Change, Convergence and Divergence among Dutch and Frisian

Wilbert Heeringa and John Nerbonne

University of Groningen, Faculty of Arts, Alfa-Informatica

Abstract

The *Algemeen Nederduitsch en Friesch dialecticon* (Winkler, 1874) (ANFD) contains 186 translations of the parable of 'the prodigal son' into dialects of the Netherlands, northern Belgium and western Germany. In 1996 Harrie Scholtmeijer repeated the work of Winkler for the Netherlandic dialects. We used both sources for diachronic research on dialects. For synchronic measurement of distances between dialects, we used the Levenshtein distance (first applied by Kessler, 1995). Using Levenshtein distance, two transcriptions corresponding to two dialectal pronunciations of the same word) are compared to each other by calculating the cost of (the least costly set of) operations mapping one string of phonetic transcription to another by inserting, deleting or replacing phones. The results of comparison are analyzed further by clustering (results in a tree structure) and by multidimensional scaling (results in a map where distances depend on the phonetic distance). For diachronic research of dialects, the Levenshtein distances are calculated and analyzed for both the 1874 variants and the 1996 variants, to see how the relation between Frisian and Dutch dialects with respect to standard Dutch has changed.

1. Introduction

In 1874 the *Algemeen Nederduitsch en Friesch dialecticon* was published by Martinus Nijhoff, a book in two parts, compiled by Johan Winkler (Winkler, 1874). It contains 186 translations of the parable of the `prodigal son' (Luke 15:11-32 of the Bible) into dialects of the Netherlands, northern Belgium and western Germany. The recordings were not made by phoneticians, but by clergyman, schoolteachers, notaries, etc. The texts are given in different spellings: German, Dutch, official Frisian (Colmjon, 1863) and a Frisian spelling made by Winkler himself. Often the texts are translations of the Statenvertaling text, but it seems that other translations were also taken as a source.

In 1996 Harrie Scholtmeijer repeated the work of Winkler. He collected 81 translations of the parable of the 'prodigal son' of dialects in the Netherlands. 74 variants are the same as in the Winkler source. The recordings are made by teachers, employees, housewives, civil servants, farmers, etc. The texts are given in spellings of dialect dictionaries or in a spelling which is based on the Dutch spelling. In most cases, the 1951 translation published by the 'Nederlands Bijbelgenootschap' was taken as a sources. In one case (Tholen) the Statenvertaling was used for religious reasons. Because our research is based on words, missing words were asked by telephone. Words were missing as a result of a free translation, or due to the fact that they usually were not used in the given context for some dialects.

In our research dialect pairs are compared. In that way old and new results can be compared. Sources can also be direct compared to each other by calculating the distances between the old variants and the corresponding new variants.

2. Data

2.1 Dialects

From the 74 variants which could be found in both the Winkler and the Scholtmeijer source we choose 42, including Standard Dutch and Standard Frisian. Because Standard Frisian is most like the dialect of Grouw, we located it at the position of Grouw (see Figure 1). Standard Dutch can likely not be located in the dialect continuum. However in the visualization of the results this is sometimes necessary. Therefore we located Standard Dutch at the position of

Dronten. According to Scholtmeijer and Kapteijn (1998) the most pure Standard Dutch is spoken in Dronten.

Sometimes the town in the Winkler source was replaced by another nearby in the Scholtmeijer. In these cases we use the area name, because the places are not exactly the same in 1874 and 1996. In the two places (nearly) the same dialect is spoken. Sometimes in the Winkler source the dialect for an area is given instead of for a place. The table below summaries the replacements used.

	1874	1996
Hunsingo	Ulrum	Zoutkamp
Stellingwerf	Noordwolde	Stellingwerf
Baronie van Breda	Rijsbergen	Etten-Leur
Oldambt	Oldambt	Westerlee
Westerwolde	Sellingen	Ter Apel

We revised the 1874 text from Ouddorp on the basis of the notes of P. Heerschap. He suggests that the Winkler text was likely drawn up by a non-native speaker because the text is still more modern than the 1996 one.

For Standard Dutch for 1874 we used the Winkler text number 115: ' De gelijkenis van de verlorenen zoon in den modern Hollandschen tongval, zoals die onder anderen te Haarlem wordt uitgesproken' . For 1996, the Winkler text number 115 was used as a basis, where the old-fashioned word ' wier' is replaced by ' werd' (became), and the pronunciations are adapted in accordance with Paardekooper (1998).

For Standard Frisian for 1874 we used the Winkler text number 87: ' Dialectus communis van de provincie Friesland' . For interpreting this text Winkler refers to Colmjon (1863). For 1996 we used the text in Dijkstra (1995).

2.2 Words

2.2.1 Choice of words

The algorithm we used for comparing dialects (the Levenshtein distance) works on the basis of words. So from the text we choose words which vary lexically or phonetically between the dialects. We did not choose words which appear only in the ' Statenvertaling' text or only in the 1951 translation of the ' Nederlands Bijbelgenootschap' . Most words must appear in almost every text, even in cases of rather free translation. We did not choose words which vary in form due to proximity to other words (e. g. cliticization, sandhi), for example: because ' daar loopt hij' is often pronounced as ' daar loopt' ie, ' hij' is not a suitable word. We did not use words which are pronounced differently stressed and unstressed, for example: the stressed word ' jij' becomes ' je' if unstressed. We did not include words which in Dutch spelling end in *en*, for example: teg*en*, lop*en*, voet*en*. This because of the fact that always in Winkler and sometimes in Scholtmeijer loopm ([lo.pm̩]) is spelled as lopen ([lo.pən]). If taking into account all these restrictions it turned out that not more than 60 words were suitable (see Table 1).

