INTELLIGIBILITY OF HIGH AND LOW GERMAN TO SPEAKERS OF DUTCH

Charlotte Gooskens[1], Sebastian Kürschner[2], Renée van Bezooijen[1]
[1]University of Groningen, The Netherlands [2]University of Erlangen-Nürnberg, Germany
c.s.gooskens@rug.nl, Sebastian.Kuerschner@ger.phil.uni-erlangen.de, r.vanbezooijen@let.ru.nl

Abstract

This paper reports on the intelligibility of Low German and High German for speakers of Dutch. Two aspects are considered. First, we assessed the relative intelligibility of Low and High German for Dutchmen in general. Second, we concentrated on Low German and compared the understanding of listeners from the Dutch-German border area with listeners from other areas in the Netherlands. The results show that Dutchmen understand more High German than Low German and that subjects from the border area understand more Low German than subjects from other parts of the country. A larger amount of previous experience with High German than with Low German seems to explain the first result while phonetic proximity plays an important role in the second result.

Key words

Intelligibility, Low German, Dutch, Levenshtein distance, language contact

1. Introduction

Dutch and German originate from the same branch of West Germanic. In the early Middle Ages these neighboring languages constituted a common Low and Middle German dialect continuum. Only when standardization came about in connection with nation building did the two distinct languages evolve into separate social units. High German spread out over the German language area as a standard language. Low German is preserved in Northern German dialects. It is considered to be structurally

closer to Dutch than to High German since the Low Franconian dialects – from which Standard Dutch originates (cf. Willemyns 2003) – just like all the other Low German dialects are not characterized by the High German sound shift.

Based on an experimental study, Ház (2005) reports that mutual intelligibility between Dutch and High German is possible to a certain extent. The intelligibility of High German among Dutch-speaking individuals is higher than that of Dutch among German individuals, which is attributed to the fact that German is an obligatory school subject in the Netherlands, whereas most Germans do not learn any Dutch. Ház reports that the knowledge of Low German enhances the intelligibility of Dutch among German subjects. Still, we do not know how well Low German is understood by Dutch individuals.

In the present investigation we address the following research questions:

1. Which language variety has the highest intelligibility among speakers of Dutch, Low German or High German?
2. Are Dutchmen from the Dutch-German border area better at understanding Low German than Dutchmen from other parts of the Netherlands?

Ad 1. Two plausible hypotheses predict different results: a) Many Dutchmen learn High German at school and they may regularly be confronted with this variety in the media. For this reason they can be expected to understand High German better than Low German. b) Dutch is structurally closer to Low German and this may result in a better understanding of Low German than of High German.

Ad 2. Although most speakers of Dutch speak (a regionally or socially marked variety of) the standard language, the local dialects spoken in the Dutch / German border area still have a strong position. The border dialects are typologically closer to the dialects in the neighboring parts of Germany than the Dutch standard language, and accordingly linguistic distances are smaller (cf. e.g. Giesbers 2008:165f.). Furthermore, people from the border region are more likely to meet speakers of Low German and hear this variety than people from other parts of the country. We

therefore expect people from the border regions to be better able to understand Low German than people from the rest of the Netherlands.

In order to test these hypotheses, we conducted intelligibility tests and compared the obtained scores with results of different linguistic distance measures. We tested the intelligibility of Low and High German isolated words among Dutch high school pupils. Intelligibility was assessed by means of an Internet-test. The subjects translated 384 High German or 369 Low German nouns into Dutch. In order to find out to what extent intelligibility is related to linguistic distance, we correlated the intelligibility scores with lexical and phonetic measures of linguistic distance. Lexical distance was measured by counting the number of cognates. Phonetic distances were obtained using the Levenshtein algorithm (Heeringa 2004). We also looked at the results for cognates (historically related words) and non-cognates (historically non-related words) separately. Since non-cognates should be unintelligible, linguistic distance can play no role in the intelligibility of this subgroup. Differences in intelligibility of non-cognates between Low German and High German (research question 1) or in border listeners and non-border listeners (research question 2) must therefore be attributed to language contact or instruction.

In Section 2 we describe the experiment that was set up to give an answer to the first research question. The intelligibility of Low and High German was tested among High school pupils from all over the Netherlands, including the border area in the North of the Netherlands close to Germany. Section 3 is concerned with the experiment which was set up to answer question 2. In this experiment the intelligibility of Low German was tested in two listeners groups, one group of high school pupils from the border area and one group from other parts of the Netherlands. In Section 4 we draw some general conclusions.

