

Interlingual text comprehension: linguistic and extralinguistic determinants

Renée van Bezooijen: Dept. of Linguistics, Radboud University Nijmegen,
P.O.Box 9103, 6500 HD Nijmegen, the Netherlands, r.v.bezooijen@let.ru.nl

Charlotte Gooskens: Dept. of Scandinavian Studies, University of Groningen,
P.O.Box 716, 9700 AS Groningen, the Netherlands, c.s.gooskens@rug.nl.

Correspondence should be sent to Charlotte Gooskens

Running head: Interlingual text comprehension

Interlingual text comprehension: linguistic and extralinguistic determinants

Renée van Bezooijen en Charlotte Gooskens*

Abstract

The three West-Germanic languages Dutch, Frisian and Afrikaans are so closely related that they can be expected to be mutually intelligible to a large extent. In the present investigation, we established the intelligibility of written Afrikaans and Frisian by Dutch-speaking subjects. It appeared that it is easier for speakers of Dutch to understand Afrikaans than Frisian. In order to explain the results, attitudes as well as linguistic distances were assessed. There was no evidence of a relationship between attitude and intelligibility. Three linguistic distances did show a relationship with reading comprehension, namely the number of non-cognates, the transparency of the lexical meaning of cognates, and the Levenshtein distance, which calculates the similarity between the written forms of words.

1 Introduction

1.1 Aim of the research

When two persons speaking different languages meet, there are three possibilities. One speaker switches to the language of the other, both persons take recourse to a third language, or both persons keep using their own language. The third type of communication, which we will refer to with the term ‘receptive multilingualism’ (Braunmüller & Zeevaert 2001), can only be used when the languages involved are sufficiently similar. Receptive multilingualism offers many advantages, especially on the production side. People can express themselves more easily and more precisely in their mother tongue than in a later acquired language. The question is, of course, what the effect is on the reception side. How much of an effort is needed to understand the spoken or written message? What is the risk of communication breaking down? What exactly are the causes of communication failure?

Research into receptive multilingualism has a long tradition. In the 1950s the mutual intelligibility of American Indian languages was studied (Hickerton, Turner & Hickerton 1952, Pierce 1952, Wolff 1959). More recently much attention has been paid to the communication among Scandinavians (e.g. Haugen 1966, Maurud 1976, Börestam Uhlmann 1994, Zeevaert 2004). Other languages that have been investigated include Spanish and Portuguese (Jensen 1989) and Slovakian and

Czech (Budovičová 1987). The present study focuses on the intelligibility of Frisian and Afrikaans for speakers of Dutch. There have been two previous studies of the intelligibility of Frisian by Dutch-speaking subjects (Van Bezooijen & Van den Berg 1999a, 1999b), but only in their spoken form, not in the written mode. The intelligibility of Afrikaans for speakers of Dutch has never been investigated experimentally. Here we focus on the relative intelligibility of written Frisian and Afrikaans.

Both Frisian and Afrikaans are related to Dutch. However, the historical backgrounds of the relationship are different (see Section 1.2). As a consequence, the present-day linguistic relationship between Dutch and Frisian deviates in many respects from that between Dutch and Afrikaans. These linguistic differences can be expected to affect the relative intelligibility of the two languages. Therefore, the first factor we will consider in order to explain intelligibility is linguistic distance. In addition, we will look at the possible role of attitudes. It is often contended in the literature that a positive attitude towards a language will motivate people to try and understand that language, whereas a negative attitude will hinder intelligibility.

So, the present study addresses the following three questions:

- (1) Which language is more difficult to understand for Dutch-speaking readers, Frisian or Afrikaans?
- (2) Can the difference in intelligibility, if any, be explained by different attitudes towards the two languages?
- (3) Can the difference in intelligibility, if any, be explained by differences in the linguistic distances to the two languages?

Before going into the research method and the results, we will first provide some background information on Frisian and Afrikaans.

