

THE BELOVED MOTHER TONGUE

*Ethnolinguistic Nationalism in Small Nations:
Inventories and Reflections*

EDITED BY

Petra BROOMANS, Goffe JENSMA,
Hans VANDEVOORDE
and
Maarten VAN GINDERACHTER



PEETERS

LEUVEN – PARIS – DUDLEY, MA
2008

LINGUISTIC DISTANCE BETWEEN ‘FLEMISH’ AND NORWEGIAN AROUND 1900

SUBJECTIVE VIEWS AND OBJECTIVE MEASUREMENTS

Charlotte Gooskens and Renée van Bezooijen

1. *Introduction*

In this paper we will take the ideas on linguistic distance as expressed by Omer Watez in 1899 in his article *Noorsch en Vlaamsch* [Norwegian and Flemish¹] as a point of departure and try to determine to what extent they have an empirical basis. We will focus on Watez’ claim that Norwegian and Flemish are remarkably similar and we will quantify the linguistic distance between these two languages by means of modern computational methods. As a reference we will also include Dutch and other Germanic languages into the comparison.

2. *Omer Watez: A man of his time*

Omer Watez (1857–1935) (see Figure 1) was born in Schorisse, not far from Oudenaarde in the province of East Flanders in Belgium. Everything he did during his life was inspired by his love for his native Flanders. This led him to write a series of nature guides with detailed descriptions of the landscape and advice on walks in the region between Oudenaarde, Ronse and Geraardsbergen.² He was very much concerned with the Flemish natural environment, and today would be considered an environmentalist. Watez also felt a deep admiration for the character and way of life of Flemish peasants, which he found to be pure, simple, honest and good. This idealized view is contrasted to everything French, which he considered artificial and perverted, insincere and wicked. This opposition constitutes the main theme of many of his writings. A clear example is the novel *Veldbloemeken* (Field Flower) from the anthology *Jonge harten* (Young Hearts, 1925). It

¹ Flemish is an umbrella term for the dialects spoken in the Dutch-speaking area of Belgium. At the time there was no standard language in Flanders, see Willemyns and Daniëls, *Verhaal*.

² For example, Watez, *De Vlaamsche Ardennen*.

describes the tragic life of the Flemish girl Rina, who grows up in the paradisiacal Flemish countryside in blissful ignorance, in unison with the animals and flowers around her. This idyll is brutally disturbed by the corrupting effects of French civilization imparted to her when sent to a Frenchified boarding school. Rina falls victim to the treacherous charms of a French-speaking nobleman, who kills his mother and takes her money to live a life of decadence in faraway Paris. Rina's life, which started off so happily, ends when she drowns herself in a lovely Flemish stream.



1. Omer Wattez.

Wattez may be considered a late exponent of Romanticism. This appears not only in his sentimentality and idealization of nature and peasantry, but also from his glorification of the past and his love for his people. Wattez was an active participant in the Flemish struggle, which aimed at giving the Flemish language and culture its own identity within a bilingual nation-state. Wattez was convinced that Flemish literature could compete with that of other, larger countries, and he assumed that there was a strong relationship between literature, language, national character, nature and landscape. Here again he sees a sharp contrast between Germanic and Romance cultures:

Ik meen, dat de Zuidervolkeren niet gevoelen, gelijk wij, Germanen, om de ziel der Germaansche balladen-poëzie te verstaan. Hunne poëzie, gelijk hunne kunst in 't algemeen, is meest uiterlijke schijn. Onder het helle licht der Zuiderzon is

hun gezichtszin meer ontwikkeld dan hun gevoel. Zij schiepen de beeldhouwkunst, gelijk in Griekenland; niet de harmonie, of het gevoel door akkoorden uitgedrukt, ene kunst, die door Germanen tot den hoogsten top werd gevoerd.³

(I think that the southern peoples are not able to understand the soul of the Germanic ballads as can we, the Germanic peoples. Their poetry, like their art in general, is mostly outward appearance. Under the bright southern sun their power of vision is better developed than their feeling. They created the art of sculpture, as in Greece; not the harmony, or the feelings expressed by chords, an art that was brought to perfection by the Germanic peoples.)

It is clear that Watez felt much more affinity with the Germanic peoples and culture than the Romance peoples and culture. Watez wrote many articles in the periodical *Germania*, which appeared between 1898 and 1906 and was established to stimulate the exchange of ideas between Flanders and Germany, stressing the unity of race and descent. According to Elias,⁴ Watez propagated race theories most strongly and took the most non-conciliatory stance with respect to the Germanic-Romance cultural opposition. In one of his articles, Watez wrote:

Naar het Oosten, waar de dageraad blinkt, waar de zon, groots en prachtig opgaat, wend ik mijn blikken met het innigste verlangen. Wij zijn Vlamingen, dus Germanen. Willen wij dat ernstig zijn of niet? Geen aarzeling of vlucht. Onze toekomst als volk staat op het spel.⁵

(Towards the east, where daybreak shines, where the sun rises, grand and gorgeous, I turn my eyes with the utmost desire. We are Flemish and therefore Germanic. Do we want that seriously or not? No hesitation or flight. Our future as a people is at stake.)

Watez was an extreme adherent of Pan-Germanism in the sense of the psychological and cultural unity of all Germanic peoples. Watez pleads, for example, for the introduction of German as a second language in Flemish and Walloon education. On the other hand, and in apparent contradiction, he maintained support for the equality of both the Germanic/Flemish and Romance/Walloon 'races' within a unitary Belgium.

³ Watez, *Germaansche balladen*, p. 7.

⁴ Elias, *Geschiedenis van de Vlaamse gedachte*, p. 323.