2. Intelligibility of High and Low German

2.1. Method

2.1.1. Intelligibility experiment

To test word intelligibility, an Internet-based experiment was conducted.[1] In this experiment, Dutch subjects were confronted with 384 High German or 369 Low German isolated singular nouns. Data for such a large number of words allow for a detailed analysis of the kind of problems that listeners meet when listening to words in a closely related language. The nouns were randomly selected from a list of 2575 highly frequent words. In a pre-test, we assured that all the nouns were known to subjects from the test group, i.e. high school pupils aged 15-18 years. The test words were read aloud by a male native speaker of High German (from Wernigerode in the district of Harz, Saxony-Anhalt) and a male native speaker of Low German (from Bremen in northwestern Germany) and recorded in professional sound studios.[2]

144 Dutch high school pupils participated in the experiment, 20 of whom listened to High German and 124 listened to Low German. The subjects were asked a number of questions about their background before they participated in the experiment. Part of the information is summarized in Table 1. We see that the mean age of the two groups of subjects is almost the same (16.5 versus 16.3). Both groups consist of approximately the same number of boys and girls. The subjects came from different places all over the Netherlands, including the border area.[3] .

	Test language of subjects	
	High German	Low German
Number	20	124
Gender	F: 10, M: 10	F: 64, M: 60
Age	15-18 (mean 16.5)	15-18 (mean 16.3)
Years German	3.4	3.7

Tab 1: Number of subjects, gender, mean age, mean number of years of German lessons at school per group of listeners.

[1] The experiment also included other Germanic language pairs. It may be found on the Internet at http://www.let.rug.nl/lrs. It is possible to participate in the test with a guest account (login: germanic, password: guest). We thank Johan van der Geest for programming the experimental interface and databases.

[2] We thank Jörg Tiedemann and Reinhard H. Golz for translating the test words into High and Low German respectively and for recording the test words for the experiment.

[3] There were subjects from the following provinces: Drenthe, Groningen, Zeeland, Flevoland, and Overijssel.

The experiment would have been too arduous if all subjects had been given all test words. Therefore, each subject heard only one word block consisting of about one quarter of the 384 High German or 369 Low German words. The choice of the words and the order of presentation were randomized so that possible tiredness effects were neutralized. Since the word blocks were automatically assigned to the subjects in random order, some word blocks were presented to more subjects than others. The lowest number of subjects who heard a word block was 3 and the highest number 38. The subjects listened to the test words via head phones and were requested to write the Dutch translation into a text field within ten seconds. Prizes were offered to the highest-scoring participants in an attempt to encourage them to complete the tasks to the best of their ability.

The results were automatically categorized as right or wrong through a pattern match with expected answers. The answers that were categorized as wrong were subsequently checked manually. Responses that deviated from the expected responses due to a mere spelling error were counted as correct identifications. Spelling errors were objectively defined as instances where only one letter had been spelt wrongly without resulting in another existing word. So, for example the mistake in *kultuur* (correct *cultuur*) 'culture' is considered a spelling mistake and therefore counted as correct (only one wrong letter without resulting in another existing word), while *ook* (correct *oog* 'eye') was not counted as correct because the spelling mistake results in an existing word meaning 'also'. Some Low German words have more than one possible translation. For example the Low German word *Laden* was sometimes translated into Dutch *winkel* and sometimes into *boetiek* both meaning 'shop'. Both translations were counted as correct. In the case of homonyms, both possible translations were accepted as correct. For example, Low German *översetten* can be translated correctly into Dutch *vertalen* 'translate' or *vertaling* 'translation'.

After this procedure, we had obtained a score of zero (word not identified) or one (word identified) per word for each subject. We then calculated the percentage of correct translations per word. This percentage was the intelligibility score per word. We also calculated the percentage of correct translations per subject, obtaining the intelligibility score per subject.

2.1.2. Linguistic distance measurements

In order to be able to relate the intelligibility results to linguistic distances, two kinds of distances were calculated, namely lexical and phonetic distances. The methods and the results of the measurements are presented in this section.