1.2 Frisian and Afrikaans

The present study deals with the variety of Frisian as spoken in the province of Friesland in the Netherlands. In the literature, this variety is sometimes referred to as West-Frisian, to distinguish it from the North-Frisian and East-Frisian varieties spoken in Germany. In this paper we will simply use the term 'Frisian'. Frisian is the second official language of the Netherlands, in addition to Dutch. It is spoken by about 350.000 people. Historically, Frisian is closely related to English. In the course of time, however, Frisian has become increasingly similar to Dutch. For centuries, Frisian was almost exclusively a spoken language. From the middle of the nineteenth century the status of Frisian slowly started to rise. This was largely due to the efforts of the Frisian movement, which strived for the acceptance of Frisian as a fully fledged language, fit to be used for all communicative purposes, including cultural expression. This goal was never reached completely, due to the pervasiveness of Dutch in Frisian society. Nowadays, Frisian is used more often in

the lower than in the higher social strata, more often in the countryside than in the towns, and more often in the informal than in the formal domains. Frisian is taught to a limited degree in education. Only a small minority of the inhabitants of Friesland (17%) can write Frisian, but most (64%) report that they are able to read Frisian (Gorter & Jonkman 1995). Due to the dominance of Dutch in the media, education, and administration, Frisian loses more and more of its typical characteristics.

Originally, Afrikaans was a dialect that developed among a small group of Dutch colonists who settled in South Africa at the beginning of the seventeenth century.¹ In the course of time its nature changed, among others because it was largely used by non-native speakers with an insufficient command of Dutch. In the beginning, Afrikaans was mainly a spoken language, but in the nineteenth century it started to be used in writing. In 1921 Afrikaans was acknowledged as a separate language from Dutch. At present, there are ten additional languages in South Africa with an official status. According to the last census of 1996, Afrikaans is spoken by about 6 million people, both black and white. This is 14% of the South African population. The majority lives in the Western Cape (39%), Gauteng (21%), the Northern Cape (10%) and the Eastern Cape (10%). Afrikaans is the mother tongue of people of all social classes. Most speakers can be found in the urbanized areas. Afrikaans is taught and used at all educational levels. However, English has more prestige and is therefore gaining ground.

2 Method

To be able to answer the research questions formulated in 1.1, three types of data are needed. First, the intelligibility of Frisian and Afrikaans texts by speakers of Dutch has to be assessed (2.1). Second, the attitudes of speakers of Dutch towards Frisian and Afrikaans have to be determined (2.2.1). Third, the linguistic distances between Dutch and Frisian and between Dutch and Afrikaans have to be measured (2.2.2).

2.1 Intelligibility

Subjects

Twenty native Dutch language students (2 men and 18 women) from the Radboud University Nijmegen and the University of Groningen were selected as subjects. They had no active knowledge (speaking or writing) of Frisian and Afrikaans and no or only very limited passive knowledge (hearing or reading). Their mean age was 23 years.

Task

Intelligibility was assessed by means of a variant of the cloze test.² As a basis we used two Dutch newspaper articles with an average level of difficulty.³ One article ('the dating text') was about dating agencies and comprised 329 words; the other ('the feminist text') dealt with the image of women created by modern music stations and consisted of 317 words. In either article, five nouns, five adverbs, five adjectives, and five verbs were selected at random. These were placed in alphabetic order above the text and replaced by blanks in the text. Next, the two texts were translated into Frisian and Afrikaans and the same words were removed and placed above the texts. The subjects were given ten minutes to put the 20 words back in the right place in the texts. The percentage of words placed back correctly was taken as a measure of text comprehension.

Design

All subjects were tested in both languages. In order to avoid order effects, subjects that got the Frisian feminist text got the Afrikaans dating text, and the other way around. Furthermore, half of the subjects started with Frisian, whereas the other half started with Afrikaans. So there were four test versions that were each administered to five subjects. All subjects first filled in a form enquiring about their origin, home language, age, and sex. Then they were given two blocks, with respect to the first and the second language tested, respectively. Both blocks contained a set of questions about the subject's attitudes towards the language at hand (see 2.2.1) and the intelligibility test.

