

Gooskens, Charlotte & Femke Swarte. 2017. Linguistic and extra-linguistic predictors of mutual intelligibility between Germanic languages. *Nordic Journal of Linguistics* 40(2), 123–147.

Linguistic and extra-linguistic predictors of mutual intelligibility between Germanic languages

Charlotte Gooskens & Femke Swarte

We report on a large-scale investigation of the mutual intelligibility between five Germanic languages: Danish, Dutch, English, German and Swedish. We tested twenty language combinations using the same uniform methodology, making the results commensurable for the first time. We first tested both written and spoken language by means of cloze tests. Next we calculated linguistic distance at the levels of lexicon, orthography, phonology, morphology and syntax. We also quantified exposure and attitudes towards the test languages. Finally, we carried out a regression analysis to determine the relative importance of these linguistic and extra-linguistic predictors for the mutual intelligibility between Germanic languages. The extra-linguistic predictor exposure was the most significant factor in predicting intelligibility in the Germanic language area. The effect of attitude was very small. Lexical distance, orthographic and phonetic distances were the most important linguistic predictors of intelligibility.

Keywords: language attitude, language exposure, linguistic distances, mutual intelligibility, receptive multilingualism

Charlotte Gooskens, Center for Language and Cognition, University of Groningen, P.O. Box. 716, 9722 AS Groningen, The Netherlands. c.s.gooskens@rug.nl

Femke Swarte, Center for Language and Cognition, University of Groningen, P.O. Box. 716, 9722 AS Groningen, The Netherlands. femkeswarte@hotmail.com

1. INTRODUCTION

Languages within a language family show various degrees of similarity. Some languages are so linguistically similar that they are mutually intelligible even if the speakers have not learned each other's languages. This paper deals with the mutual intelligibility of five Germanic languages: Danish, Dutch, English, German and Swedish, and the variables that predict the mutual intelligibility among these languages. Previous research has shown that the Scandinavian languages are mutually intelligible to a considerable extent (see Haugen 1966, Maurud 1976, Bø 1978, Börestam Uhlmann 1991, Jörgensen & Kärrlander 2001, Lundin & Zola Christensen 2001, Delsing & Lundin Åkesson 2005). They are so similar that with some effort speakers of Danish, Swedish and Norwegian can communicate sufficiently well to

sustain a meaningful exchange of information each using their own native languages. We use the term RECEPTIVE MULTILINGUALISM (Braunmüller 2007) for this manner of communication. An asymmetry has been found in spoken intelligibility between Danish and Swedish. It is easier for Danes to understand spoken Swedish than vice versa. These results are usually explained by extra-linguistic factors such as asymmetric attitudes towards the (speakers of the) languages involved and unequal experience with the languages. Danes generally have a more positive attitude towards Swedes and often have more exposure to Swedish through the media and on vacation than the other way around (Delsing & Lundin Åkesson 2005). In addition to these non-linguistic explanations of asymmetry, strong evidence has been found for linguistic explanations of the asymmetric Danish–Swedish intelligibility (Schüppert 2011).

The mutual intelligibility between Dutch and German, too, is asymmetric. This is often attributed to the fact that speakers of Dutch learn German in school (see Ház 2005). However, Gooskens, Van Bezooijen & Van Heuven (2015) discovered that this asymmetry also exists among younger speakers of Dutch who have not yet learned German at school and young speakers of German. Therefore they conclude that the asymmetry between Dutch and German may also have a linguistic basis.

There are also a few studies that investigate the mutual intelligibility across the North and West Germanic languages, like the (mutual) intelligibility of Danish and Swedish on the one hand and Dutch on the other (Hedquist 1985, Hedquist & Strangert 1989), the intelligibility of English and Danish for Norwegians and Swedes (Gooskens 2006b), the intelligibility of Danish for speakers of Dutch and Frisian (Van Bezooijen, Gooskens & Kürschner 2012) and the intelligibility of Swedish for Germans (Vanhove 2014).