2.2.2 Choice of translations

In 2.2.1 we explained how we made the word list. Now for each dialect in the corresponding text we search for the translations of the words in the word list. A translation is useful if it is synonymous in the given context with the word in the word list, for example: ' to run' is synonymous with ' to dash' . A translation consisting of a more general word with modification was not acceptable, for example: ' to run' is not synonymous with ' to walk quickly' . In some dialect texts no translation could be found for some words due to "freer" translation.

In the 42 1874 texts, the minimum number of missing words is 0, the maximum number is 16 and the mean is 4.63. For the 1996 texts the missing words were asked by telephone to the translator or someone who speak the same dialect as the translator. In that way, in the 1996 texts the minimum number of missing words is 0, the maximum number could be reduced to 3 and the mean is 1.05.

If comparing two dialects on the basis of the translations of the words, and in one of them or in both no translation is present, the word is not taken into account.

2.2.3 Transcriptions of words

2.2.3.1 Vowels

The [ʊ]is pronounced between [u] and [o]. The [ʏ] is pronounced between [ø] and [o]. In the feature system of Hoppenbrouwers (Hoppenbrouwers, 1988) [ʊ] get the values of [o], and [ʏ] get the values of [ø]. The results in this paper are made on the basis of the feature system of Vieregge (Vieregge (1984) and Vieregge (1987)). However, we changed the feature values of the vowels analogous to the system of Hoppenbrouwers (see Nerbonne and Heeringa (1998) for more details). If a <o> in a closed syllable precedes the <m>, <n> or <ng>, we noted it as [ʊ]. For example: <bom> is noted as [bʊm].

In Saxon and Franconian dialects the <oa> and the <ao> are pronounced between the [o] and the [ɑ]. If nothing else is indicated by Winkler (1874) or in relevant dictionaries (1996) we suppose the <oa> tends towards the <o> and the <ao> tends towards the <a>, as is suggested by Winkler. We note the <oa> as [ʊ.] and the <ao> as [ɔ.].

In Frisian dialects the <oa> usually represents a diphthong. We note the <ao> as [ʊə]. In texts of other dialects the same sound is also spelled as <oë> (Ouddorp 1874), <ôô> (Gorinchem, Land van Axel, Laren, Ouddorp and Tholen 1996), <ôo> (Walcheren 1996), <oô> (Hoorn and Schagen 1996) and <óó> (Baronie van Breda and Texel 1996).

In Frisian dialects the <ea> represents a diphthong. We note the <ea> as [ɪə]. In texts of other dialects the same sound is also spelled as <èeë> (Scheveningen 1874), <èë> (Ouddorp, Tholen and Walcheren 1874), <ae> (Scheveningen 1996), <êê> (Gorinchem, Land van Axel, Ouddorp and Tholen 1996), <êe> (Walcheren 1996), <eê> (Hoorn and Schagen 1996) and <éé> (Baronie van Breda and Texel 1996).

For diphthongs we follow the distinction between real diphthongs and pseudo-diphthongs as given in Rietveld and van Heuven (1997). Real diphthongs are <ei/ij>, <ui> and <ou/au>. Here we transcribe for example <bij> as [bɛi], <bijen> as [bɛiə], <bui> as [bœy], <buien> as [bœyə], <bouw> as [bɔu] and <bouwen> as [bɔuə].

For the dialects of Woltersum, Oldambt and Westerwolde, Winkler describes the <ei> as a sound which tends towards the <ai>. If in 1996 in the same word at the same position an <ai> is noted, we interpreted the <ei> in the 1874 text as [æi].

Pseudo-diphthongs are arbitrarily vowels followed by the [j] or [u]. Here we transcribe for example <nieuw> as [niu], <nieuwe> as [niuə], <meeuw> as [meu], <meeuwen> as [meuə], <haai> as [haj], <haaien> as [hajə], <hooi> as [hʊj], <hooien> as [hʊjə], <boei> as [buj] and <boeien> as [bujə]. Following the choice of Hoppenbrouwers (1988) most pseudo-diphthongs will be treated as a sequence of two separated vowels.

Besides real diphthongs and pseudo-diphthongs, the centering diphthongs are also recognized as a special class in Hoppenbrouwers (1994). Centering diphthongs are arbitrary vowels followed by the [ə]. We also recognize them as a special class.

Special cases are vowels preceding the <r>. We transcribe <ier> as [iər], <eer> as [ɪər], <uur> as [yər], <eur> as [ʏər], <oer> as [uər] and <oor> as [ʊər]. However, we transcribe

<ir> as [ɪr], <er> as [ɛr], <ur> as [ʏr], <' r> as [ɐr], <aar> as [ar], <ar> as [ɑr] and <or> as [ɔr]. For the [æ] and the [ɒ] no fixed spelling is known. However, if they occur followed by the <r>, we transcribe them respectively as [ær] and [ɒr].

2.2.3.2 Consonants

In the texts no distinction is made between [r] (alveolar trill) and [ʀ] (uvular trill). We always transcribe the <r> as [r]. In the texts no distinction is made between [ɣ] and [g]. We always transcribe the <g> as [ɣ]. We transcribe <v> as [v], <f> as [f], <z> as [z] and <s> as [s].

For some consonants the transcription depends on the position -- at the beginning or at the end of a syllable. At the beginning we transcribe <g> as [ɣ], <w> as [ʋ], <d> as [d] and <b> as [b]. At the end we transcribe <g> as [x], <w> as [w], <d> as [d] and <b> as [b].

For the dialects of Walcheren, Tholen and Land van Axel a <h> is pronounced where a <g> is spelled in Standard Dutch. In our word list it concerns the following words (given in Standard Dutch): *g*ebracht, be*g*on, *g*ing, *g*aan and *g*ezond. If in the 1874 texts a <g> is noted where in the 1996 a <h> is noted, we replaced these <g> with an <h>.