2.2 Measures to explain intelligibility

2.2.1 Attitudes

Before carrying out the intelligibility test, subjects first answered some questions that aimed at gaining insight into their attitudes towards the two languages. Not only the attitudes towards the languages themselves were probed, but also the attitudes towards the speakers of the languages and the countries where the languages are spoken. The idea behind this is that language attitudes may be influenced by social connotations (Trudgill & Giles 1978). Subjects noted down their responses on six five-point scales. As an example we present the questionnaire related to Frisian; the same questionnaire had to be filled in for Afrikaans.

Does Friesland appeal to you as a vacation destination?						
very much	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	not at all
What is your impression of Frisians?						
likeable	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	not likeable
intelligent	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	not intelligent
reliable	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	not reliable
What do you think of the Frisian language?						
beautiful	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	ugly
Would you like to learn Frisian?						
very much	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	not at all

2.2.2 Linguistic distances

To calculate the linguistic distances, we first aligned the Dutch texts with the Frisian translations and with the Afrikaans translations. In a few cases we adapted the word order to obtain a better alignment. The aligned word pairs formed the basis for seven distance measures. Measures A, B, and C express the nature of the relationship between the Dutch and the Frisian/Afrikaans words, whether they are cognates or not. Measures D, E, and F express the transparency of the relationship from the viewpoint of the Dutch reader. And measure G, the so-called Levenshtein distance, expresses the degree of phonetic similarity between corresponding words. Because we assumed that lexical words (nouns, adjectives, numerals, main verbs) are more important for intelligibility than function words (articles, conjunctions, prepositions, pronouns, auxiliaries, modals, particles, adverbs), distances were calculated separately for these two word categories. The distance measures will now be explained in further detail.

A. Percentage of cognates

A large proportion of cognates, i.e. words in the two corresponding texts with a common root, can be expected to facilitate comprehension. However, a direct relationship is not a necessary condition for mutual intelligibility. In some cases, the meaning of a word can be deduced via a cognate synonym. For example, the Dutch word *samenleving* ('society') in the original newspaper article was translated in the Frisian text with *maatskippij*. These two words are non-cognates. Nevertheless, the Dutch reader can easily understand the Frisian word *maatskippij* because of the existence of the Dutch synonym *maatschappij*. The original Dutch text could just as well have contained this word. The percentage of cognates, either related directly or via a synonym, constitutes the first linguistic distance measure.

B. Percentage of cognates related via a paradigm

It is also possible to deduce the meaning of a word paradigmatically. For example, the Frisian

translation of the Dutch word *zijn* (third person plural present tense of the verb *to be*) is *binne*. These two words are not related, not directly and not via a synonym. However, a Dutch reader may nevertheless understand the meaning of the Frisian word because it is related to the Dutch word *ben* (first person singular present tense of the verb *to be*), which belongs to the same paradigm as *zijn*. The percentage of cognates related via a paradigm constitutes the second distance measure. Most of the words in this category are function words.

C. Percentage of non-cognates

It should be impossible to deduce the meaning of a word in an unknown language if it bears no formal relationship with the corresponding word in the mother tongue. The percentage of non-cognates should therefore be an important indicator of (the impossibility of) mutual intelligibility. It is the complement of the first two measures, but for the sake of completeness we present all three.

D. Transparency of the lexical meaning

As stated above, it is impossible to deduce the meaning of a non-cognate (measure C). In the case of cognates, however, deducibility varies. In the course of time, the phonology or spelling of cognate words may have changed to such an extent that to a reader with no historical-linguistic background the relationship is no longer transparent. We therefore included a measure expressing the transparency of lexical meaning. The scoring was done by the first author and compared to the independent scoring of a second linguist. The percentage of identical scores was 82.2% (Pearson's $r = .92$) for the two Frisian texts and 90.1% (Pearson's $r = .92$) for the two Afrikaans texts. On the basis of these results we concluded that the scores were sufficiently reliable. In the scoring the following four grades of transparency were distinguished.