The investigations discussed above used various methods to establish the level of mutual intelligibility between the languages making it difficult to compare the results. The aim of this paper is to report on a large-scale investigation on the predictors of mutual intelligibility between five Germanic languages, two North Germanic (Danish and Swedish) and three West Germanic (Dutch, English and German). We tested twenty language combinations using the same uniform methodology, making the results commensurable for the first time. We tested both written and spoken language.¹ The listeners were assigned a test language randomly and some of them had learned the language (or another closely related language) at school or were familiar with it from various kinds of exposure. In that sense, the results are an ecologically valid representation of the mutual intelligibility between five Germanic languages among young educated people. We refer to this kind of intelligibility as ACQUIRED INTELLIGIBILITY. We also had a closer look at the intelligibility results of a subgroup of participants who had not learned the test language and had only minimum of exposure to it. This allowed us to investigate how well participants are able to understand the language on the basis of structural linguistic similarities between their own language and the test language, so-called INHERENT INTELLIGIBILITY.

Intelligibility research of the past decade often focused on uncovering the variables that influence mutual intelligibility. The results are interesting from a theoretical linguistic perspective since it is of importance to gain a better understanding of the mechanisms underlying intelligibility. However, from a more practical point of view, the results may also be of importance. For example, they may aid in resolving issues that concern language planning and policies, second-language learning, and language contact. Predictors of intelligibility are usually divided into two groups: linguistic and extra-linguistic variables. The more linguistic features two languages share, the higher their degree of mutual intelligibility will be. In previous research, relationships between intelligibility and distance at various linguistic levels have been established: lexicon (Van Bezooijen & Gooskens 2005), phonology (Gooskens 2006a, Kürschner, Van Bezooijen & Gooskens 2008, Gooskens et al. 2015) and orthography (Van Bezooijen & Gooskens 2005, Kürschner et al. 2008, Doetjes & Gooskens 2009, Gooskens et al. 2015). The role of morphology and syntax has hardly been investigated (Hilton, Gooskens & Schüppert 2013). In this paper we will for the first time correlate syntactic and morphological distances with intelligibility and include them in a statistical model of mutual intelligibility between Germanic languages together with lexical, phonetic and orthographic distances.

Extra-linguistic variables may also play a role in intelligibility. The amount of exposure to the language of the speaker has been shown to correlate positively with intelligibility (Bø 1978, Jörgensen & Kärrlander 2001, Lundin & Zola Christensen 2001, Delsing & Lundin Åkesson 2005). Previous research (e.g. Hedquist 1985, Golubović 2016) has shown that in the case of closely related languages only very little exposure or a short language course that makes speakers conscious of the most important differences and similarities between their native language and the language of the speaker can improve mutual intelligibility considerably. Through exposure the participant will get used to the sounds of the language and how these sounds correspond to those in his own language. He or she is also likely to learn some of the vocabulary. Of course also people who have learned the language in a formal setting will understand the language better than participants who have not. Furthermore, a number of researchers found high intelligibility scores to correlate with positive attitudes towards the test language (Kuhlemeier, Van den Bergh & Melse 1996, Jörgensen & Kärrlander 2001, Lundin & Zola Christensen 2001, Delsing & Lundin Åkesson 2005, Gooskens 2006a, Impe 2010, Schüppert 2011). Positive attitudes toward a test language, its speakers and the country where it is spoken can be expected to motivate a participant to make a greater effort to understand the language than negative attitudes. However, the link demonstrated in experimental settings has often been rather weak. We will include measurements of exposure, language learning and attitude in our statistical model.

In [Section 2](#) we will present the set-up and results of an online investigation on the mutual intelligibility of five Germanic languages, in [Section 3](#) we will show

how we quantified the linguistic distances and extra-linguistic predictors of mutual intelligibility between these languages, and in [Section 4](#) we will determine how well the degree of mutual intelligibility between the languages can be predicted.

2. MEASURING MUTUAL INTELLIGIBILITY BETWEEN GERMANIC LANGUAGES

2.1 Method

2.1.1 Cloze test

In order to measure intelligibility a cloze test was developed. This kind of test has proven to be a reliable method to measure text comprehension (Bormuth 1969, Abraham & Chapelle 1992, Keshavarz & Salimi 2007) and has been used in previous studies on mutual intelligibility (Van Bezooijen & Gooskens 2005, Gooskens & Van Bezooijen 2006). For the test we selected four English texts at the B1-level of difficulty, as defined by the Common European Framework for languages (Council of Europe 2001)² and adjusted them to a length of approximately 200 words each. Their contents are culturally neutral and concerned with everyday subjects (children and sports, having a cold, driving in winter and riding a bike). The texts were translated into each of the five test languages and recorded by four female native speakers (one per text) who had been selected from a larger group by a panel of native speakers for being representative speakers of the standard languages.