⁵ Cited in Elias, *Geschiedenis van de Vlaamse gedachte*, p. 325.

3. Watez' views on the relationship between Flemish and Norwegian

The article *Noorsch en Vlaamsch* (Norwegian and Flemish), written by Watez in 1899, is a typical example of his vision of Flemish language and literature. In his opinion, modern writers are verbose and conventional. He considers that young talented people are spoiled through the suppression of their individuality, which has a disastrous effect on the quality of their art. Watez makes a plea for the use of an unpolished, original and natural literary style. This kind of writing can be found in Norwegian literature, which he sees as 'a refreshing bath' for civilized culture. He mentions writers such as Ibsen and Bjørnson, who are not afraid of 'provincialisms'. Watez offers the Norwegian text *Arne* by Bjørnson as an example. The sentences are short and the vocabulary is uncomplicated. This text had previously been translated into Dutch in verse form by Pol De Mont under the title *Distelvink* (Goldfinch). While the original Norwegian text consists of only one page, this translation amounted to as many as twelve pages. Watez also presents his own translation to demonstrate the unadorned and simple style which he prefers. To Watez, the Norwegians and their language represent everything that is authentic and unspoiled. He wants to show that the Flemish language has the same qualities.

Watez has a positive attitude towards Norwegian as a pure representative of the Germanic language family, and towards the Flemish vernacular as a central element of traditional Flemish culture. By linking Flemish to Norwegian and stressing the points they have in common, the value of Flemish within the Germanic language group is raised. Moreover, by pointing out that simple elements enhance the quality of Norwegian literature, Watez states that what some people would consider to be 'provincialisms' are suitable for use in a cultured, elevated style. If this applies to Norwegian, why would it not apply to Flemish? In contrast, Watez has an aversion to the Dutch, whom he finds to be arrogant, presumptuous, condescending and patronizing. His central views are reflected in the following statement:

Doch vooral is het voor den Vlaming opmerkenswaardig, zeg ik, dat hij in een taal, die vijf, zes honderd mijlen hooger in 't Noorden gesproken wordt, woorden en wendingen kan hooren, die door ons Vlaamsche volk gebruikt worden. En als wij die woorden in de letterkundige taal gebruiken, zegt men ons dat wij 'provincialismen' invoeren, die men in Noord-Nederland niet verstaat. In Noord-Nederland zou men niet verstaan of liever niet willen verstaan, wat men in 't hooge Noorden begrijpt!⁶

⁶ Watez, 'Noorsch en Vlaamsch', p. 91.

(But it is especially noticeable for the Flemish, I say, that in a language spoken five, six hundred miles further to the north, they hear words and phrases used by our own Flemish people. And when we use those words in the literary language, we are accused of importing 'provincialisms', which are not understood in the northern Netherlands. In the northern Netherlands, people would not understand or rather choose not to understand what is understood even in the far North!)

It is likely that Watzel's ideals and attitudes influenced his perception of language distance. This can be compared with claims about the *intelligibility* of a language (after all, intelligibility can be considered to be the communicative consequence of language distance). If one has a positive attitude towards a people and its culture, one will be motivated to try and understand its language, minimizing the differences and maximizing the similarities. On the other hand, a negative attitude will have an adverse effect on intelligibility. The relationship between attitude and perceived linguistic distance can be illustrated by means of two examples from the literature.

The first example is taken from Wolf,⁷ discussing communication between speakers of Nembe and Kalabari in the eastern Niger Delta. Nembe and Kalabari are so closely related that they can be considered dialects of the same language. Speakers of Nembe acknowledge the similarity and claim to have no problems understanding Kalabari. On the other hand, speakers of Kalabari state that Nembe differs from their own language to such an extent that, except for a scattered word, it is completely unintelligible to them. Wolf explains the asymmetry in terms of attitudinal factors. The Kalabari are the most prestigious, largest and economically prosperous group in the eastern Delta. They regard the Nembe, who live in small fishing villages, as inferior, poor country cousins. According to Wolf 'the intelligibility evidence merely seems to underscore Kalabari ascendancy.'⁸

A similar situation can be found in Scandinavia. Norwegian, Swedish and Danish are so closely related that one would expect them to be mutually intelligible. However, results of various investigations⁹ show that Danes understand spoken Swedish better than Swedes understand spoken Danish. Sweden is often called the 'big brother' of Denmark because of its political and economic dominance. For this reason Danes are assumed to have a positive attitude towards Swedish and to be willing to make an effort to understand Swedish. In contrast, to Swedes, successfully communicating with

⁷ Wolf, 'Intelligibility and interethnic attitudes'.

⁸ *Ibidem*, p. 442.

⁹ For example, Maurud, *Nabospråksforståelse i Skandinavia*; Bø, *Ungdom og naboland*; Börestam, *Dansk-svensk språkgemenskap på undantag*.

Danes would be of less importance. Studies have indeed confirmed that both Swedes and Danes have a more negative attitude towards the Danish language than towards Swedish.¹⁰

4. *Empirical evidence provided by Wattez*

To convince the reader of the close relationship between Flemish and Norwegian, Wattez wants to show that the two languages share many words and expressions which are lacking in northern Dutch. He gives twenty-five examples to prove his point. Ten examples involve a lexical difference, whereby Flemish and Norwegian share a word that is missing in Dutch. Seven examples are semantic in nature, which means that a word occurs in more or less the same form in all three languages, but with a deviant meaning in Dutch. Six examples pertain to the phonetic/phonological level, Flemish and Norwegian having a similar sound, which differs from Dutch. Finally, there is one syntactic and one morphological example.