- Meaning completely transparent (0 points)

When two cognates have an identical form, there is no recognition problem. In this case a score of 0 points is assigned. Example: Afr. *uitbuiting* vs. Du. *uitbuiting* 'exploitation'. Cognates with a small spelling difference not leading to a difference in pronunciation were also assigned a score of 0 points. Example: Fri. *buro's* vs. Du. *bureaus* 'bureaus'.

- Meaning fairly transparent (1 point)

A score of 1 point is assigned whenever two cognates are so similar that the reader can be assumed to recognize the relationship fairly easily. In most cases there is a difference in only one letter. Example: Afr. *sewentig* vs. Du. *zeventig* 'seventy'.

- Meaning rather untransparent (2 points)

A score of 2 points is assigned whenever two cognates have so little in common that it will be quite difficult for a Dutch reader to recognize the relationship. Usually several letters will be different. Example: Fri. *jierren* vs. Du. *jaren* 'years'.

- Meaning completely untransparent (3 points)

A score of 3 points is assigned if the two cognates bear so little resemblance that it must be (virtually) impossible for a Dutch reader to see the relationship. In most cases the majority of letters will differ. Example: Afr. *hê* vs. Du. *hebben* 'have'.

Distance measure D was calculated by averaging the total number of points over the total number of word pairs; it can vary between 0 and 3.

E. Transparency of the grammatical meaning

In addition to the lexical meaning, a correct interpretation of the grammatical meaning of a word is also a prerequisite for text comprehension. Is it clear which word category (noun, verb, etc.) is involved? Is it clear what tense, number, gender, person, etc. is involved? An example of the first type of problem is Afr. *die*, which Dutch readers will be inclined to interpret as a demonstrative pronoun rather than as a definite article. An example of the second type of problem is Afr. *is*, which will be interpreted by Dutch readers as a singular form of the verb 'to be', whereas in Afrikaans the form is also used for the plural. If the grammatical meaning was transparent, 0 point was assigned; if it was not transparent 1 point was assigned. The two types of grammatical transparency were scored separately and summed. The scores thus range between 0 and 2.

F. Transparency of the total meaning

To determine the total transparency of the word meaning, measures D and E were summed. The scores can range between 0 and 5.

G. Levenshtein distance

Measures D, E, and F are to some extent subjective and also very time consuming to compute, as the transparency has to be estimated separately for each pair of corresponding words. The advantage of the Levenshtein distance is that similarity between word forms can be computed objectively and automatically. In the present study, the simplest variant of the Levenshtein measure was opted for. Distance was determined on the basis of the minimum number of letters that need to be inserted, deleted, or substituted in order to transform the word in the one language into the other. All three operations were given an equal weight of 1 point. Word length was compensated for by dividing the total sum of costs by the number of alignments of letters (see Heeringa (2004) for an extensive explanation of the nature and application of Levenshtein distances). As an example we present the calculation of the distance between the Dutch word *zoeken* and the Frisian word *sykje* 'search':

Alignment	1	2	3	4	5	6	7
Dutch	z	o	e	k	-	e	n
Frisian	s	y	-	k	j	e	-
Cost	1	1	1		1		1

The sum of costs (1+1+1+1+1=5) is divided by the number of alignments (7). The result is a distance of 0.71 or 71%. The total distance between two texts is the mean distance over all word pairs. The Levenshtein distances were calculated on the basis of the texts in which the original source word had been replaced by a synonym if a synonym was available (see Measure A).

3 Results

3.1 Which language is more difficult to understand, Frisian or Afrikaans?

To determine the relative intelligibility of Frisian and Afrikaans we counted the number of words that were placed correctly in the original sentence context. For greater stability we summed the data for the two texts. The mean percentage of correct responses appears to be considerably higher for Afrikaans (81.8% correct) than for Frisian (50.3%). The difference is significant ($t=36.12$, $df= 1,38$, $p=.00$). So, the answer to the first question is clear, written Frisian is more difficult for Dutch-speaking readers than Afrikaans.