From each text twelve words were deleted at roughly equidistant intervals and replaced by a gap of uniform length. In the written version of the cloze test, the passage was presented on a computer screen in its entirety. A grid with response alternatives was continually shown at the top of the screen, listing four nouns, four adjectives and four verbs in three columns. The task of the participants was to drag and drop each of the twelve response alternatives in the appropriate slot in the text. The task had to be completed within ten minutes. In the oral version of the test the same twelve target words were replaced by a beep (1000 Hz sine wave) of uniform length (1 second). Each sentence was made audible twice in a row, with 1 second between first and second presentation. Each sentence contained just a single beep. Within 30 seconds, the participants had to click one of the twelve response alternatives in the grid presentation on screen such that it best fitted the missing word in the utterance just heard. Once clicked the response alternative was greyed out. The participants were told that each word could be used only once, but if they thought they had made a mistake, they could select a word they already used. If the participants moved their cursor over a word, a translation of the word in their native language was shown. This was done because the purpose of the cloze test was not to investigate whether the participants knew the twelve words in the grid, but whether they understood the text

as a whole. Responses were automatically scored as correct or incorrect. [Figure 1](#) shows a screen shot of the user interface employed in the presentation of the written cloze test.

2.1.2 Participants

In order to keep the participant groups as comparable as possible, we made a careful selection of participants for further analysis. We only selected participants with a university education, who had grown up in the countries included in the project (Denmark, Germany, the Netherlands, Sweden or the United Kingdom), had lived there most of their lives, and who indicated that they only had one native language and that this native language was the only language they spoke at home. After the selection procedure, 954 participants (39% male and 61% female) remained who were suitable for the analysis, 528 of these took part in the written cloze test and 426 who took part in the spoken cloze test. We tested participants with five different native languages (Danish, Dutch, English, German or Swedish). Each participant was confronted with one of five related languages (Danish, Dutch, English, German or Swedish, but not their native language) and was tested with a written or a spoken cloze test. In total, this results in 40 participant groups. In the final selection, each of the 40 participant groups contained between 15 and 41 participants. The average number of participants per group was 23.9. All selected participants were between 18 and 33 years old. The mean age was 23.6 years ($sd = 2.82$).

2.1.3 Procedure

The intelligibility, attitude and exposure data were gathered with an online application (www.micrela.nl/app). Participants were mainly recruited through Facebook, online newspapers and university mailing lists during the spring of 2014. Participants first selected their native language and then filled out a list with background questions concerning age, sex, level of education, the country they grew up in and the country where they spent most of their lives. Also, questions about which languages they spoke at home, how many years they had learned English and how well they thought they could understand each of the other Germanic languages were included in this part of the questionnaire. Next, they were asked to indicate how ugly or beautiful they found each of the five languages on a five-point scale (from 1 'very ugly' to 5 'very beautiful'). Participants gave their ratings once without having heard a spoken sample of the languages. They then listened to each of the five test languages and rated the samples for perceived beauty. For each language, they heard the first paragraph of the declaration of human rights spoken by one of the four speakers (random choice) who had also recorded the stimulus materials for the intelligibility tests.

Participants were then informed which language they would hear (or see) as the test language. The test mode (written or spoken), text (one of four) and the

lange	Schnee	fährst
vereisten	Auto	passieren
nass	Kaffee	vermeiden
geschmolzenem	Mal	rutschen