Wattez does not define what exactly is being compared. With respect to Flemish, he alternately uses 'wij' (we), 'wij in de streek van Oudenaarde' (we in the area of Oudenaarde), 'Zuid-Vlaanderen' (southern Flanders), 'Vlaanderen' (Flanders), or 'hier' (here). For Dutch, he uses the terms 'Noord-Nederlands' (northern Netherlands), 'Nederlandsche boekentaal' (Netherlandic literary language) and the language in *Van Dale*, that is, the official dictionary for Standard Dutch. Norwegian is referred to with the general terms 'Noorwegen' (Norway) and 'Noors' (Norwegian) or illustrated by means of concrete examples from the contemporary literature, specifically tales and plays by Bjørnson and Ibsen.¹¹ In addition to being vague, the overview is incomplete in that Wattez often limits himself to giving the corresponding words or expressions in Flemish and Norwegian, leaving it to the reader to establish the relationship with Dutch. He states, for example: '*Krop* is bij hen de strot, gelijk hier.' (They [the Norwegians, CG and RvB] use *krop* in the sense of throat, just like we do).¹²

Despite the lack of clarity we have nevertheless attempted to assess the validity of Wattez' observations. As for Flemish, we trust that the informa-

¹⁰ Haugen, 'Semicommunication'; Maurud, *Nabospråksforståelse i Skandinavia; Bø, Ungdom og naboland*.

¹¹ Just like other contemporary Flemish writers, Wattez was probably wrong about Norwegian. Bjørnson and Ibsen's works are written in Dano-Norwegian, which was in fact Danish with Norwegian influences. For this reason we also looked at the distances between Flemish and Danish and between Dutch and Danish (see Section 5.2). See Doorns, *Een gezonde Noordervorst. Vlaamse visies op Noorse literatuur tussen 1830 en 1914*.

¹² Wattez, 'Noorsch en Vlaamsch', p. 93.

tion provided is correct.¹³ After all, here Wattez could draw on his own experience. We have checked his statements for Norwegian in the *Metaordboken*,¹⁴ *Nynorskordboka* and *Bokmålordboka*.¹⁵ The examples given for Dutch were checked in the *Woordenboek der Nederlandsche Taal* (WNT) and the *Etymologisch Woordenboek*.

Generally speaking, Wattez' observations concerning Dutch and Norwegian appear to be correct, at least from a contemporary point of view. However, there are a few exceptions. For example, Wattez claims that Norwegian *krop* means 'throat'. However, in the dictionary the only meaning given is 'body'. Also, Wattez states that Norwegian *stiv* has the same meaning as Flemish *stijf* (very) – as in Flemish *stijf schoon* (very handsome). However, we think that the example given 'Se mig stivt i øjene' (Look straight into my eyes)¹⁶ does not support this view. *Stivt* is not used here as an intensifier but rather as an adverb of manner, meaning 'intently' or 'hard'.

Even though most of Wattez' observations are correct, we nevertheless argue that his conclusion that Flemish is more similar to Norwegian than Dutch needs to be restricted in two ways. In the first place, it may only hold for the small set of selected items which he presents. We will come back to this below. In the second place, a distinction has to be made between a synchronic and a diachronic perspective. Wattez presents his examples as an absolute truth. Although he does not state so explicitly, his formulations suggest that what holds true in the present also held true in the past. In some cases Wattez might be right in this respect. For example, in Flemish the word *snel* has the meaning 'friendly' or 'attractive', just like Norwegian *snil*, while in Dutch *snel* means 'fast'. We have not found it to mean 'friendly' or 'attractive' in northern Dutch at any stage of its development. In addition, we have not been able to trace a Dutch word corresponding to Flemish *pijnten* and Norwegian *at pynte*, 'decorate'.

However, in many other cases the difference between Flemish/Norwegian and Dutch may apply to the period around 1900 but not to earlier stages of Dutch. We will give two examples. The first example pertains to Flemish *bescheid weten* 'to know the ins and outs'. Wattez notes that this expression exists in Norwegian in a similar form and meaning but is lacking in Dutch. However, in the *WNT* several examples can be found of the use of *bescheid weten* in northern Dutch. For example the Dutch author

¹³ As Wattez fails to make clear which (local) variety of Flemish his examples are taken from it is impossible to verify whether they are correct.

¹⁴ <<http://www.dok.hf.uio.no/perl/search/search.cgi?appid=7&tabid=571>>.

¹⁵ <<http://www.dokpro.uio.no/ordboksoek.html>>.

¹⁶ From Ibsen, *Et dukkehjem*.

G.A. Bredero, who was born in Amsterdam and lived there his entire life, wrote in *Klucht van Symen sonder soeticheyt* (Symen's Farce without Sweetness) from 1620: *Maer lieve Symen, segh mijn doch eens het rechte bescheyt ...* (But dear Symen, for once give me the true information ...).¹⁷ In modern Dutch we say *zekerheid hebben* or *de waarheid weten*. Secondly, Watez compares Flemish *taoke* (take) with Norwegian *at tage*, suggesting that a related form does not occur in northern Dutch. However, the verb *taken* did exist in the past. The Amsterdam playwright A. Leeuw, for example, used the word in his farce *Broers Hert* (Brother's Heart) from 1668: *De nikker mag die meidt wel taken* (The devil may well take that girl).¹⁸ The regular form in modern Dutch is *nemen*. In these two examples Flemish and Norwegian have preserved a meaning and use which once existed in northern Dutch as well. The same pattern emerges from other examples. It looks as if Flemish and Norwegian are more conservative, more stable than Dutch. Yet again, this may only apply to the small number of examples selected by Watez with the specific aim of showing correspondences between Flemish and Norwegian lacking in northern Dutch.