Lexical distances

We express the lexical distances as the percentage of non-cognates. The larger the number of non-cognate relationships between two languages, the larger the lexical distance. In our material, most of the test words (295 = 76.8% for High German and 302 = 81.8% for Low German) are cognates, i.e. they are historically related to their Dutch equivalents. Examples of such cognates are High German *Weg* – Dutch *weg* 'road' and Low German *Stimm* – Dutch *stem* 'voice'. There are 89 (23.2%) High German non-cognates, for example High German *Ausbildung* – Dutch *opleiding* 'education', and 67 (18.2%) Low German non-cognates, for example Low German *Dern* Dutch – *meisje* 'girl'. This means that the Dutch subjects listening to Low German may have a small lexical advantage compared to the subjects listening to High German, since they are less often confronted with a non-cognate.

Phonetic distances

The phonetic distances were measured by means of Levenshtein distances. The Levenshtein algorithm is a measure of string edit distance based on the smallest number of operations necessary to map a given string on another string. Applied in linguistics, a string of sounds from one variety can be mapped on the corresponding string in another variety (cf. Heeringa, 2004). Insertions, deletions, and substitutions are possible operations. The example in Table 2 shows the calculation of the string edit distance between Dutch *maand* and High German *Monat* 'month', pronounced as [maːnt] in Dutch and as [moːnat] in High German.

.

Alignment	1	2	3	4	5
High German	m	oː	n	a	t
Dutch	m	aː	n		t
Operations		substitution		deletion	
Cost	0	1	0	1	0

Tab 2: Calculation of Levenshtein distance.

First, the two strings are aligned, with identical sounds being matched with each other (e.g. [m] and [m]). Subsequently, the number of operations necessary to transform the one string into the other is calculated. Each operation is assigned a cost of one point. In our example three sounds are identical and therefore they do not add any costs. In contrast, operations are necessary for the first vowel, which has to be substituted, and for the second vowel, which has to be deleted in order to change the High German pronunciation into the Dutch pronunciation. Since operations have to be performed at two slots, the Levenshtein distance is 2. To relate the distance to word length, we divide by the number of alignments, i.e. 5 in the example. The normalised distance is 2/5 = 0.4, i.e. 40 per cent for our example.

The test words and their Dutch equivalents were transcribed by an experienced phonetician who is a mother tongue speaker of Dutch. Levenshtein distance was calculated automatically for all pairs of cognates in both test languages as a fraction representing a percentage. 28 High German and 34 Low German words had a distance of zero to Dutch, for example Dutch *blad* - Low German *Blatt* 'leaf', which are both pronounced as [blat]. Four High German and four Low German cognate words had the maximum distance of 100 per cent to Dutch, for example Dutch *oog* [o:x] - High German *Auge* [awgə] 'eye'. In Levenshtein terms these cognate word pairs behave like non-cognate pairs.

The results of the measurements reveal that Low German cognates are slightly more similar to Dutch than High German cognates (40.6% versus 43.3%). We tested whether these results differ significantly. Since the results were not normally distributed, we applied a Mann-Whitney Test. The results show that the difference is not significant ($p > .05$, $df = 584$, Mann-Whitney Test).

2.2. Results

In Table 3 the results of the word intelligibility tests are presented. We will first discuss the overall results and next we will focus on the intelligibility of the cognates and the non-cognates since this can give an impression of the role that linguistic distance and experience with the language play in the intelligibility.

Words	N words High German	Intelligibility High German	N words Low German	Intelligibility Low German	Sign. ($df = 142$)
Cognates	295	71.4	302	65.6	$p < .01$
Non-cogn.	89	26.6	67	10.3	$p < .001$
Total	384	60.2	369	55.7	$p < .05$

Tab 3: The results of the High German and Low German intelligibility tests, for cognates, non-cognates and totals as well as the results of a Mann-Whitney-Test.

2.2.1. Overall intelligibility

The results in Table 3 show that overall the subjects listening to the High German words were able to translate more words correctly (60.2%) than the subjects listening to Low German (55.7%). The results of a Mann-Whitney Test show that the means differ significantly ($p<.05$).

We correlated the binary variable coding the cognate/non-cognate distinction with the intelligibility scores per word in both groups. The correlation between the Dutch-Low German lexical distance and the intelligibility of Low German is significant ($r = .55$, $p < .001$) and so is the correlation between the Dutch-High German lexical distance and the intelligibility of High German ($r = .49$, $p < .001$). The correlations between the phonetic distances and the intelligibility results are also significant for both test languages ($r = -.52$, $p < .001$ for High German and $r = -.61$, $p < .001$ for Low German). It is therefore reasonable to expect lexical and phonetic distances to play a role in the intelligibility. Since the linguistic distances between Low German and Dutch are smallest, the Dutch subjects listening to Low German may have some linguistic advantage. But even though the linguistic distances to Low German are smaller, the subjects still understand High German better. This points to prior experience with the test language as a more important factor than linguistic distances.