Der Winter ist gefährlich, da es schwierig einzuschätzen ist, was passiert. Unfälle _____ so schnell. Hinter einer Hügelkuppe kann Nebel warten. Unter _____ Schnee kann sich Eis verstecken, das einen von der Straße abbringt. Das Auto, das dir entgegenkommt, kann auf einmal über die Straße _____ . Regel Nummer Eins für das Fahren auf _____ Straßen ist gleichmäßiges Fahren. Ruckhafte Bewegungen sorgen dafür, dass ein _____ schwierig zu kontrollieren ist. Jedes _____ , wenn man lenkt, bremst oder beschleunigt, muss man dies so sanft und langsam tun wie möglich. Stelle dir vor, dass auf dem Beifahrersitz eine volle heiße Tasse _____ steht. Fahre so, dass du sie nicht umschütten würdest. Regel Nummer Zwei ist es aufmerksam zu sein, was passieren könnte. Je mehr Eis sich auf der Straße befindet, desto weiter muss man voraus blicken. Teste, wie _____ man braucht um anzuhalten, indem du sanft bremst. Denk daran, dass du vielleicht schneller _____ als du denkst. Sorge generell für den doppelten Bremsabstand, wenn die Straße _____ ist. Nimm dreimal mehr Abstand bei _____ , und sogar noch mehr bei einer vereisten Straße. Probiere jederzeit die Kontrolle über dein Auto zu behalten, so wirst du Probleme _____ .

Figure 1. Screen shot of a German written cloze test.

test language were randomly assigned to the participant. They were asked to indicate whether they had learned the test language, and to estimate how much exposure to the language they had had by listening to speakers through the media (television, DVD, movies), playing computer games, talking to people (live, or through chatting/skyping over the internet) and by reading books, newspapers or text on the internet. The intensity of the non-native language exposure was rated on a five-point scale between 1 'never' and 5 'daily'. This extra-linguistic information was collected immediately before the participant did the intelligibility test.

Many of the participants described above had some degree of acquired intelligibility of the test language. We were also interested in inherited intelligibility. This condition could be approximated by making a further selection from our set of participants, such that those who were included had (virtually) no experience with the stimulus language. In the ideal case we would select participants with no prior exposure to the test language whatsoever. However, that would leave too few participants for an analysis. We therefore selected a subset of participants who had indicated that their mean exposure with the test language across the six scales was below 2 on the scale from 1 'never' to 5 'daily' and who had not learned the test language. This not only reduced the number of participants to 292 but it also reduced the number of language combinations represented by at least six participants. For instance, there are no Dutch participants older than 18 years who have not learnt English in school. As a consequence, six out of the original twenty language combinations are no longer represented in the sub-set of inherited intelligibility.

2.2 Results

In the cloze test, participants have to fill words into twelve gaps. These gaps can be correctly or incorrectly filled. In [Figures 2](#) and [3](#) we present the percentages of correctly filled gaps for the spoken and the written cloze test. We present two sets of data. The black bars show the results of all 954 participants. The grey bars show the results of the subgroup of participants who had had a minimum of exposure to the test language (inherited intelligibility). The data are ordered from lowest to highest percentage of correct answers in the spoken cloze test. Each bar represents the intelligibility of one test language in one group of participants (a language combination). The participant language is given first and the test language second. So, for example we see that Dutch participants taking the English test perform best, both in the spoken and the written version of the cloze test.

2.2.1 Overall results

We first present the overall results of the two tests (black bars in [Figures 2](#) and [3](#)). These results include acquired intelligibility since many participants (69%) had learned the test language or had been exposed to it. In all language combinations it

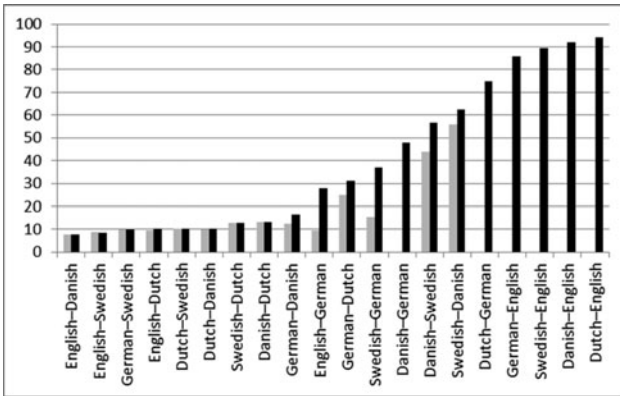


Figure 2. The results of the spoken cloze test. For each language combination, participant language is given first, test language second, so for example ‘English–Danish’ means ‘English native speakers listening to Danish’. The black bars show the mean results of all participants ordered from lowest to highest percentage of correct answers and the grey bars show the mean results for the participants with a minimum of exposure (inherited intelligibility).