3.2 Can the difference in intelligibility between Frisian and Afrikaans be explained by differences in language attitudes?

The mean ratings on the attitude scales are presented in Table 1. In a number of cases the difference between Frisian and Afrikaans is significant. The subjects

- find South Africa a significantly more attractive vacation destination than Friesland,
- judge South Africans to be significantly more intelligent than Frisians,
- are significantly more motivated to learn Afrikaans than Frisian,
- think Frisians to be significantly more reliable than South Africans.

Table 1. Mean attitude scores for Frisian and Afrikaans. The lower the score, the more positive the attitude.

Scale	Mean		<i>t</i> -test		
	Frisian	Afrikaans	<i>t</i>	<i>df</i>	sig.
Does Friesland/South Africa appeal to you as a vacation destination?	3.3	2.4	-2.7	19	.01
What is your impression of Frisians/South Africans?	2.6	2.4	-1.07	19	.30
likeable – not likeable					
intelligent – not intelligent	2.8	2.6	-2.18	19	.04
reliable – not reliable	2.6	3.1	3.25	19	.00
What do you think of the Frisian/Afrikaans language?	3.2	2.9	-1.67	19	.11
beautiful – ugly					
Would you like to learn Frisian/Afrikaans?	3.8	2.9	-4.50	19	.00

In general, the attitudes towards Afrikaans are more positive than those towards Frisian. This is in line with the results of the intelligibility test, as the subjects had fewer problems with

understanding Afrikaans than Frisian. To examine the relationship between attitudes and comprehension at the individual level, we correlated the two types of data of each subject. None of the correlation coefficients proved to be significant at the 5% level. We also computed (at the level of the subject) the correlation between the difference scores for the Frisian and Afrikaans texts and the difference scores between the attitude scores towards Frisian and Afrikaans. Again, no significant correlations were found. It can thus be concluded that there is no relationship between the attitudes of individual subjects towards Frisian and Afrikaans and their success in understanding the respective languages.

3.3 Can the difference in intelligibility between Frisian and Afrikaans be explained on the basis of the linguistic distance?

Assuming that intelligibility is inversely related to linguistic distance, we expected the linguistic distances between the Dutch and Frisian reading texts to be larger than between the Dutch and Afrikaans reading texts. In Table 2, the results of the first three linguistic distance measures are presented. We see that for the content words the percentage of cognates (related directly or via a synonym) is almost identical (94.1% versus 94.6%). However, there is a marked difference in the function words. Whereas in Frisian almost all function words (93.4%) are related directly to their Dutch counterparts, Afrikaans has relatively many function words (23.7%) that are related to Dutch via a paradigm. Furthermore (not shown in the table), we want to mention the fact that the nature of the cognates differs between Frisian and Afrikaans, in the sense that the Afrikaans texts have a larger proportion of words that are related via a synonym. The meaning of these synonyms sometimes slightly deviates from that of the original Dutch word. For example, the Dutch word *opeens* ‘suddenly’ was translated by Afrikaans *skielik*. Dutch readers probably interpret this word correctly via the Dutch synonym *schielijk*. However, Dutch *schielijk* is less frequent and slightly archaic compared to *opeens*. In quite a few cases the relationship with the Dutch counterpart appears to be more direct in Frisian than in Afrikaans. However, we do not know to what extent this affects intelligibility.

What we do know for sure is that a large proportion of non-cognates must have a negative influence on intelligibility. In this case the term non-cognates is used in the strict sense. It refers to those words in the Afrikaans or Frisian text which are not related to the corresponding words in the Dutch text and which have no related synonyms in Dutch. The percentage of non-cognates is higher for Frisian than for Afrikaans, both for function words (2.0% and 0.8%) and for content words (5.9% and 3.7%). This does not only hold for the number of tokens (24 and 14, respectively) but also for the number of types (20 and 10, respectively). These differences between Afrikaans and Frisian might seem small, but it should be recalled that one single unintelligible word can make a whole sentence or even a complete paragraph unintelligible.