2.2.3.3 Diacritics

When converting words from spelling to phonetic script, we do not use diacritics, except length marks. Furthermore we take into account whether symbols are written smaller and/or in superscript (this occurs only in the 1874 texts). In that case the significance of the segment is halved. The reduction in diacritics is justified because the level of detail in transcription varies. It is better to compare on the basis of the same level of detail.

If no special information is given about the length, we chose default lengths. The lengths of monophthongs are borrowed from Rietveld and van Heuven (1997). Short are: [ɪ], [ɛ], [ɑ], [ɔ], [ʊ], [ʏ], [ə] and [æ]. Half long are: [i.], [y.] and [u.]. Long are: [e:], [a:], [ø:] and [o:].

For some special notations in spelling we deviate from the default lengths. <oa> is transcribed as [ʊ.], <ao> as [ɔ.], <ii> as [i:], <ææ> as [ɛ:], <òò> as [ɔ:], <èè> as [ɛ:], <ê> as [ɛ:] and <ô> as [ɔ:]. Furthermore, the length of all types of diphthongs (real, pseudo or centering) is long.

3. Comparison of dialects

As mentioned in 2.2.1 60 words are chosen from the text. On the basis of these words, in transcribed form, we compare dialects by calculating their mutual distances. For this purpose we use Levenshtein distance for measuring the phonetic distances between dialects. The distance measure is explained in Kruskal (1999) and was first applied to dialect data by Kessler (1995).

The Levenshtein distance may be understood as the cost of (the least costly set of) operations mapping one string to another. The basic costs are those of (single-phone) insertions, deletions and substitutions. Insertions and deletions cost half that of substitutions. The principle can be illustrated by a small example. In Standard American ‘saw a girl’ is pronounced as [sɔ:əgIrl] while in the dialect of Boston this is pronounced as [sɔ:rəgø:l]. Now we change the first pronunciation into the other.

sɔəgIrl	delete r	1
sɔəgIl	replace I/ø	2
sɔəgøl	insert r	1
sɔrəgøl		

Many sequence operations map [sɔ:əgIrl] in [sɔ:rəgø:l], but the Levenshtein distance is always equal to the cost of the cheapest mapping.

The simplest versions of this method are based on a notion of phonetic distance in which phonetic overlap is binary: nonidentical phones contribute to phonetic distance, identical ones do not. Thus the pair [a,p] counts as different to the same degree as [b,p]. In more sensitive versions phones are compared on the basis of their feature values, so the pair [a,p] counts as more different than [b,p].

In this research we used the feature system of Vieregge (1987). Vieregge' s system consists of 4 multi-valued features for vowels, and 10 multi-valued features for consonants. Vieregge' s system was developed for a similar comparison task, that of checking the quality of phonetic transcription. This involves comparison to consensus transcriptions.

Nerbonne et al. (1996), Nerbonne and Heeringa(1998) and Nerbonne et al. (1999) show the application of the Levenshtein distance to Dutch dialects.

4. Classification of dialects

On the basis of the Levenshtein distances the dialects are classified by clustering (Jain and Dubes, 1988) or multidimensional scaling (Kruskal and Wish, 1984). The final result of clustering is a dendrogram which is a hierarchically structured tree in which the dialects are the leafs. The result of multidimensional scaling is a map, where the geographic distance between kindred dialects is small, and between different dialects great.

4.1 Clustering

The basis for clustering is a matrix in which the Levenshtein distances between dialects are arranged. Assume we get the following matrix:

	Assen	**Delft**	**Kollum**	**Nes**	**Soest**
Assen	**0**	***73***	***64***	***67***	***79***
Delft	**73**	**0**	***81***	***74***	***68***
Kollum	**64**	**81**	**0**	***43***	***91***
Nes	**67**	**74**	**43**	**0**	***86***
Soest	**79**	**68**	**91**	**86**	**0**

Clustering here is most easily understood procedurally. In the matrix only the upper half is used (values in italics). At each iteration of the procedure, we select the shortest distance in the matrix. At the first iteration the shortest distance is between Kollum and Nes (43). Then we fuse the two data points which gave rise to it. Kollum and Nes are removed from the matrix, while a new cluster Kollum/Nes is inserted. To iterate, we have to assign a distance from the newly formed cluster to all other points. In the first iteration, the distances from Kollum/Nes to all remaining points have to be calculated. For example, the distance from Kollum/Nes to Assen can be calculated by taking the average of the distance Kollum-Assen and the distance Nes-Assen: (64+67)/2=65.5. Besides the average, there are several more alternatives (Jain and Dubes, 1988). Now the next iteration is executed until one cluster is left.

Figure 2 shows the results for the 1874 distances, while figure 3 shows the results for the 1996 distances.

4.2 Multidimensional scaling

On the basis of geographic coordinates the distances between locations can be determined. The reverse is also possible: on the basis of the mutual distances, an optimal coordinate system can be determined with the coordinates of the locations in it. The last is realised by a technique known as ‘multidimensional scaling’. On a multidimensional scaling plot, strongly

related dialects are located close to each other, while strongly different dialects are located far away from each other (Kruskal and Wish, 1984).

As input each dialect is defined as a range of distances, namely the distance to itself and the distances to the other dialects. The distances correspond to dimensions. If we have 84 variants (42 old and 42 new ones), we get 84 dimensions. With multidimensional scaling, the dimensions can be reduced to 2, 3 or more dimensions, so for each variant we get a coordinate in 2, 3 or more dimensional space.