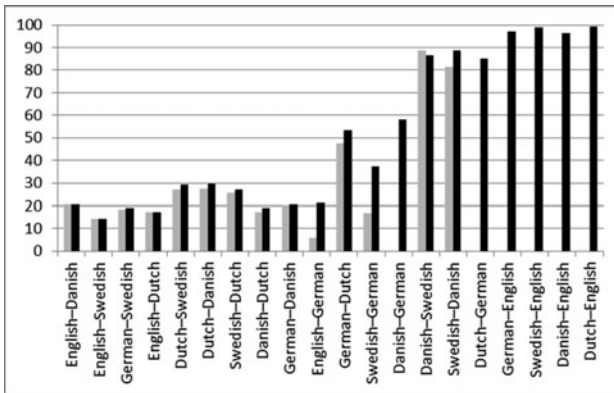


Figure 3. The results of the written cloze test. For explanations see Figure 2 caption. For easy comparison between spoken and written intelligibility the results are presented here in the same order as in Figure 2.

was easier for the participants to understand a closely related language in the written form (mean across all language combinations is 51%) than in the spoken form (40%). The same general trend can be seen for spoken and written intelligibility: language combinations with low intelligibility in writing also show low intelligibility when spoken and language combinations with high intelligibility when tested with the spoken cloze test also get high scores in the written cloze test. The correlation between the two tests is significant at the .01 level ($r = .97$).

There are large differences in intelligibility among the various language combinations. Nine of the participant groups in the spoken test and four groups

in the written test scored below 20% and there are four groups in the spoken and seven groups in the written test that have high mean scores (above 80%). The highest scores all involve English as the test language. These high scores were expected because of the role of English as a lingua franca and the importance of English in the education system. The lowest scores are found in combinations of languages where the participants and the test languages are from different branches of the Germanic language tree (West or North), such as Swedish–Dutch, as well as intelligibility of Dutch and German for English participants. Finally there is a middle group of language combinations with intelligibility scores between 20% and 80%. Some of these combinations involve German as the test language. German is a school language for many children in the other countries, which explains why German is so well understood even among participants from the North Germanic group. Furthermore, we see that German participants understand some Dutch (31% correct responses in spoken test and 53% in written test) and that the Danish–Swedish mutual intelligibility is fairly high (57% and 63% correct) in the spoken test and even higher (86% and 89%) in the written test.

2.2.2 Inherited intelligibility

The results of the sub-group with a minimum of exposure to the test language are shown by the grey bars in Figures 2 and 3. A comparison of the black bars to the grey bars makes clear that the overall intelligibility is lower as a result of the more stringent selection criteria. The mean intelligibility within the Germanic group has decreased from 27% to 18% in the case of the spoken test and from 39% to 31% in the case of written intelligibility. Swedes and Danes reach the highest degree of mutual understanding (44% and 56% in the spoken cloze test and 89% and 82% in the written cloze test). The Danish–Swedish mutual intelligibility is good enough to allow for successful communication between Swedes and Danes when they engage in receptive multilingualism, though with some effort (see Section 1 above). Germans understand some Dutch (25%) and we know from the literature that receptive multilingualism is used as a means of communication, especially in the Dutch–German border regions (Ház 2005). However, at a first confrontation speakers of Dutch and German will mostly only be able to communicate at a very basic level.

3. MEASURING PREDICTORS OF INTELLIGIBILITY

3.1 Linguistic distances

3.1.1 Lexical distance

The most common way of calculating lexical distance is to count the number of non-cognates (words that are not historically related) in the native language of the

participant and the test language. This way of quantifying lexical distances has been used previously to measure lexical distances between varieties within different language areas (e.g. Goebel 1993, Heeringa & Nerbonne 2006, Giesbers 2008). Generally, the larger the percentage of cognates shared between two languages, the better the mutual intelligibility between them. Lexical distance may be asymmetrical. A word in language A may have a cognate in language B but a synonym in B need not have a cognate synonym in A. The Dutch word *dienst* ‘service’, for example, has no cognate in English. The equivalent for *dienst* in English is *service*. For the English word *service*, however, there is a cognate in Dutch, namely *service*, which is a synonym of the Dutch stimulus word *dienst*. For this reason the lexical distances were calculated both ways for each language combination. We calculated the distances on the basis of all words in the four texts used for the intelligibility experiments.

