, depending on the skills, or idiosyncratic habits, of the transcriber. When several transcribers are involved, the data may reveal ‘dialect’ differences that are actually merely differences between transcribers. For example, Heeringa (2005), found the Frisian dialect area to be divided in a northern and southern part, which reflected the work areas of the two transcribers. The effect of different transcribers on the transcriptions was also seen in an analysis of the whole Dutch dialect area (see Heeringa et al., 2004, pp. 235–266).

A first attempt to measure dialect distances directly on the basis of the acoustic signal was made by Heeringa and Gooskens (2003). But some information from the transcriptions was still used. The number of segments in a pronunciation was used for the purpose of speech rate normalization. We will refer to this methodology as the semi-acoustic approach.

The goal of this paper is to go one step further and to find a fully acoustic measure which approximates the quality of the semi-acoustic measure of Heeringa and Gooskens (2003). We test the hypothesis that varieties of Norwegian can be classified on the basis of acoustic features only, without the filter of a given listener (i.e. the transcriber). The classification scheme obtained in this way correlates significantly with both the traditional dialectology criteria, which classifies the dialects according to a number of relevant linguistic features (phonological, lexical, etc.) and the results of a perceptual classification experiment. We will experiment with different representations of the acoustic signal to investigate which representation gives the best results.

The basis of the research presented in this paper is a database which contains recordings of Norwegian dialect varieties compiled by Jørn Almberg and Kristian Skarbø.<sup>1</sup> The database comprises recordings of translations of the fable ‘The North Wind and the Sun’. In this paper, we will compare our results to the results of a perception experiment reported by Charlotte Gooskens. When the perception experiment was carried out, recordings of only 15 varieties were available. Therefore, we use the same 15 varieties. Today more than 50 recordings are available, giving much better possibilities to pick a representative selection of varieties.

Section 2 gives a brief overview of the main linguistic phenomena that play a role in traditional Norwegian dialectology. In Section 3, we describe the perception experiment. In Section 4, we describe our acoustic model and its parameters. In Section 5, we validate the results of our methodology. Section 6 shows results. The results will be compared to both the results of traditional Norwegian dialectology and the results of the perception experiment. In Section 7, some conclusions will be drawn.

## 2. Traditional Norwegian dialectology

In traditional Norwegian dialectology the dialect map of Skjekkeland (1997) is an authoritative map. It divides the Norwegian dialect area in two main groups, Vestnorsk (northern and southwestern varieties) and Austnorsk (southeastern varieties). Vestnorsk is divided in Nordnorsk (north) and Vestlandsk (southwest). Nordnorsk, Vestlandsk and Austnorsk in turn are divided in three smaller groups, giving a total of nine groups. In our set of 15 dialect varieties, six groups are represented. On the map in Fig. 1 the six groups as represented by the 15 varieties are shown.

Skjekkeland’s map is based on 24 single linguistic features. In this section, we classify the 15 varieties on the basis of especially those features which are represented in our data, i.e. the transcriptions of the Norwegian translations of ‘The North Wind and the Sun’. The text consists of 58 different words. Since we have a translation of the text for each of the 15 varieties, we have a translation of each of the 58 words in (nearly) all dialects. Due to the free translation of some phrases for certain varieties a few of the expected words were missing. If the same word appears more than once in a text, we consider only the first occurrence.

### 2.1. Features

In this section, we give an overview of the features which have contributed to Skjekkeland’s dialect classification and which are represented in our data. In some cases we discuss features which are closely related to the ones mentioned by Skjekkeland.

*Apocopation of verb endings.* Skjekkeland shows that the endings of infinitive verbs and weak feminine nouns might have different pronunciations or they have been apocopated. In our data we found three infinitive verbs which show a clear distinction between forms with final vowel and apocopated forms: *kunne* ‘could’, *gjelde* ‘count’ and *innrømme* ‘admit’. The same distinction was found for third person singular of three verbs given in the past tense: *skulle* ‘would’, *bläste* ‘blew’ and *mätte* ‘had to’. The distributions of the six verbs are shown in Table 1. The table shows that we find more apocopated verbs when we go further to the north. In the Trøndsk varieties of Bjugn, Stjørdal, Trondheim

<sup>1</sup> Department of Linguistics, University of Trondheim. The recordings are available at <http://www.ling.hf.ntnu.no/nos/>.

Download English Version:

<https://daneshyari.com/en/article/566251>

Download Persian Version:

<https://daneshyari.com/article/566251>

[Daneshyari.com](https://daneshyari.com)