5. Alternative measurements on the basis of single words

Watez indeed succeeded in finding a number of Norwegian words which corresponded with Flemish but not with Dutch words. However, does this indeed prove that Flemish is more closely related to Norwegian than Dutch? Another selection of words might have shown the opposite – that Dutch is closer to Norwegian than Flemish. We constructed an alternative, unbiased set of words to investigate linguistic distances in a more objective way. These words were translated into East Flemish, Dutch and Norwegian, as well as other Germanic languages, which served as a reference. We sought answers to the following questions:

1. Is East Flemish more similar to Norwegian than Dutch?
2. Is East Flemish more similar to Norwegian than to other Germanic languages?
3. Is East Flemish more similar to Norwegian than other Flemish and Dutch dialects?

In Section 5.1 we will describe how we composed the speech material and measured the linguistic distances. The results will be presented in Section 5.2.

¹⁷ Source: *Woordenboek der Nederlandsche Taal*, II-13/1 (1899).

¹⁸ Source: *Woordenboek der Nederlandsche Taal* (Dictionary of the Dutch language), XVI-6 (1928).

5.1 *Method*

5.1.1 *Single words*

The twenty-five words presented by Watez were selected with the specific aim of demonstrating the similarity between Flemish and Norwegian. It is a heterogeneous and small sample, which is undoubtedly biased. We used a more neutral and larger set of words than had been collected in a previous study. These all referred to elementary objects and concepts of everyday life. Eleven domains were represented: parts of the house, objects in the living room and bedroom, objects in the kitchen, birds and insects, domestic animals, animals in ditches, fields and forests, food and drink, parts of the body, clothing, flowers, plants and trees, weather and landscape. Twenty-five women and men with varying socioeconomic backgrounds (mean age 39.6 years, range: 22–61) were asked to note down (in Dutch) the first ten objects and concepts that came to their mind for each domain. All the words mentioned by at least 10 of the 25 subjects (40%) were included in the sample. This yielded 88, mostly monomorphemic, nouns. To this set, nine nouns were added: 'house', 'bird', 'animal', 'ditch', 'field', 'forest', 'flower', 'plant' and 'tree'. These were not mentioned by the subjects because they were part of the headings of the domains presented to them. Thus, the total set of nouns that we constructed consisted of 97 words.

The 97 words were first recorded on tape using a native speaker of Standard Dutch. Subsequently, they were translated and recorded in native Standard Norwegian and in the dialect of Ouwegem, a village situated in the province of East Flanders approximately 18 km from Schorisse, the place where Watez was born (see Figure 2). These three versions constituted the core of the sample. In addition, nine reference versions were recorded, namely the five standard varieties of English, German, Frisian, Danish and Swedish, and the four non-standard varieties of De Panne (in the province of West Flanders in Belgium), 's-Heerenbroek (in Zeeland, a southwestern province of the Netherlands), Maastricht (in Limburg, a southeastern province of the Netherlands) and Uithuizen (in Groningen, a northeastern province of the Netherlands). The locations of the Flemish and Dutch varieties can be seen in Figure 2, where the Netherlands and the Dutch-speaking part of Belgium are shown. All recordings were transcribed using SAMPA.¹⁹ This is a machine-readable phonetic alphabet, which maps IPA symbols to seven-bit printable ASCII/ANSI characters.

¹⁹ See <<http://www.phon.ucl.ac.uk/home/sampa/>>.

5.1.2 Linguistic distance measurements

Linguistic distances between the twelve languages and language varieties were measured at the phonetic/phonological and lexical levels. Details are presented below.

Phonetic/phonological distance

The phonetic/phonological distance was established by means of the so-called Levenshtein algorithm. The Levenshtein distance measure, which can be computed automatically, expresses the phonetic/phonological dissimilarity between strings of symbols, in this case SAMPA-symbols denoting the pronunciation of words with the same semantic content in different languages and language varieties. The degree of dissimilarity is derived from the number of symbols that need to be inserted, deleted or substituted in order to transform one word into the other.

We applied the Levenshtein distance measure to all pairs of corresponding words in all combinations of languages and language varieties in our study. Insertions and deletions were assigned a ‘cost’ of 1.0 point, substitution of identical symbols 0 points, substitutions of a vowel by a vowel or a consonant by a consonant 0.5 of a point, and substitutions of a vowel by consonant or of a consonant by a vowel 1.0 point. Diacritics were joined with the preceding symbol, adding an extra 0.25 of a point. So, for example, the distance between [a] and [a:] was 0.25, that between [a] and [o] 0.5, and that between [o] and [a:] 0.75. The unwanted effect of word length was compensated for by dividing the total sum of costs by the number of symbol alignments. We refer to Heeringa²⁰ for a more extensive explanation of the Levenshtein algorithm.

As an example we present the calculation of the distance between the cognate words for ‘head’ in East Flemish [y:əft] and Dutch [ho:ft]. Converting the one word into the other involves the insertion of [h] with a cost of 1 point, the substitution of the lengthened vowel [y:] by the lengthened vowel [o:] with a cost of 0.5 of a point, and the deletion of [ə] with a cost of 1 point. The two phonemes /f/ and /t/ are identical, so no costs are involved.

| | | | | | |
|------------|---|-----|---|---|---|
| alignments | 1 | 2 | 3 | 4 | 5 |
| | | y: | ə | f | t |
| | h | o: | | f | t |
| costs | 1 | 0.5 | 1 | 0 | 0 |

²⁰ Heeringa, *Measuring dialect pronunciation differences using Levenshtein distance*.

The sum of costs (2.5 points) is divided by the number of alignments (5). The result is a distance of 50% between the two words in question. The total Levenshtein distance between two varieties is the mean distance over all 97 word pairs.