3.1.2 Orthographic stem distance

For measuring orthographic distances, the same corpus was used as for measuring lexical distances. For each cognate pair in a test language and participant language, the orthographic distance was calculated by means of the Levenshtein algorithm, which computes the smallest number of string edit operations needed to convert the orthographic string in language A to the cognate string in language B (Nerbonne & Heeringa 2010). Possible string operations are deletions, insertions and substitutions of symbols. For example, if we calculate the orthographic distance between Dutch *oord* and its German cognate *Ort* ‘place’, we see that the *o* is deleted and the *d* is replaced by a *t*. This comes down to a total of two string operations. If we divide the number of operations (2) by the total number of alignments (4), we get a distance of 50% (see Table 1). The overall orthographic distance between two languages is the arithmetic mean of the normalised distances for all cognate word pairs in a corpus. We calculated distances for the word stems only.

Alignment	1	2	3	4
Dutch	o	o	r	d
German		O	r	t
Penalty points	1			1

Table 1. Example of the Levenshtein procedure.

In the weighing of the differences between two letters, a distinction was made between differences in the base of the letter and differences in diacritics. Comparing Danish *hånd* with English *hand*, for example, we see that the *å* in the Danish form contains a diacritic, while the *a* in the English form does not. Differences like this were weighed 0.3. Differences in the base of the character, like the difference

between German *Prozent* and Dutch *procent* ‘percent’ were weighed 1. The maximum weighing was 1, so if two characters differed in both base and diacritics (e.g. *o* versus *ô*), they got a weighing of 1 and not of 1.3.

3.1.3 Affix distance

The above orthographic distance is computed on the stem of words. Heeringa et al. (2014) found low correlations between stem and affix distances and therefore concluded that both measures should be included in a model of intelligibility. The affix distance is calculated in the same manner as the stem distances, in principle, but the aligned strings contain the affix portions of the cognate words only. This measure serves as an estimate of the degree of morphological similarity between two related languages (and excludes the similarity between the word stems).

3.1.4 Phonological distance

Phonological distance was measured on the basis of word lists that were used in a word translation task that was also part of the project. First, broad phonological transcriptions of each word in the stimuli lists of each language were made. Distances were measured between the phonetic transcriptions of all cognate pairs in the lists by means of the Levenshtein algorithm. The weights attributed to each operation vary between 0 and 1 using the gradual weighing for the phonological distances determined by Heeringa (2004).

3.1.5 Syntactic distance

Although syntactic differences between related languages are generally small, hearing (or seeing) words in unusual positions in a sentence may compromise intelligibility. In order to measure syntactic distances, we used the text corpus, based on the same four texts that we used for the cloze tests. These texts consist of 66 sentences in total. We translated the stimulus texts in each of the languages as literally as possible, into the other four languages (i.e. changing the word order in the test language as little as possible to conform to the grammar of the participant language without making the sentence ungrammatical). As a second step, we tagged the words in each sentence with respect to syntactic word class. The beginning and the end of a sentence was marked with a \$ sign. Since syntactic distances were for the first time used to predict intelligibility in this project we used three different syntactic measurements: (i) the movement measure, (ii) the indel measure and (iii) the trigram measure to see which one reflects intelligibility best. These three methods will be explained below. A more detailed discussion of the three measurements can be found in Heeringa et al. (2017).

3.1.5.1 The movement measure

The movement measure is based on the number of words that are moved when translating a sentence from language A into language B. A movement consists of a deletion of a word in one position in the sentence and an insertion of that word at another position in the sentence. An example of a movement can be found in (1) where the infinitives of 'to do' *doen/tun* are in different positions in the Dutch and German translations of the sentence *Therefore it is often difficult to know what to do*.

(1) Dutch: Daarom is het vaak moeilijk te weten wat je moet **doen**.

German: Darum ist es oft schwierig zu wissen, was man **tun** muss.