Figure 4 shows the result. Y-coordinates represent the first and X-coordinates represent the second dimension. The labels of the 1874 variants are in lower case, the labels of the 1996 variants in capitals. In scaling to 2 dimensions, both dimensions turned out to be geographic: the first dimension corresponds to the west-east axis, the second dimension to the north-south axis.

5. Convergence and divergence

We calculated the mutual Levenshtein distances between the old variants, the mutual distances between the new dialects, and the distances between the old and the new variants. The correlation between the old phonetic distances and the geographic distances turned out to be equal to 0.5218, while the correlation between the new phonetic distances and the geographic distances is equal to 0.4954. Furthermore the correlation between the old and the new distances is equal to 0.9186. All correlation coefficients mentioned here are significant.

5.1 Examining classification results

Examining the 1874 dendrogram (Figure 2), we see that there is a clear division between Frisian dialects and other dialects. Within Frisian, there is a division in rural Frisian and town Frisian. Within other dialects, we get a division in the dialects of Zeeland, Franconian dialects and Saxon dialects. Usually Zeeland is grouped under the Franconian dialects. Some properties which make the Zeeland dialects apart from the Franconian dialects are h-deletion (<huis> is pronounced as [y.s], <hemel> is pronounced as [e:məl]) and the rather frequent appearance of the diphthong [ɪə] (<één> is pronounced as [ɪə]). Further we note that at the end of <ring> always a [k] is added, which was not found in the immediately adjoining dialects.

For 1996 (Figure 3) the division between Frisian dialects and other dialects is still maintained. However, town Frisian has moved to ' the other dialects' . Maybe this can be explained from the fact that the ' afsluitdijk' was built in 1932, which caused a stronger contact between Holland and Friesland. Roughly speaking, the other dialects are divided in a western and an eastern group. The western group contains mainly Franconian dialects while the eastern group contains chiefly Saxon dialects. Note that in that view the dialects of Limburg should belong to the Saxon dialects. One of the joining properties of the eastern dialects is the word <ver> which is always pronounced as [ʋi:t].

The multidimensional scaling plot clarifies how relations among the old and the new variants have changed (Figure 4). The position of a dialect on the plot depends on the relations to the other points on the plot. So the shift of a dialect doesn' t say anything about the dialect itself, but rather about the relations to the dialects in their context. Note the small shift of Standard Dutch and the big shift of Helmond from the west to the east. Within rural Frisian West-Terschelling shifted at most, while Hindeloopen turned out to be very stable. Within town Frisian, Bolsward and Leeuwarden shifted at most. As the following section will show, this change in relations does not imply that all the varieties involved underwent diachronic change. For example West-Terschelling changed very little. Because most neighbouring varieties did change, its relations changed a great deal.

5.2 Convergence and divergence with respect to standard languages

By subtracting the old distances from the new distances, we get negative and positive values. Negative values correspond with convergence, while positive values corresponds with divergence.

Figure 5 shows that 23 of the 41 variants converged to Standard Dutch. Especially Bolsward, Leeuwarden, Oldambt and Zweeloo strongly converged to Standard Dutch. Figure 6 shows that 18 variants diverged from Standard Dutch. It concerns notably dialects along the South-West coast line (Zaankant, Scheveningen, Ouddorp and Walcheren) and dialects in the Middle-East (Oldenzaal, Deventer and Varsseveld). For the Frisian dialects no systematic tendencies could be found.

Turning our attention to Friesland, we also added Texel and Grijpskerk. Grijpskerk belongs to the dialects of ' Westerkwartier' which are an form intermediate between the Groningen dialects and Frisian. As mentioned in 2.1 we place Standard Frisian at the position of Grouw in the maps produced for this paper. Figure 7 shows which dialects converged to Standard Frisian. Only Het Bildt and Stellingwerf rather strongly converged to Standard Frisian. Figure 8 shows which dialects diverged. Texel converged most strongly. Examining both figures we could not find that a particular group (rural Frisian, town Frisian or Stellingwerf/Grijpskerk) converged or diverged.

For each dialect we also calculated the distance between the old and the new variant. Then we measure the change of a dialect to itself, not the changes in relations to other dialects. In the Figure' s 5, 6, 7 and 8 lighter dots represent greater distances between the old and the new variant. We see that Standard Dutch, West-Terschelling and Het Bildt were very stable, while Zweeloo, Helmond, Baronie van Breda and Aardenburg changed a lot.

5.3 Dialects in and near Friesland: Dutch or Frisian?

Goossens(1977) tries to answer the question: What are Dutch dialects? (pp. 11-30). Among other things he discusses the definition which recognises a Dutch dialect as a variant which shows more characteristics of Dutch than of any other standard language. Goossen' s objection against this definition is that applying it is not feasible. One has to know the complete vocabulary of all speakers of the dialect you want to research, and besides, the vocabulary of a speaker can change any time.

Ignoring this objection, and assuming that the informants are representative of the dialects, and that the word list is representative of the vocabulary of the informants, we can determine if the dialects in and around Frisian are Dutch or Frisian using Levenshtein distance. Applying this definition all rural Frisian dialects turned out to be Frisian (as expected). The dialects just outside Friesland (Grijpskerk, Stellingwerf and Texel) are Dutch dialects. The town Frisian dialects (Ameland, Bolsward, Dokkum, Het Bildt, Leeuwarden, Midsland) are Dutch. So we could affirm Kloeke' s characterisation of town Frisian: "Hollands in friese mond" and reject the characterisation of Gosses who like to regard town Frisian as Frisian "met wat Hollands erbij" (Goossens, 1977). We obtain these findings on both the 1874 and the 1996 source.