Lexical distance

The Levenshtein distances were calculated on the basis of the complete set of 97 words, including cognates (etymologically related words) and non-cognates (etymologically unrelated words). This means that some word pairs differed at the phonetic/phonological level, whereas others differed at the lexical level. Since most of Watzel's examples involved a lexical difference, we decided to also measure distances separately for this type of word pair, counting the number of non-cognates for all combinations of languages. As an illustration we give two examples of non-cognates:

- East Flemish Ouwegem *talure* corresponds with Dutch *bord* (plate)
- Dutch *rok* corresponds with Norwegian *skjørt* (skirt)

The lexical distance between two languages was expressed as the percentage of non-cognates across the 97 words.

5.2 Results

Phonetic/phonological distance

In Table 1 the results of the Levenshtein distance measurements are presented for all combinations of languages and language varieties. The results which bear directly on the three research questions formulated in Section 5 are marked in grey.

The first and most central question is whether East Flemish is indeed more similar to Norwegian than is Dutch. In our study, East Flemish is represented by the dialect of Ouwegem, and Table 1 shows that our results do not support Watzel's claim. On the contrary, the distance between Ouwegem and Norwegian is larger (49.0%) than the distance between Dutch and Norwegian (45.5%). However, the difference is not significant ($t = 1.835$, $df = 96$, $p = .070$).²¹ Thus, it must be concluded that statistically speaking East Flemish and Dutch are equally different from Norwegian at the phonetic/phonological level. The same pattern emerges when the distances are measured between Flemish and Danish and between Dutch and Danish (see note 11).

In order to place the results in perspective, we also assessed the phonetic/phonological distance between East Flemish and other Germanic lan-

²¹ The statistical significance was assessed with a post hoc analysis at the 0.05 level for the whole investigation.

guages. Was Watez right in focusing on Norwegian? Is there perhaps another language which is more similar? Table 1 reveals that there are indeed two languages with a smaller Levenshtein distance to Ouwegem than Norwegian (49.0), namely German (48.7) and Frisian (42.9). The distance to German is not significantly different, but the distance to Frisian is significantly different at the .01 level ($t = 2,790$, $df = 96$, $p = .006$). The other Germanic languages, namely Danish (58.2), Swedish (52.0) and English (56.4), show a larger distance. In two cases the difference is significant, namely for Danish ($t = -6.021$, $df = 96$, $p = .000$) and for English ($t = 4.336$, $df = 96$, $p = .000$). These results suggest that Watez could just as well have given examples of correspondences between Flemish and German or Flemish and Swedish. On the other hand, we do not know whether in such a case he would have been able to find examples which deviated from Dutch.

Finally, we also compared the phonetic/phonological distances between Ouwegem and Norwegian to the distances of other Dutch and Flemish dialects to Norwegian. Is Watez right in suggesting that East Flemish is a special case in being exceptionally similar to Norwegian? The data in Table 1 do not confirm Watez' view. None of the dialects are significantly more distant or less distant to Norwegian than the dialect of Ouwegem.

An insightful way of analysing the results is cluster analysis. The goal of clustering is to identify groups of similar variables or objects, in our case language varieties. These groups are called clusters. Clusters may consist of subclusters, which may in turn consist of subsubclusters. The result is a hierarchically structured tree (a dendrogram), in which the language varieties are the leaves. There are several ways of clustering the data. Our dendrograms (see Figure 3) are based on squared Euclidian distances with median linkage.

Figure 3 visualizes what was already concluded above. There are no traces of a close relationship between East Flemish and Norwegian. East Flemish Ouwegem is linked at a low level with West Flemish from De Panne and 's-Heerenbroek in Dutch Zeeland, and at a higher level with the other Dutch varieties. At a much later stage German was also added to this group of West Germanic languages. On the other hand, Norwegian is linked with Swedish and, at a higher level, with Danish, thus constituting the group of North Germanic languages. Apparently, the phonetic/phonological properties of English are so idiosyncratic that this language constitutes a branch of its own.

Lexical distance

Although the percentage of non-cognates is a rougher and arithmetically much simpler measure than the Levenshtein distance, it reveals the same pattern of relationships. First, Dutch is slightly more similar to Norwegian

than East Flemish ($t = 1,835$, $df = 96$, $p = .07$). Second, German and Frisian are more similar to Ouwegem than Norwegian, whereas Danish, Swedish and English are less similar. However, in this case none of the differences is significant. Third, as for the Levenshtein distances, none of the Dutch and Flemish dialects are significantly closer or more distant from Norwegian than the dialect of Ouwegem.

The correlation between the mean Levenshtein distances (Table 1) and the percentages of non-cognates (Table 2) is high ($r = 0.87$). As with the phonetic/phonologically based dendrogram, the lexically based dendrogram (Figure 4) shows a division into a North Germanic branch, a West Germanic branch, and English. Ouwegem is linked with De Panne at the same level, and Norwegian with Danish. Some other varieties, however, have changed places. For example, Frisian is now clustered with Uithuizen, whereas previously it was not linked with any of the Flemish or Dutch varieties in particular.

6. *Alternative measurements on the basis of a running text*

The list of 97 single words that formed the basis of our first set of linguistic distance measurements has the advantage of being already available in a transcribed form for a great number of languages. We only had to collect additional versions for Norwegian, Danish and Swedish. However, there are some disadvantages. In the first place, the recordings were relatively recent and the speakers relatively young, so that the present language situation rather than the situation described by Watzel is reflected. Moreover, connected speech phenomena such as assimilation processes at word boundaries are left out of consideration, for example, Dutch *dat zal >* [dAt sAl] (that will). Finally, on the basis of single words – mostly monomorphemic nouns in the singular – only lexical and phonological-phonetic differences can be taken into account, whereas languages may differ at the morphological (for example, noun plurals, diminutives, adjective inflection, verb endings) and syntactic levels as well. A typical syntactic feature of East Flemish, for example, is the use of double negation such as *dat is n'ier den emel nie* (that is [NEG] here the heaven not) and reduplication of the personal pronoun such as *'k ebbe 'k ik* (I have I I).