2.2.2. Cognates

When looking at the cognates separately, we see the same trend as for the overall results: the Dutch subjects translate more High German than Low German words correctly (71.4% versus 65.6%). This difference is significant at the one percent level. Again this points to prior experience with the test language as a more important factor than linguistic distances. However, this does not mean that linguistic distances do not play a role. The difference may have been even larger if the subjects listening to Low German had not had a phonetic advantage.

In Section 2.1.2 we already saw that there is a significant inverse correlation between phonetic distance and intelligibility. In order to confirm that smaller phonetic distances do indeed lead to a higher percentage of correctly translated words, we split up the results into three groups applying the following procedure. We first subtracted the phonetic distances to Low German from the distances to High German per word. 92 words turned out to have a smaller distance to Low German than to High German (group 1, see examples in Table 4). For 75 words it was the other way round, in this group the distance to Low German was larger than to High German (group 2, see examples in Table 5). Finally, 101 words showed no difference in distance (group 3, see Table 6).

Dutch	Low German	High German	meaning
[stɔrm]	[stœɐm]	[ʃtʊɐm]	‘storm’
[as]	[aʃ]	[aʃə]	‘ashes’
[vlax]	[flax]	[flagə]	‘flag’

Tab 4: Examples of Dutch words that have a smaller phonetic distance to Low German than to High German (group 1).

Dutch	Low German	High German	meaning
[haːvən]	[hɔːbm]	[haːfn]	‘harbor’
[blut]	[blɔwt]	[bluːt]	‘blood’
[zaːl]	[zɔːl]	[zaːl]	‘saloon‘

Tab 5: Examples of Dutch words that have a smaller phonetic distance to High German than to Low German (group 2).

Dutch	Low German	High German	meaning
[blɔk]	[blɔk]	[blɔk]	‘block’
[kʏnst]	[kʊnst]	[kʊnst]	‘art’
[pla:ts]	[plats]	[plats]	‘place’

Tab 6: Examples of Dutch words that have the same phonetic distance to High German and to Low German (group 3).

If phonetic distance does indeed play a role, the subjects can be expected to have an advantage when listening to the Low German words with a smaller distance to Dutch than the High German words and they also can be expected to have phonetic advantage when listening to the High German words that are more similar to Dutch than the Low German words. However, it is an additional advantage for this latter selection of words that the subjects have experience with High German. For the words with the same phonetic distance to Low and High German there is no phonetic advantage for either of the two languages and only an advantage of experience for the High German words. The advantages for the three selections of words are summarized in Table 7 together with the intelligibility results.

	Group 1:		Group 2:		Group 3:	
Phonetic distance	Low German < High German		High German < Low German		High German = Low German	
Test language	Low German	High German	Low German	High German	Low German	High German
Phonetic similarity	+	-	-	+	-	-
Experience	-	+	-	+	-	+
% correct	71.7	68.7	58.4	75.6 **	71.0	71.7

Tab 7: phonetic and experience advantages and intelligibility results for three subgroups of words, 1 (Dutch words with a smaller phonetic distance to Low German than to High German), 2 (Dutch words with a smaller phonetic distance to High German than to Low German), and 3 (Dutch words with the same distance to High German and Low German). ** indicates that the result for the test language is significantly higher than for the other test language at the .01 level.

Table 7 shows that the intelligibility of the Low German words that have a smaller distance to Dutch than their High German equivalents (group 1) are not significantly better understood (Low German 71.7% correct translations versus 68.7% for High German, $p > .05$, Wilcoxon Signed Ranks Test). This result shows that the advantage of the greater phonetic similarity of Low German is counterbalanced by the advantage of the greater experience with High German.

Looking at the words that have a smaller phonetic distance to High German than to Low German (group 2), we see a large difference between Low German (58.4% correct) and High German (75.6% correct). The difference is significant at the .01 level (Wilcoxon Signed Ranks Test). This large difference can probably be explained by the fact that both experience and linguistic distance favor the intelligibility of High German as compared to Low German. It is furthermore noticeable that there is a large difference between the intelligibility scores of Low German in group 1 and 2 (71.7% vs. 58.4%) while this difference is small for High German (68.7% vs. 75.6%). This points to an important role of phonetic similarity for the intelligibility of Low German and a smaller role for the intelligibility of High German.