6. Conclusions

The research shows that in 1874 there was a rather sharp division in Frisian, Saxon en Franconian dialects, although the dialects of Zeeland took a special place. In 1996 we get a division in Frisian, Western Dutch and Eastern Dutch dialects. The town Frisian dialects shifted from the Frisian dialects to the Dutch dialects.

Tendencies in convergence and divergence seem to be rather arbitrary. We observe however, that the majority of dialects converged to Standard Dutch. Only the dialects along the South-West coast line and in the Middle-East diverged somewhat from Standard Dutch. In the

course of time Standard Dutch, West-Terschelling and Het Bildt turned out to be very stable, while Zweeloo, Helmond, Baronie van Breda and Aardenburg changed a lot to itself. If examining whether dialects in and near Friesland are Frisian or Dutch, only the rural variants (West-Terschelling, Oost-Terschelling, Schiermonnikoog and Hindeloopen) are Frisian dialects. The town Frisian dialects are Dutch dialects.

Acknowledgements

We thank Harrie Scholtmeijer for making available his 1996 texts and for giving help in interpreting the 1874 texts. We thank Simon Reker, Hermann Niebaum, Tjeerd de Graaf and Wouter Jansen for giving advises in converting the words in the texts from spelling to transcription. We thank Peter Kleiweg for his graphic programs seen in all the picture in this paper. We thank Frits Steenhuisen for providing the coordinates of the map of the Netherlands and Friesland.

References

Colmjon, G. (1863)
Beknopte Friesche spraakkunst voor den tegenwoordigen tijd, Gezelschap voor Friesche taal en letterkunde. Leeuwarden: H. Kuipers.

Dijkstra, F.D. et al. (1995)
Bibel, út de oarspronklike talen op 'e nij yn it Frysk oerset, Oersetting yn opdracht fan it Nederlands Bijbelgenootschap en de Katholieke Bijbelstichting en ûnder ferantwurdlikens fan de Stifting Yntertsjerklike Kommisje foar de Nije Fryske Bibeloersetting. Haarlem: Nederlands Bijbelgenootschap (3th edition).

Goossens, J. (1977)
Inleiding tot de Nederlandse Dialectologie. Groningen: Wolters-Noordhoff (2nd edition).

Hoppenbrouwers, C. (1994)
` De indeling van de zuidoostelijke streektalen' , in: *TABU: Bulletin voor taalwetenschap* **24**(2), pp. 37-63.

Hoppenbrouwers, C. and G. Hoppenbrouwers (1988)
` De featurefrequentiemethode en de classificatie van Nederlandse dialecten' , in: *TABU: Bulletin voor taalwetenschap* **18**(2), pp. 51-92.

Jain, A. K. and R. C. Dubes (1988)
Algorithms for clustering data. Englewood Cliffs and New Yersey: Prentice Hall.

Kruskal, J. B. (1999)
` An overview of sequence comparison' , in: D. Sankoff and J. Kruskal (eds.), *Time Warps, String edits, and Macro molecules: The Theory and Practice of Sequence Comparison*. Stanford: CSLI (2nd edition), pp. 1-44.

Kruskal, J. B. and M. Wish (1984)
Multidimensional Scaling. Beverly Hills and London: Sage Publications.

Nerbonne, J. and W. Heeringa (1998),
` Computationele Classificatie van Nederlandse Dialecten' , in: *Taal en Tongval* **50**(2), pp. 164-193. Available as: http://www.let.rug.nl/~nerbonne/papers/tetv99.ps.

Nerbonne, J., W. J. Heeringa and P. Kleiweg (1999),
` Edit distance and dialect proximity', in: Sankoff, D. and J. Kruskal (eds.), *Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison*. Stanford: CSLI (2nd edition), pp. v-xv. Available as: http://www.let.rug.nl/~nerbonne/papers/timewarp.ps.

Nerbonne, J., W. Heeringa, E. van den Hout, P. van der Kooi, S. Otten and W. van de Vis (1996)
` Phonetic Distance between Dutch Dialects' , in: G. Durieux, W. Daelemans and S. Gillis (eds.), *CLIN VI, Papers from the sixth CLIN meeting*. Antwerp: University of Antwerp, Center for Dutch Language and Speech, pp. 185-202. Available as: http://www.let.rug.nl/~heeringa/dialectology/clin96.ps.

Paardekooper, P. C. (1998)
ABN-uitspraakgids. Den Haag: Sdu Uitgevers (3th edition).

Rietveld, A. C. M. and V. J. van Heuven (1997)
Algemene fonetiek. Bussum: Dick Coutinho.

Scholtmeijer, H. and M. Kapteijn (1998)
' Het Nederlands van jongeren in Oost-Flevoland' , in: *Taal en Tongval*, **50**(1), pp. 63-80.

Vieregge, W. H., et al. (1984)
A Distinctive Feature Based System for the Evaluation of Segmental Transcription in Dutch, Nijmegen, in: M. P. R. van den Broecke and A. Cohen (eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences*. Dordrecht and Cinnaminson: Foris Publications, pp. 654-659.

Vieregge, W. H. (1987)
` Basic aspects of phonetic segmental transcription' , in: A. Almeida and A. Braun (eds.), *Probleme der phonetischen Transkription, Zeitschrift für Dialektologie und Linguistik, Beihefte*. Stuttgart: Franz Steiner Verlag Wiesbaden GMBH, pp. 5-55.

Winkler, J. (1874)
Algemeen Nederduitsch en Friesch dialecticon. ' s-Gravenhage: Martinus Nijhoff.