Table 2. Percentage of Dutch-Frisian and Dutch-Afrikaans cognates, cognates via a paradigm, and non-cognates, for function words and content words separately and total.

		Frisian			Afrikaans		
		Function	Content	Total	Function	Content	Total
A.	Cognates	93.4	94.1	93.8	75.5	94.6	84.1
B.	Cognates via a paradigm	4.6	-	2.5	23.7	1.7	13.8
C.	Non-cognates	2.0	5.9	3.7	0.8	3.7	2.1

As stated above, not all cognates are equally transparent for the Dutch reader. In Table 3 the transparency scores are presented, separately for the lexical meaning (measure D), the grammatical meaning (measure E) and the complete meaning (F). The scores have also been converted to percentages to facilitate comparison among the three measures. For example, the deducibility of the lexical meaning of Frisian function words is $(1.24/3)100 = 41.3\%$.

Table 3. Mean transparency of the meaning of cognates, split up for function words and content words. Higher values denote lower transparency. Between brackets the scores converted to percentages. All differences between Frisian and Afrikaans are significant at the 1% level.

Transparency		Frisian			Afrikaans		
		Function	Content	Total	Function	Content	Total
D.	Lexical meaning (range 0-3)	1.24 (41.3)	1.45 (48.3)	1.33 (44.3)	0.40 (13.3)	0.63 (21.0)	0.50 (16.7)
E.	Grammatical meaning (range 0-2)	0.03 (1.5)	0.07 (3.5)	0.05 (2.5)	0.35 (17.5)	0.21 (10.5)	0.29 (14.5)
F.	Total meaning (range 0-5)	1.28 (22.7)	1.52 (30.4)	1.39 (27.8)	0.77 (15.4)	0.82 (16.4)	0.79 (15.8)

It can be seen in Table 3 that the lexical meaning of the Frisian cognates (1.33) is more difficult to deduce than of the Afrikaans cognates (0.50). This holds both for content words and function words. On the other hand, it is easier to deduce the grammatical meaning of Frisian words (0.05) than of Afrikaans words (0.29). However, the latter difference is smaller than the former. Also, it has to be assumed that the transparency of the lexical meaning contributes more to text comprehension than the transparency of the grammatical meaning. We therefore think that measure D forms part of the explanation of why Frisian is more difficult to understand than Afrikaans.

Finally, we also calculated the linguistic distance of Afrikaans and Frisian to Dutch by means of the Levenshtein distance (measure G). The Levenshtein measure expresses how many letters the Afrikaans and Frisian words differ from the corresponding Dutch words as a proportion of the total number of alignments (see 2.2.2). We only show the results for the cognates that are related directly or via a synonym (measure A). Any correspondence between the cognates that are related via a paradigm (measure B) and between the non-cognates (measure C) is largely based on chance and therefore not interesting. It appears from Table 4

that the Levenshtein distance (total distance) is higher for Frisian (34.2%) than for Afrikaans (20.9%). The difference is significant at the 1% level. This means that with respect to cognates the form of the Frisian words deviates significantly more from Dutch than the form of the Afrikaans words.

Table 4. Mean Levenshtein distance between Frisian-Dutch and Afrikaans-Dutch cognates (in %).

	Frisian			Afrikaans		
	Function	Content	Total	Function	Content	Total
G.	37.6	30.1	34.2	22.5	19.4	20.9

4 Conclusion and discussion

This study of interlingual text comprehension reveals that Dutch-speaking readers with no previous exposure to Frisian and Afrikaans are able to some extent to read Frisian and Afrikaans newspaper articles. There is, however, a considerable difference between the two languages. Reading Afrikaans presents few problems. In a cloze test, four out of the five words were placed back correctly in their original sentence context. Reading Frisian appears to be much more difficult, as not more than half of the words are placed back correctly. The results of this study suggest that in the written communication with Dutch-speaking people, South Africans can use (to a large extent) their own language, whereas this is not possible for Frisians. This pertains to informative texts of average difficulty read by linguistically trained students. Further research will have to show whether the same results are obtained for other types of texts and subjects.