Finally, when there is no differences in the phonetic distance to the two test languages, there is also no significant difference in intelligibility (71.0% for Low German and 71.7% for High German, $p >.05$, Wilcoxon Signed Ranks Test). This is remarkable. Because of prior experience one would expect High German to be easier to understand. .

2.2.3. Non-cognates

Finally we look at the intelligibility of the non-cognates. We see a large difference in intelligibility: 26.6% for High German and 10.3% for Low German. This difference is significant at the .001 level (Mann Whitney-Test). As for both language the phonetic distance is maximal for this subset of words, phonetic differences cannot play a role in explaining the difference. So on the basis of the intelligibility of non-cognates we conclude that experience does play an important role in explaining the difference in intelligibility of High and Low German. Non-cognates can only be understood through previous exposure.

3. Intelligibility of Low German

The previous section showed that speakers of Dutch understand more High German than Low German words. This must be due to previous exposure to High German via lessons at school and via the media. In the Netherlands television programs in which a foreign language is spoken are never dubbed and always subtitled. This means that Dutch people are often auditorily confronted with foreign speech. They are often confronted with Standard (High) German, but Low German programs are rare in the German television. We hypothesized that a smaller linguistic distance to Low German may give Low German a head start, but previous exposure to High German seems to counterbalance this effect.

In this Section we will have a closer look at the intelligibility of Low German in the Netherlands. We compare the intelligibility scores of Dutch subjects from the north-eastern regions close to the border with Germany to the performance of subjects from non-bordering regions. Traditionally, the Low-Saxon dialects close to the border have been linguistically transitional between Low German and the dialects of Dutch that are spoken more to the west (cf. Hinskens 1993). Although nation building and standardization has had a large impact on these varieties (cf. de Vriend et al. in press, Giesbers 2008, Kremer 1990, 1996, Niebaum 1990), they are still closer to Low German than the non-Saxon varieties of the Netherlands to which Standard Dutch belongs (cf. Reker 1996). We therefore expected subjects from the border area in the North-Eastern part of the Netherlands to be better at understanding Low German words than subjects from non-bordering regions.

3.1. Method

3.1.1. Intelligibility experiment

The Low German test was identical to the test that was described in Section 2.1.1. i.e. 369 highly frequent Low German nouns were tested among Dutch high school pupils via an Internet experiment. In total 124 subjects were tested. Since we wanted to compare the intelligibility among subjects living close to the German border to the

intelligibility of subjects living in other parts of the countries, the subjects were divided into two groups, referred to as the border group and the non-border group. To determine if the subjects living close to the border could have an advantage from their knowledge of the North-Eastern Dutch dialects, we asked the subjects if they knew the local dialect. Only 65 of the 97 subjects from the border region answered positively to this question. We excluded the remaining 32 subjects from further analysis. There were now 65 subjects in the border group and 27 in the non-border group, adding up to 92 subjects. The 65 subjects from the border group all came from the provinces of Groningen and Drenthe in the Northern part of the Low Saxon area. The dialects spoken in the places where the speakers come from are part of a coherent dialect group. They all belong to the dialect group that Hoppenbrouwers & Hoppenbrouwers (2001, p. 65) refer to as the 'Low Saxon central group'. The 27 subjects in the non-border group came from the provinces of Zeeland (Hulst) and Flevoland (Almere) that are not adjacent to the German border. Information about the subjects is summarized in Table 8. We see that the border subjects are slightly older than the non-border subjects and that they have had German at school for a longer period. In the border group more male than female subjects participated (37 males versus 28 females) and in the non-border group it was the other way round (17 females and 10 males).

	Subjects	
	Border	Non-border
Number	65	27
Gender	F 28, M 37	F17, M10
Age	15-19 (mean 16.5)	15-19 (mean 16.1)
Years German	4.1	3.3

Tab 8: Number of subjects, gender, mean age, number of years of German lessons at school per group of listeners.