1	had	had	11
2	twee	two	11
3	jongste	younger	12
4	vader	father	12,18,21
5	deel	portion	12
6	en	and	13,15,29
7	naar(prep)	to	13,18,28
8	ver	far	13
9	land	country	13,14
10	doorgebracht	squandered(part)	13,14,30
11	toen	when, then	14,17
12	kwam	came	14,25
13	begon	began	14
14	ging	went	15,20,28
15	stuurde	sent(past)	15
16	om	to, around	15,20
17	brood	bread	17
18	honger	hunger	17
19	zal	will	18
20	gaan	go	18,28
21	heb	have	18,21,29
22	hemel	heaven	18,21
23	niet	not	19,21,28
24	meer(adv)	more	19,21
25	waard(adj)	worthy	19,21
26	zoon	son	19,21
27	één	one	19,26
28	zag	saw	20
29	liep	ran(past)	20
30	kuste	kissed(past)	20
31	maar	but	22
32	breng(com)	bring(com)	22,23
33	beste	best	22
34	aan	on	22
35	ring	ring	22
36	kalf	calf	23,27,30
37	blij	glad	23,32
38	was	was	24
39	dood	dead	24,32
40	levend	alive	24,32
41	oudste	oldest	25
42	dicht(adv)	close, near	25
43	huis	house	25,28
44	riep	called(past)	26
45	vroeg(verb)	asked(past)	26
46	zei	said(past)	27,31
47	broer	brother	27,32
48	heeft	has	27,30
49	gezond	healthy	27
50	werd	get	28
51	boos	angry	28
52	wilde	wanted(past)	28
53	nooit	never	29
54	met	with	29,30
55	nu	now	30
56	die	which	30
57	voor(prep)	for	30
58	bent	are	31
59	altijd	always	31
60	bij(prep)	near, with	31

Table 1. The words chosen from the text of the parable of ' the prodigal son' .

Figure 1. The locations of the 41 dialects.

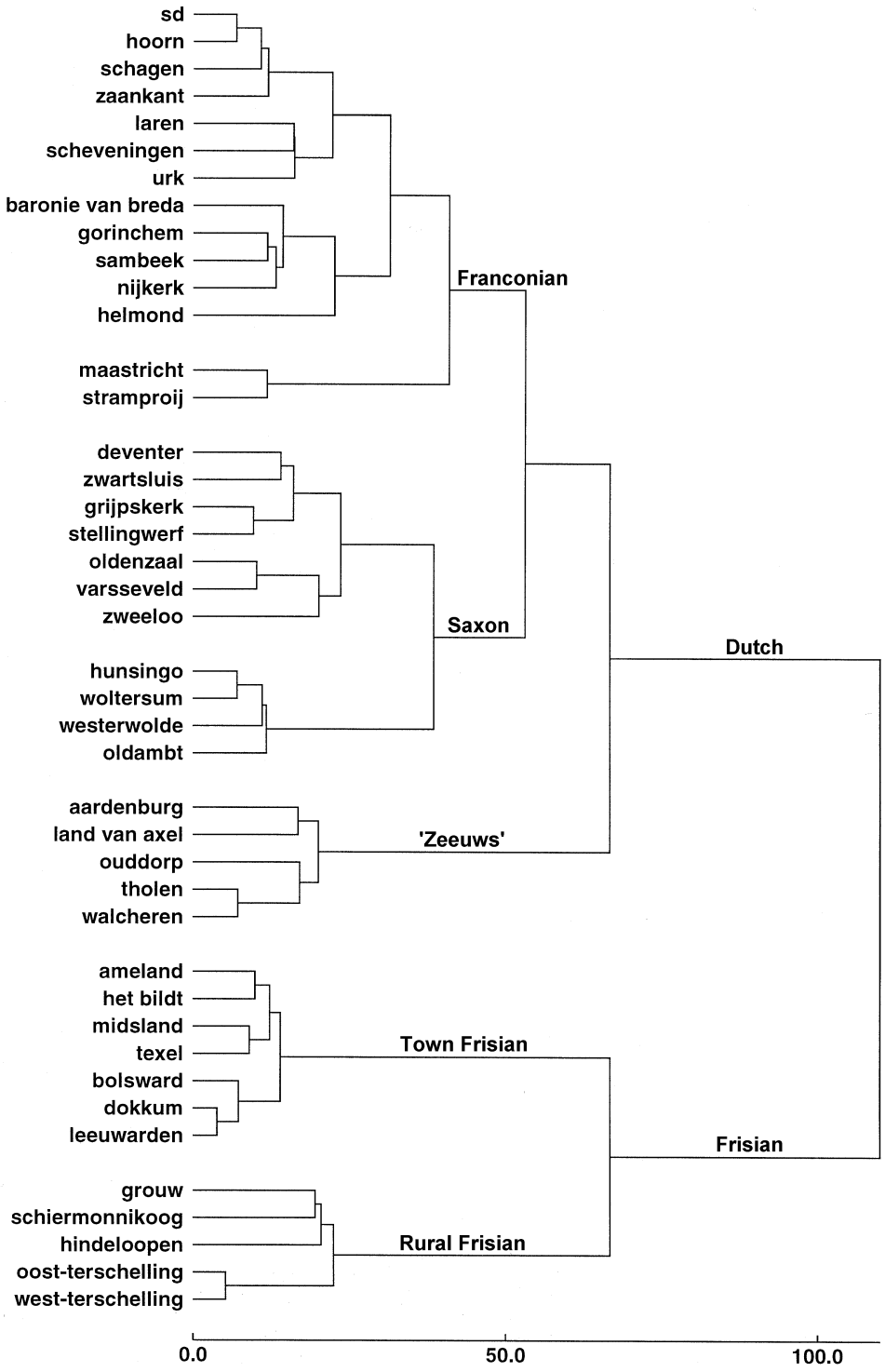

Figure 2. A dendrogram derived from the mutual distances of the 1874 variants. Some labels might be controversial.