In view of these shortcomings, we decided to compute a second set of linguistic distance measurements, this time based on a running text reflecting language use around 1900. We limited the material to East Flemish, Dutch and Norwegian, which means that we focused on the question of whether Flemish is more similar to Norwegian than Dutch and abstained from considering the relationship with other Germanic languages and dia-

lects. In Section 6.1 we will describe the nature of the text and the linguistic distance measurements. In Section 6.2 we will present the results.

6.1 *Method*

6.1.1 *Running text*

Fortunately, for East Flemish we could make use of an existing and easily accessible collection of recordings, namely 'de Taalkamer' (the Language Room).²² This collection comprises speech samples of 64 dialects spoken in the Dutch-speaking part of Belgium, northern France (French Flanders) and the southern part of the Netherlands (Zeeland Flanders). In selecting the most suitable East Flemish dialect fragment, we applied four criteria. Ideally the recording should:

1. be of a dialect spoken not too far from the place where Watez grew up (Schorisse, see Figure 2)
2. reflect dialect use around the beginning of the twentieth century
3. have a good technical quality allowing a reliable transcription
4. consist of a coherent story translatable into Standard Dutch and Norwegian.

The application of these criteria led to the recording of the dialect of Geraardsbergen, entitled *Is dat hier den hemel?* (Is this heaven?). Geraardsbergen is situated in the province of East-Flanders, approximately 19 km from Watez' birthplace (see Figure 2). The speaker, named Jules, had been selected because he was considered in the local community to be a good dialect speaker. Both his parents originated from Geraardsbergen. He was born in 1896 and recorded in 1969 at the age of seventy-three. In line with common practice in sociolinguistics (cf. Chambers 1995:194) we assumed that his language remained fairly stable from the time he was young and thus would reflect the dialect characteristics at the beginning of the twentieth century. The recording is of remarkably good quality, and the speaker articulates well, so that the transcription presented very few problems. Moreover, a rough transliteration of the text was available in a special issue of the *Oost-Vlaamse Zanten. Tijdschrift voor Volkscultuur in Vlaanderen* (East-Flemish Saints. Journal of Popular Culture in Flanders), entitled

²² The audio recordings and written versions of the dialect fragments are available on cd-rom and can be obtained at De huis van Alijn, Kraanlei 65, 9000 Gent, Belgium or via e-mail: <huis.aliyn@gent.be>.

‘Taalkamer’.²³ Finally, the speaker tells a highly structured and humorous story that recounts a dream in which he was on his way to heaven. The tale is easy to follow and all words and expressions can be identified. For the purpose of the present study we selected three passages, one from the beginning, one from the middle, and one from the end of the story. In total, the selected speech sample consists of 465 words.

The text of the tale was translated into northern Standard Dutch by the second author (who was born in Flanders) and checked by one of the editors of the *Flemish Dialect Dictionaries* produced at Ghent University.²⁴ To make sure that the text contained no words or expressions atypical of the beginning of the twentieth century, we looked up all content words in Van der Sijs²⁵ or the *Woordenboek der Nederlandsche Taal* to see when they first appeared in the Dutch language. Moreover, three other books were a great help to us, namely Zwaardemaker and Eijkman,²⁶ the first complete manual on the phonetics of Standard Dutch, Van de Velde,²⁷ a meticulous phonetic analysis of older and newer radio recordings placed within a historical framework, and Hulshof,²⁸ a detailed linguistic description of the nature of both the written and spoken language around the turn of the century. The Dutch text served as the basis for the Norwegian version and a similar procedure was adopted to verify that the translation was congruent with the language situation around 1900.²⁹ The three versions of the text were transcribed into SAMPA in the same way as the single word lists.

6.1.2 Linguistic distance measures

The three versions of the text were aligned as well as possible for all pairs of languages, that is, East Flemish with Dutch, East Flemish with Norwegian, and Dutch with Norwegian. In most cases, words were aligned with words, but sometimes one word in one language was matched with more than one word in the other language, for example in the case of the double negations and reduplicated pronouns mentioned in Section 6. At some points the structure of the texts diverged so much that no match could be

²³ *Oost-Vlaamse Zanten. Tijdschrift voor Volkscultuur in Vlaanderen*, pp. 219-221.

²⁴ We thank Jacques van Keymeulen for his helpful suggestions.

²⁵ Van der Sijs, *Chronologisch woordenboek*.

²⁶ Zwaardemaker and Eijkman, *Leerboek der fonetiek*.

²⁷ Van de Velde, *Variatie en verandering in het gesproken Standaard-Nederlands (1935-1993)*.

²⁸ Hulshof, ‘Nieuwnederlands (circa 1880-1920)’.

²⁹ We thank Arne Torp for translating the Flemish text into Norwegian.

found, thus yielding incomplete word pairs for which we developed a special procedure.³⁰

6.2 Results

The results of the distance measurements of the running text are presented in Table 3. Just as for the single words, the distance between East Flemish, this time represented by Geraardsbergen, and Norwegian is not significantly different from the distance between Dutch and Norwegian, either at the phonetic/phonological level as measured by means of the Levenshtein distance or at the lexical level, as measured by means of the percentage of non-cognates. So, again we find no support for Watzek's claim about the close relationship between Flemish and Norwegian.