3.1.2. Linguistic distance measurements

In Section 2.1.2, we explained how we measured the linguistic distances between Standard Dutch and Low German. In order to be able to compare the linguistic basis that the border group had for understanding the Low German words to that of the non-border group we also needed to measure the distances between the Low Saxon dialect spoken in the area where the subjects from the Dutch border group live and Low German. We therefore had a speaker of the Dutch Low Saxon variety translate the Low German words into his variety and read them aloud. The border group of subjects came from different places in the border area. In order to achieve transcriptions that would represent the whole area we made recordings of a speaker who is born and raised in Uithuizen and now lives in Bedum. These places are in the middle of the area where the listeners came from.[4] On the basis of the recordings, phonetic transcriptions were made by a German linguist (the second author).

Lexical distances

The lexical distances to Low German are very similar for Standard Dutch and the border dialect. Of the 369 Low German test words, 302 (81.8%) were Standard Dutch cognates and 300 (81.3%) were border dialect cognates. Examples are Low German [poli'tsaj] - Dutch border dialect [po'litsi] - Standard Dutch [po'litsi] 'police' and Low German [hɔnt] - Dutch border dialect [hɔ:nt] - Standard Dutch [hɔnt] 'dog'. There were 67 (18.2%) Standard Dutch and 69 (18.7%) border dialect non-cognates, for example Low German [deɐn] - Dutch border dialect [ʊɪxjə] - Standard Dutch [mɛjʃə] 'girl'.

Phonetic distances

The results of the phonetic distance measurements show that the Low German cognates are only slightly closer to the border dialect (39.9%) than to Standard Dutch (40.6%). This difference is not significant ($p > .05$, $df = 290$, Wilcoxon Signed Ranks Test).[5] However, it should be kept in mind that the border group is bilingual and is able to use phonetic information from both the border dialect and Standard Dutch. Therefore we calculated a new Levenshtein distance selecting the smallest distance to

[4] We thank Siemon Reker for translating and recording the test words.

[5] A few cognates had to be left out of the analysis because the transcriptions were missing in the database. Therefore the df is smaller than the number of cognates –1.

Low German per word pair. 49 of the border dialect words had a smaller distance to Low German than their Standard Dutch equivalent and for 38 words the situation was reverse, the Standard Dutch words having a smaller distance to Low German than the border dialect equivalent words. For 130 words the distance was the same for both varieties. The new phonetic distances calculated on the basis of the smallest distance to the border dialect per word pair was 35.5 and this is significantly smaller. So it is likely that the border group subjects have a phonetic advantage above the Standard Dutch group when confronted with Low German.

3.2. Results

In Table 9 the results of the word intelligibility tests are shown. Like in the previous section we will first discuss the overall results and next we will focus on the intelligibility of the cognates and the non-cognates separately.

The correlation between the lexical distance between Low German and the Dutch border dialect and the intelligibility scores of the border group was higher (r = .54, $p < .001$) than the correlation between the lexical distance between Low German and Standard Dutch and the intelligibility of the non-border group (r = .50). However, the correlation is even higher when the intelligibility scores of the border group are correlated with the lexical distance between Low German and Standard Dutch (r = .56, $p < .001$). This might suggest that the border group relates the Low German words to Standard Dutch words in the first place.

The correlation between the linguistic distance from the own variety (border dialect or Standard Dutch) to Low German and intelligibility is significant at the .001 level for both groups (r = –.47 for the border group and r = –.59 for the non-border group). For the border group the correlation becomes larger when intelligibility is correlated with the distance between Low German and Standard Dutch rather than the distance between Low German and their own dialect (r = –.60). Also when correlating with the new distances calculated on the basis of the smallest distance to the border dialect (see above), the correlation is higher (r = –.52). This again suggests that the border group does indeed use Standard Dutch when understanding Low German, and even that the route via Standard Dutch is preferred.

Words	N words	Intelligibility border group	Intelligibility non-border group	Sign. ($df = 90$)
Cognates	302	67.0	61.9	$p < .01$
Non-cogn.	67	11.6	8.1	$p = .07$
Total	369	57.9	51.9	$p < .01$

Tab 9: The results of the Low German intelligibility test for the border group and the non-border group, for cognates, non-cognates and totals as well as the results of a Mann-Whitney-Test.

3.2.1. Overall intelligibility

The border group understands Low German significantly better (57.9% correct answers) than the non-border group (51.9% correct). This result can be due to a linguistic advantage or more experience in the group of the border subjects. In the following two sections we will try to get an impression of the role that these two factors play in the intelligibility of Low German in the two groups of subjects.