In the literature, problems in interlingual communication are often attributed to the attitude of the receiver (listener or reader). It is assumed that the reported or measured comprehension problems are not so much due to a lack of transparency of the meaning of the language, but rather to a lack of motivation (see Wolff (1959) for various examples from West-Africa). Unfortunately, it is very difficult to determine experimentally whether it is a question of lack of ability or lack of willingness. In the present study the attitudes towards Afrikaans tended to be more positive than towards Frisian. It is unlikely that this has led to greater success in the cloze tests. First, there were no significant correlations at the individual level. So, subjects with a relatively positive attitude did not perform any better in the intelligibility test than subjects with a relatively negative attitude. Second, the subjects had had (virtually) no personal contact with (speakers of) Frisian and Afrikaans. Their reported attitudes probably reflect general, non-emotionally based, stereotypes. Third, the test was administered as part of a course in sociolinguistics. The students saw it as an interesting and challenging assignment. We observed no signs of an aversion towards the task.

We think that the difference in intelligibility between the two languages has mainly been caused by linguistic factors. A first factor is the number of 'true' non-cognates, i.e. Frisian or Afrikaans words that have no relation to the corresponding word in the Dutch text nor to a Dutch synonym. These occur more often in the Frisian texts than in the Afrikaans

texts. Not understanding a single, central term may have disastrous effects on text comprehension. A second factor which may have played a role is the fact that the Frisian cognates diverge more from Dutch than the Afrikaans cognates. This appears both from the objective Levenshtein measurements and from the subjective transparency scores. Finally, a third factor which may have worked to the advantage of Afrikaans is the spelling. The Afrikaans spelling is more similar to the Dutch spelling than the Frisian spelling. Especially, the spelling of the Frisian vowels, with a great number of unfamiliar diacritics and letter combinations (e.g. *ii, û, ú, ô, â, ê, ea, eo, oa, ue, iu, oai, uoi*) may be confusing to the Dutch reader (for more examples of differences between the Frisian, Afrikaans, and Dutch spelling systems, see Van Bezooijen & Gooskens 2005).

There are many morphological differences between Afrikaans and Dutch, but these are mainly simplifications. For a Dutch reader the simplified morphological system may be unusual, but in the end it may have little effect on text comprehension. A text generally contains so much redundancy grammatically, that the absence of explicit marking of, for example, number in the verb system presents no problems. After all, it does not present any problems to the Afrikaans-speaking readers either. The morphological differences between Frisian and Dutch are of a different nature. Here the Dutch readers are confronted with meaningful endings that they have never seen. Both the Frisian and Afrikaans texts contained few syntactical differences with Dutch, so in this respect both languages must have presented few problems.

Notes

* Renée van Bezooijen: Dept. of Linguistics, Radboud University Nijmegen, P.O. Box 9103, 6500 HD Nijmegen, the Netherlands, r.v.bezooijen@let.ru.nl

Charlotte Gooskens: Dept. of Scandinavian Studies, University of Groningen, P.O. Box 716, 9700 AS Groningen, the Netherlands, c.s.gooskens@rug.nl.

We would like to thank Vincent van Heuven, Sulette Bruwer, Jelske Dijkstra and Bart Alewijnse for their help in various stages of the investigation.

¹ The data for Afrikaans have been taken from Webb (2002).

² The cloze test was developed in 1953 in the United States by William Taylor and has been used extensively for measuring text comprehension. Sometimes the words are placed above the text, like in the present study. Alternatively, it may be left to the subjects to think of suitable words.

³ We determined the readability of the text by means of the so-called LIX-index (Björnsson 1968). This index consists of the mean number of words per sentence plus the percentage of words exceeding seven letters. Texts with a value of between 35 and 44 have a mean degree of difficulty. The mean LIX-value for the two Dutch texts was 42.