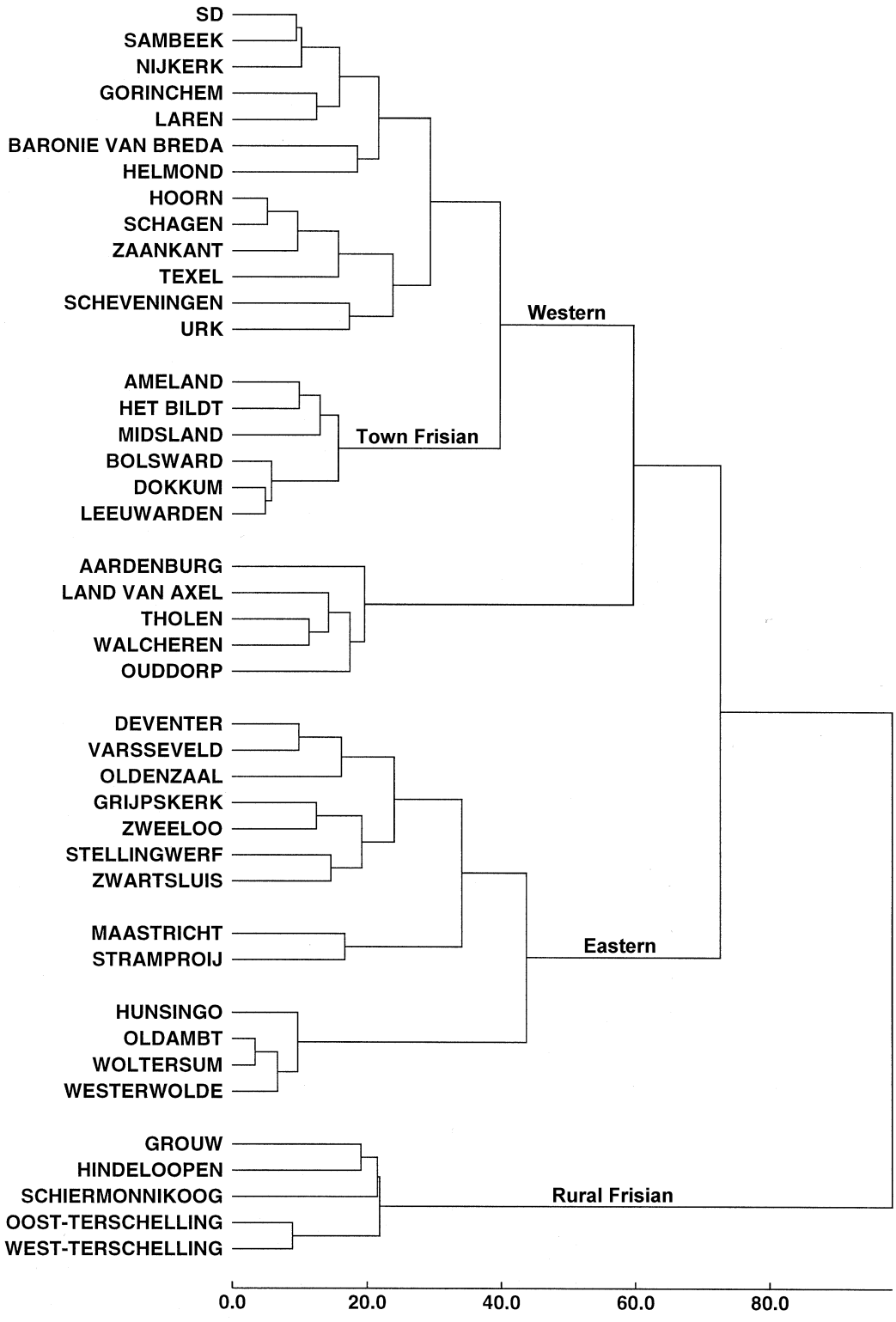

Figure 3. A dendrogram derived from the mutual distances of the 1996 variants.