3.2.2. Cognates

When only looking at the intelligibility of the cognates we also find that the border group translates significantly more words correctly (67.0%) than the non-border group (61.9%). As explained in Section 3.1.2, the fact that no significant differences are found between the linguistic distances to Low German does not mean that linguistic distances are not part of the explanation for the difference in intelligibility. It is possible that the border group has a linguistic advantage because they can get help from both the border dialect and Standard Dutch. On the other hand the border group is also likely to have had more contact with Low German than the non-border group. We will now have a look at three subgroups of the cognates, namely 49 word pairs where the phonetic distances between the border dialect and Low German are smaller than between Standard Dutch and Low German (group 1, see examples in Table 10), 38 word pairs where the phonetic distances between the Standard Dutch and Low German are smaller than between the border dialect and Low German (group 2, see

Table 11), and 130 word pairs where the distances are the same (group 3, see Table 12).

Low German	Dutch border dialect	Standard Dutch	meaning
[knɛj]	[knɛj]	[kni]	‘knee’
[buːtn̩lant]	[butənlant]	[bœjtn̩lant]	‘abroad’
[ʋiːn]	[ʋiːn]	[ʋɛjn]	‘wine’
[mɔnt]	[mɔːnt]	[maːnt]	‘month’

Tab 10: Examples of Low German words that have a smaller phonetic distance to the border dialect than to Standard Dutch

Low German	Dutch border dialect	Standard Dutch	meaning
[bɛt]	[bɛːr]	[bɛt]	‘bed’
[bibliote:k]	[bibəltajk]	[bibliote:k]	‘library’

Tab 11: Examples of Low German words that have a smaller phonetic distance to Standard Dutch than to the border dialect

Low German	Dutch border dialect	Standard Dutch	meaning
[bɪlt]	[beːlt]	[beːlt]	‘picture’
[modɛl]	[modɛl]	[modɛl]	‘model’

Tab 12: Examples of Low German words that have the same phonetic distance to Standard Dutch and to the border dialect

In Table 13 we summarize the advantages of phonetic proximity and experience for the three subgroups of cognates as well as the intelligibility results for the two groups of subjects. In the group 1 words, the border group subjects have a phonetic advantage as well as an advantage from the fact that they probably have had more contact with Low German and we therefore expect a higher percentage correct answers for the border group. This is indeed the case. The border group translated 48.4% of the words correctly and the non-border group only 36.8%. The difference is significant at the .01 level.

In the group 2, both groups are likely to have the same phonetic advantage, since they both know Standard Dutch. The border group has an advantage from experience. Since the difference in the percentages of correct translations in the two groups of subjects is not significant, it can be concluded that contact only seems to play a minor role in the intelligibility of the border group. We furthermore notice that both subject groups perform better when Standard Dutch words are more similar to Low German (group 2 words) than when the border dialect words are more similar (group 1 words), $p < .001$, $df = 85$ in both cases (Mann-Whitney test). This points to both groups having a larger advantage of phonetic proximity to Standard Dutch than to the border dialect. Also for the border group, the route via Standard Dutch seems to be the most favored one.

In the group 3 words neither of the two groups of subjects has a phonetic advantage but the border group has an advantage of experience. If the border group translates more words correctly it can therefore be concluded that contact plays a role. This is indeed the case.

Based on these results we conclude that contact plays a small role while phonetic advantage is probably more important for the intelligibility. Unfortunately a subgroup of words where contact plays no role does not exist. This makes it difficult to draw stronger conclusions about the role of phonetic similarity compared to experience.

Phonetic distance	Group 1: border < non-border		Group 2: Non-border < border		Group 3: border = non-border	
Test group	border	non-border	border	non-border	border	non-border
Phonetic similarity	+	-	-	-	-	-
Experience	+	-	+	-	+	-
% correct	48.4 **	36.8	74.5	76.8	71.4 *	67.1

Tab 13: Phonetic and experience advantages and intelligibility results for the three subgroups of words, 1 (Low German words with a smaller phonetic distance to border dialect than to Standard Dutch, 2 (Low German words with a smaller phonetic distance to Standard Dutch than to border dialect), and 3 (Low German words with the same distance to border dialect and Standard Dutch). ** indicates that the result

for the test group is significantly higher than for the other test group at the .01 level, and * that it is significant at the .05 level.