References

- Björnsson, C. H. 1968. *Läsbarhet* [Readability]. Stockholm: Liber.
- Börestam Uhlmann, U. 1994. *Skandinaver samtaler. Språkliga och interaktionella strategier i samtal mellem danskar, norrmän och svenskar* [Conversations between Scandinavians].

- Linguistic and interactional strategies in conversations between Danes, Norwegians and Swedes]. Uppsala: Uppsala University.
- Braunmüller, K. & Zeevaert, L. 2001. *Semikommunikation, rezepitive Mehrsprachigkeit und verwandte Phänomene. Eine bibliographische Bestandsaufnahme*. Arbeiten zur Mehrsprachigkeit – Folge B, nr. 19. Hamburg University.
- Budovičová, V. 1987. “Literary languages in contact (A sociolinguistic approach to the relation between Slovak & Czech today).” In *Reader in Czech sociolinguistics*, J. Chloupek, J. Nekvapil et al. (eds), 156-175. Amsterdam/Philadelphia: John Benjamins.
- Gorter, D. & Jonkman, R. 2005. *Taal yn Fryslân op 'e nij besjoen* [Language in Friesland reconsidered]. Leeuwarden: Fryske Akademy.
- Haugen, E. 1966. “Semicommunication: The language gap in Scandinavia.” *Sociological Inquiry* 36: 280-297.
- Heeringa, W. 2004. *Measuring dialect pronunciation differences using Levenshtein distances*. Groningen: Groningen dissertations in linguistics (Grodil).
- Hickerton, H., Turner, G.D. & Hickerton, N.P. 1952. “Testing procedures for estimating transfer of information among Iroquois dialects and languages.” *International Journal of American Linguistics* 18: 1-8.
- Jensen, J.B. 1989. “On the mutual intelligibility of Spanish and Portuguese.” *Hispania* 72: 849-852.
- Maurud, Ø. 1976. *Nabospråksforståelse i Skandinavia. En undersøkelse om gjensidig forståelse av tale- og skriftspråk i Danmark, Norge og Sverige* [Mutual intelligibility of languages in Scandinavia. A study of the mutual understanding of written and spoken language in Denmark, Norway and Sweden]. Stockholm: Nordiska rådet.
- Pierce, J.E. 1952. “Dialect distance testing in Algonquian.” *International Journal of American Linguistics* 18: 203-210.
- Trudgill, P. & Giles, H. 1978. “Sociolinguistic and linguistic value judgments: correctness, adequacy, and aesthetics.” In *Functional studies in language and literature*, F. Coppieters and D.L. Goyvaerts (eds), 167-190. Gent: Story-Scientia.
- Van Bezooijen, R. & Van den Berg, R. 1999a. “Word intelligibility of language varieties in the Netherlands and Flanders under minimal conditions.” In *Linguistics in the Netherlands 1999*, R. van Bezooijen and R. Kager (eds), 1-12. Amsterdam/Philadelphia: Benjamins.
- Van Bezooijen, R. & Van den Berg, R. 1999b. “Hoe verstaanbaar is het Fries voor niet-Friestaligen? [How intelligible is Frisian for non-Frisian speakers?]” *Philologia Frisica anno 1999*: 9-26. Ljouwert: Fryske Akademy.
- Van Bezooijen, R. & Gooskens, C. 2005. “Intertalig tekstbegrip. De begrijpelijkheid van Afrikaanse en Friese teksten voor Nederlandse lezers [Interlingual text comprehension. The intelligibility of Afrikaans and Frisian texts for Dutch readers].” *Nederlandse Taalkunde* 10: 129-152.
- Webb, V. 2002. *Language in South Africa. The role of language in national transformation, reconstruction and development*. Amsterdam: John Benjamins.
- Wolff, H. 1959. “Intelligibility and inter-ethnic attitudes.” *Anthropological Linguistics* 1: 34-41.

Zeevaert, L. 2004. *Interskandinavische Kommunikation. Strategien zur Etablierung von Verständigung zwischen Skandinaviern im Diskurs*. Hamburg: Dr. Kovač.