7. Conclusion

Omer Watzek, born in East Flanders in 1857, was a true Romanticist, a zealous Pan-Germanist, and a confirmed supporter of the Flemish cause, who lived at the end of the nineteenth and the beginning of the twentieth century. He suffered deeply from French and northern Dutch dominance and wished to glorify a mighty, prosperous and culturally highly developed Flanders of the past, and very much wanted to raise the status of Flemish to a fully accepted language suitable for use in all domains of society, including literature. He saw his ideals reflected in the writings of famous Norwegian authors, who employed the stylistic naturalness which he preferred and who were not afraid of using simple words and regional expressions which reminded him of his beloved Flemish. The correspondences which he saw – and sought – between Flemish and Norwegian matched his ideal of a fully

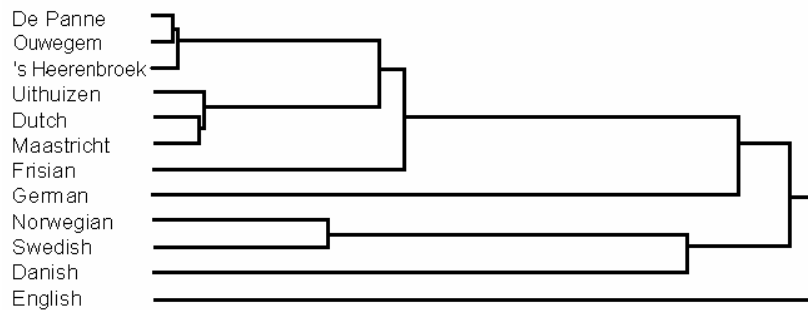
³⁰ The percentage of non-cognates was computed as the proportion of all complete word pairs. As for the Levenshtein distance, the fact that some words in one text have no equivalent in the other increases the distance between the two texts. In principle, the Levenshtein distance between a missing word and a word in a word pair has a maximum value of 100%, because all symbols of the one word are deleted (or inserted depending on the point of view), see Section 5.1.2. However, we thought that a cost of 100% would have too much of an effect on the overall mean. Moreover, as long words (mostly content words) in general contain more information than short words (mostly function words), we thought that the absence of the latter should have a smaller weight than the absence of the former. We therefore applied the following procedure. We divided 1 by the number of symbols in the word and subtracted this value from 1. For example, if the Dutch word *kijk* [keik] (look), with four symbols, has no equivalent in Flemish or Norwegian, this gives a distance of $100 \cdot (1 - 1/4) = 75\%$. If the Dutch word *de* [də] (the), with two symbols, has no equivalent in the other language, this has a smaller distance of $100 \cdot (1 - 1/2) = 50\%$.

blown Flemish language having equal status to the other Germanic languages which he admired and respected.

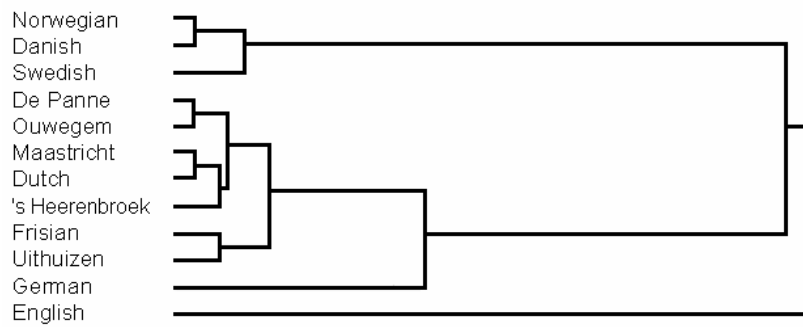
In our study we have tried to show that Watzel's idea of a particularly close linguistic relationship between Flemish and Norwegian was an illusion, based on selective perception and fed by his ideals. We measured phonetic/phonological and lexical similarity on the basis of a neutral set of 97 nouns and found the linguistic distance between Flemish and Norwegian not to be any smaller than the distance between northern Dutch varieties, including Standard Dutch, and Norwegian. Moreover, Flemish was not any closer to Norwegian than to other Germanic languages such as Frisian, German, or Swedish. In addition, when basing the linguistic measurements on a running text, thus including connected speech processes, morphology and, to a limited degree, syntax, Flemish was not found to be any closer to Norwegian than Standard Dutch.



2. The locations of the Flemish and Dutch varieties in the study with single words (indicated by ●). Geraardsbergen (indicated by ■) represents East Flemish in the running text study (see Section 6). The birth place of Watez, Schorisse, is indicated by a star. The thick line represents the border between the Netherlands (to the north) and Belgium (to the south), the thin lines represent the borders of provinces.



3. Dendrogram showing the phonetic/phonological similarity of the twelve language varieties on the basis of the Levenshtein distances in Table 1.



4. Dendrogram showing the lexical similarity of the twelve language varieties on the basis of the percentages of non-cognates in Table 2.

| | Norwegian | Danish | Swedish | English | German | Frisian | Uithuizen | Maastricht | 's Heerenbroek | De Panne | Dutch |
|----------------|-----------|--------|---------|---------|--------|---------|-----------|------------|----------------|----------|-------|
| Ouwegem | 49.0 | 58.2 | 52.0 | 56.4 | 48.7 | 42.9 | 42.9 | 41.3 | 27.8 | 23.9 | 35.8 |
| Dutch | 45.5 | 56.0 | 47.5 | 52.0 | 44.7 | 34.9 | 23.6 | 23.8 | 27.0 | 36.6 | |
| De Panne | 50.5 | 58.9 | 52.6 | 55.9 | 46.7 | 41.8 | 39.8 | 40.5 | 24.7 | | |
| 's Heerenbroek | 50.8 | 60.1 | 52.3 | 56.3 | 46.3 | 39.7 | 35.5 | 35.7 | | | |
| Maastricht | 48.3 | 57.4 | 50.5 | 52.9 | 47.0 | 40.2 | 32.1 | | | | |
| Uithuizen | 48.3 | 57.3 | 49.4 | 55.2 | 48.6 | 35.6 | | | | | |
| Frisian | 46.3 | 55.7 | 49.8 | 53.0 | 47.9 | | | | | | |
| German | 51.1 | 56.4 | 51.4 | 52.9 | | | | | | | |
| English | 53.5 | 59.5 | 55.6 | | | | | | | | |
| Swedish | 30.5 | 48.4 | | | | | | | | | |
| Danish | 38.0 | | | | | | | | | | |