3.2.3. Non-cognates

When it comes to understanding non-cognates, phonetic distances can play no role. Non-cognates are in principle unintelligible unless the subjects know them from previous experience with the language. If the border group subjects are able to translate more non-cognates correctly than the non-border group subjects, it can be concluded that subjects living at the Dutch side of the border indeed come into contact with the German dialects spoken at the other side of the border. From Table 9 it becomes clear that the border group translated slightly more non-cognates correctly (11.6%) than the non-border group (8.1%). However, this difference is not significant at the .05 level. On the basis of this result we can therefore again conclude that previous experience does not play a role.

4. Conclusions

The Dutch language is linguistically related more closely to Low German than to High German. The fact that the results of the first experiment presented in this paper show that speakers of Dutch understand more High German than Low German words must therefore be attributed to the more intensive contact with High German via lessons at school and via the media.

Dutchmen from the Dutch-German border area understand more Low German words than Dutchmen from other part of the Netherlands. In this case contact seems to play a less important role and the difference in intelligibility should probably be attributed to the phonetic distances to Low German that are smaller for the border dialect than for Standard Dutch. However, the results also suggest that even though listeners from the border area have a phonetic advantage from their local dialect when listening to Low German, Standard Dutch plays a dominant role for this group as well in the decoding of Low German words.

References

DE VRIEND, Folkert, Charlotte GIESBERS, Roeland VAN HOUT & Louis TEN BOSCH (in press) "The Dutch-German border. Relating linguistic, geographic and social distances", in John NERBONNE, Charlotte GOOSKENS, Sebastian KÜRSCHNER & Renée VAN BEZOOIJEN (eds.): *Computing and language variation.* Special issue of *International Journal of Arts and Humanities Computing* 2/1-2, 119-134.

GIESBERS, Charlotte (2008) *Dialecten op de grens van twee talen: een dialectologisch en sociolinguïstisch onderzoek in het Kleverlands dialectgebied.* Groesbeek: Uitgeverij Reijngoudt-Giesbers, C.

GOLTZ, Reinhard H. & Alastair G. H. WALKER (1990) "North Saxon", in Charles V. J. RUSS (ed.), *The dialects of Modern German.* London: Routledge, 31-58.

HÁZ, Eva (2005) *Deutsche und Niederländer. Untersuchungen zur Möglichkeit einer unmittelbaren Verständigung.* Hamburg: Dr. Kovač. (Philologia 68).

HEERINGA, Wilbert (2004) *Measuring dialect pronunciation differences using Levenshtein distances.* Groningen: Grodil.

HINSKENS, Frans (1993) "Dialect als lingua franca? Dialectgebruik in het algemeen en bij grensoverschrijdend contact in het Nederrijnland en Twente", in Ludger KREMER (ed.), *Diglossiestudien. Dialekt und Standardsprache im niederländisch-deutschen Grenzland.* Vreden: Landeskundliches Institut Westmünsterland, 209-245.

HOPPENBROUWERS, Cor & Geer HOPPENBROUWERS (2001) *De indeling van de Nederlandse streektalen. Dialecten van 156 steden en dorpen geklasseerd volgens de FFM*. Assen: Van Gorcum.

KREMER, Ludger (1990) "Kontinuum oder Bruchstelle? Zur Entwicklung der Grenzdialekte zwischen Niederrhein und Vechtegebiet", in Ludger KREMER & Hermann NIEBAUM (eds.), *Grenzdialekte. Studien zur Entwicklung kontinentalwestgermanischer Dialektkontinua.* Hildesheim etc.: Olms (Germanistische Linguistik 101-103), 85-123.

KREMER, Ludger (1996) "Standardisierungstendenzen und die Entstehung sprachlicher Bruchstellen am Beispiel der niederländisch-deutschen Kontaktzone", in *Niederdeutsches Wort* 36, 59-74.

NIEBAUM, Hermann (1990) "Staatsgrenze als Bruchstelle? Die Grenzdialekte zwischen Dollart und Vechtegebiet", in Ludger KREMER & Hermann NIEBAUM (eds.), *Grenzdialekte. Studien zur Entwicklung kontinentalwestgermanischer Dialektkontinua.* Hildesheim etc.: Olms (Germanistische Linguistik 101-103), p. 49-83.

REKER, Siemon (1996) "Das Groningische auf dem Wege vom niederdeutschen Dialekt zur Lautvariante des Niederländischen", in *Niederdeutsches Wort* 36, 19-29.

WILLEMYNS, Roland (2003) "Dutch", in Ana DEUMERT & Wim VANDENBUSSCHE (eds.), *Germanic standardizations. Past to present.* Amsterdam/Philadelphia: John Benjamins, 93-125.