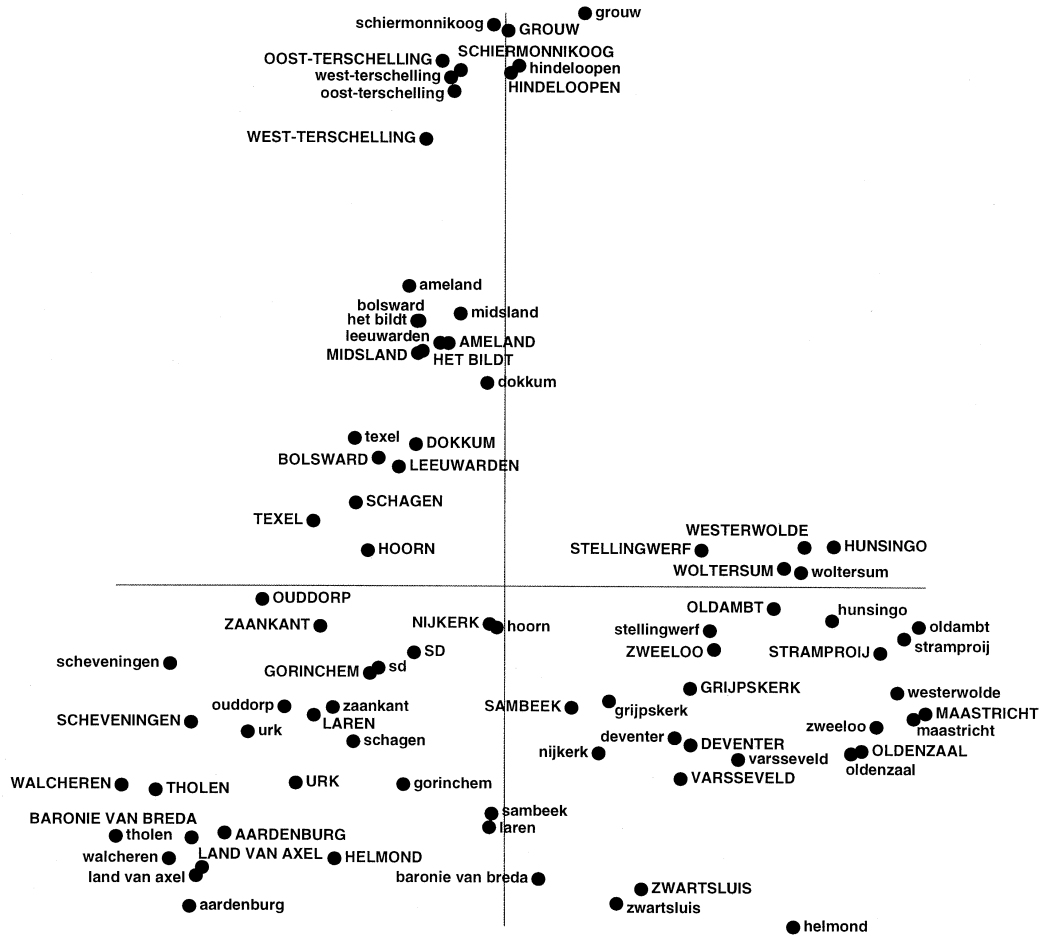

Figure 4. Using multidimensional scaling 84 dimensions are reduced to 2 dimensions. Y-coordinates represent the first and X-coordinates represent the second dimension. The labels of the 1874 varieties are in lower case, the labels of the 1996 varieties in capitals. Note the shift of Helmond from the west to the east.. The plot is comparable with the geographic map. Note the central positions of old and new Standard Dutch (sd/SD).

Figure 5. Darker lines indicate stronger convergence. Lighter dots represent greater distances between the old and the new variant. 23 of the 41 variants converged toward Standard Dutch.

Figure 6. Darker lines indicate stronger divergence. Lighter dots represent greater distances between the old and the new variant. 18 of the 41 variants diverged from Standard Dutch.

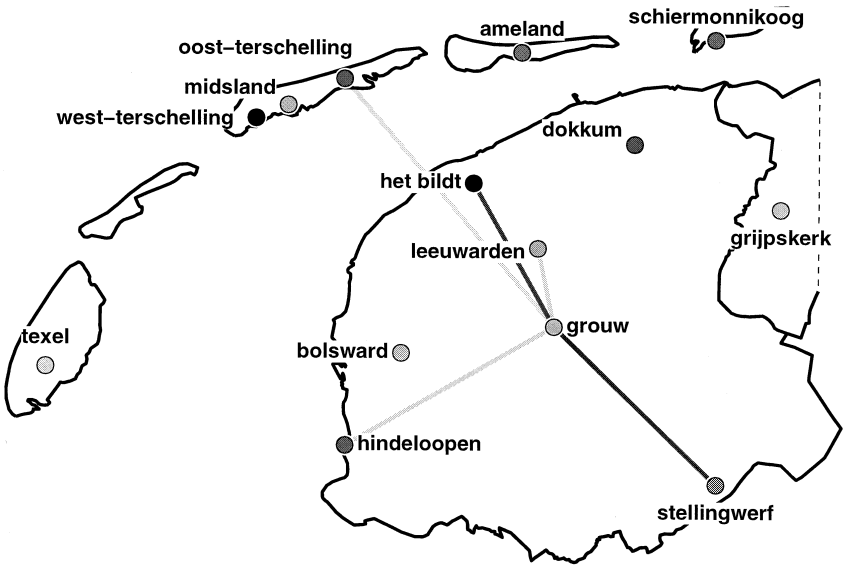

Figure 7. Convergence with respect to Standard Frisian. Darker lines indicate stronger convergence while lighter dots represent greater distances between the old and the new variant.

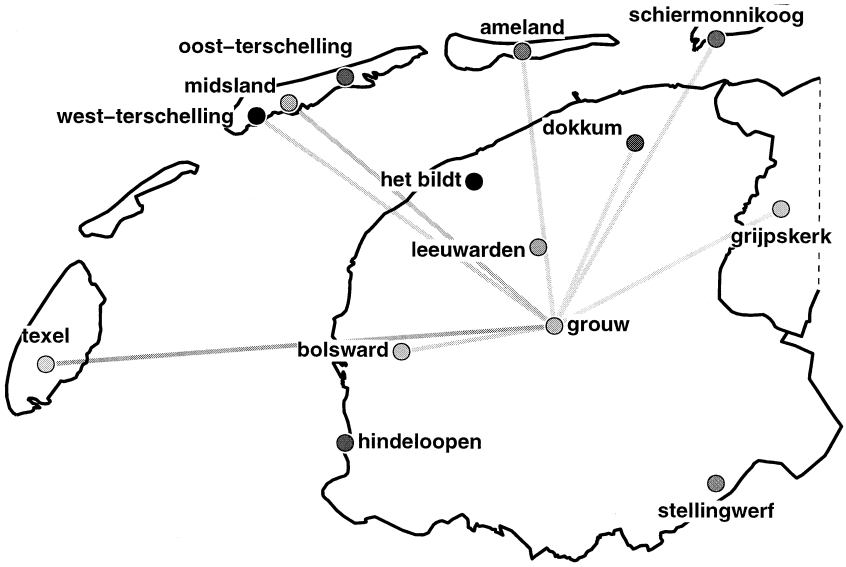

Figure 8. Divergence with respect to Standard Frisian. Darker lines indicate stronger convergence while lighter dots represent greater distances between the old and the new variant.