Table 1. *Single words. Levenshtein distances (%) between all combinations of languages and language varieties. Higher values indicate larger phonetic/phonological distances. The results directly relevant to the three research questions (see Section 5) are marked in grey. Ouwegem represents East Flemish. See Figure 2 for the locations of the Dutch and Flemish dialects.*

| | Norwegian | Danish | Swedish | English | German | Frisian | Uithuizen | Maastricht | 's Heerenbroek | De Panne | Dutch |
|----------------|-----------|--------|---------|---------|--------|---------|-----------|------------|----------------|----------|-------|
| Ouwegem | 38.7 | 40.7 | 40.7 | 48.4 | 30.4 | 24.7 | 25.3 | 13.4 | 13.4 | 9.3 | 15.5 |
| Dutch | 37.6 | 37.6 | 40.7 | 47.4 | 27.8 | 16.0 | 11.9 | 8.2 | 11.3 | 13.9 | |
| De Panne | 40.2 | 40.2 | 42.3 | 49.5 | 26.3 | 22.2 | 21.1 | 10.9 | 13.9 | | |
| 's Heerenbroek | 43.3 | 43.3 | 45.4 | 52.1 | 34.0 | 21.6 | 20.1 | 11.3 | | | |
| Maastricht | 39.7 | 39.7 | 41.7 | 46.9 | 24.2 | 18.0 | 16.0 | | | | |
| Uithuizen | 43.3 | 43.3 | 46.4 | 51 | 28.3 | 16.0 | | | | | |
| Frisian | 47.4 | 47.4 | 48.0 | 51.5 | 34.5 | | | | | | |
| German | 37.6 | 36.6 | 38.7 | 50.5 | | | | | | | |
| English | 45.9 | 51.0 | 50.0 | | | | | | | | |
| Swedish | 17.5 | 18.6 | | | | | | | | | |
| Danish | 6.7 | | | | | | | | | | |

Table 2. *Single words. Percentage of non-cognates for all combinations of languages and language varieties. Higher values indicate larger lexical distances. The results directly relevant to the three research questions (see Section 5) are marked in grey. Ouwegem represents East Flemish. See Figure 2 for the locations of the Dutch and Flemish dialects.*

| | Geraardsbergen – Norwegian | Dutch – Norwegian | Geraardsbergen – Dutch |
|-----------------------|----------------------------|-------------------|------------------------|
| phonetic/phonological | 62.3 | 61.3 | 30.7 |
| lexical | 54.6 | 53.4 | 18.7 |

Table 3. *Running text. Levenshtein distances and percentages of non-cognates between the dialect of Geraardsbergen (East Flemish), Dutch and Norwegian. Higher values express larger distances.*

CONTENTS

| | |
|--|------|
| Preface and Acknowledgements | vii |
| Introduction <i>Petra Broomans, Goffe Jensma, Hans Vandevoorde and Maarten Van Ginderachter</i> | ix |
| Contributors | xiii |
| How Useful is the Concept of Ethnolinguistic Nationalism? On Imagined Communities, the Ethnic-Civic Dichotomy and Banal Nationalism <i>Maarten Van Ginderachter</i> | 1 |
| Linguistic Geopolitics and the Problem of Cultural Nationalism <i>Joep Leerssen</i> | 15 |
| The Concept of Ethnolinguistic Nationalism and Cultural Transfer. The Reception of Scandinavian Literature in Flanders and the Netherlands <i>Petra Broomans</i> | 37 |
| From Language to Nationality. The Case of the Dutch-speaking Belgians in the Nineteenth Century <i>Bruno De Wever</i> | 49 |
| Minorities and Kinships. The Case of Ethnolinguistic Nationalism in Friesland <i>Goffe Jensma</i> | 63 |
| Language Policy as a Cornerstone of Nation-Building. Linguistic Identities in the Nordic Countries <i>Godelieve Laureys</i> | 79 |
| Language, National Culture and the Clergy in Nineteenth-Century Denmark <i>Nikolaj Bijleveld</i> | 93 |

| | |
|--|-----|
| The Song of the Mother Tongue. Swedophone Finland as Text <i>Clas Zilliacus</i> | 109 |
| Linguistic Distance between 'Flemish' and Norwegian around 1900. Subjective Views and Objective Measurements <i>Charlotte Gooskens and Renée van Bezooijen</i> | 121 |
| The Power of Illusion. Flemish Writers on Germany, 1870-1914 <i>Greet Draye</i> | 139 |
| The Netherlands: an Ethnic Nation in spite of Itself <i>Marnix Beyen</i> | 157 |
| Racial-Linguistic Nationalism. Race, Language and Nation at the Antwerp Universal Exhibition of 1894 <i>Hans Vandevoorde</i> | 171 |
| Conclusion | 187 |
| Bibliography | 191 |
| Index | 209 |



PEETERS

PEETERS - BONDGENOTENLAAN 153 - B-3000 LEUVEN