Each word in both sentences was given a code identifying its ordinal position from left to right in the source sentence. Corresponding words got the same code. Thereafter, the distance between the sentence pairs was calculated with help of the Levenshtein procedure as shown in Table 2.

1	2	3	4	5	6	7	8	9	10	11	12
Daarom	is	het	vaak	moeilijk	te	weten	wat	je		moet	doen
Darum	ist	es	oft	schwierig	zu	wissen	was	man	tun	muss	
0	0	0	0	0	0	0	0	0	1	0	1

Table 2. Alignment of the Dutch and the German sentences in (1).

The number of positions a word moved is counted. This means that the more positions a word is moved, the more negative the effect on intelligibility is. In example (1) the word *doen* has moved two positions. The distances between each language combination is the average of the syntactic distance of all the 66 sentences in the corpus.

3.1.5.2 The indel measure

The indel measure is similar to the movement measure. The difference is that we are not looking at words that have moved in a sentence, but we are looking at words that have been inserted or deleted when translating a sentence from language A into language B. Compare the Danish sentence and its English translation in (2).

(2) Danish: Jeg får en masse breve på denne tid af året.

English: I get a lot of letters at this time of year.

Aligning the sentences according to the Levenshtein procedure is shown in Table 3.

1	2	3	4	5	6	7	8	9	10	11
Jeg	får	en	masse		breve	på	denne	tid	af	året
I	get	a	lot	of	letters	at	this	time	of	year
0	0	0	0	1	0	0	0	0	0	0

Table 3. Alignment of the Danish and the English sentences in (2).

In the English sentence only the word *of* is inserted at position 5. So the indel measure is $1/11 = .09$. In the indel measure, the aggregate distance is defined as the mean syntactic indel distances for the 66 sentences in the corpus.

3.1.5.3 The trigram measure

In the trigram measure, the syntactic distance is measured by computing the correlation between syntactic trigram frequencies between two languages (see Nerbonne & Wiersma 2006, Hirst & Feiguina 2007). For each corpus, a frequency vector is created. How high each frequency is depends on how often each trigram occurs in the corpus and how large the corpus is. For example, in the English sentence in Table 4 we find six trigrams and in the sentence in Table 5 we find seven trigrams.

Comparing the two sentences of this small corpus we find 11 different trigrams. Two of these occur twice (pronoun – modal verb – verb and modal verb – verb – adverb) the other trigrams occur only once each. We calculated the correlations between the frequencies of every occurring trigram in the language corpora of each language combination. The syntactic distance between the two languages is then computed as 1 minus the Pearson's correlation coefficient between the pairs of trigram frequencies.

	It	would	be	difficult	to	cycle	
\$	pronoun	modal verb					
	pronoun	modal verb	verb				
		modal verb	verb	adverb			
			verb	adverb	to		
				adverb	to	verb	
					to	verb	#

\$ marks the beginning of a sentence; # marks the end of a sentence

Table 4. Trigrams in the sentence *It would be difficult to cycle*.

	After	a	while	it	will	become	easier	
\$	preposition							
	preposition	noun						
		noun	determiner					
		noun	determiner	pronoun				
			determiner	pronoun	modal verb			
				pronoun	modal verb	verb		
					modal verb	verb	adverb	
						verb	adverb	#

\$ marks the beginning of a sentence; # marks the end of a sentence

Table 5. Trigrams in the sentence *After a while it will become easier.*

3.2 Extra-linguistic predictors

We included three extra-linguistic predictors in our analysis: attitude, exposure and the number of years the participants had learnt the test language at school.

Before the intelligibility test we asked the participants twice about their attitude towards the test language, first without a speech fragment and next after having played a sample of the test language (see Section 2.1.3 above). The two attitude tests correlate highly ($r = .85, p < .001$). We therefore used the mean across the two scales as our measurement of attitude.

The exposure questions contained six five-point scales (see Section 2.1.3). We correlated the scores on the six scales. Since all scales correlated significantly we created a total exposure index by computing the mean across the six scales. This is the variable that we will use in our analysis in Section 4. In addition to information on amount of exposure we also included the number of years the participants had learned the language at school as a predictor.

3.3 Results

For each of the language combinations, Table 6 lists the scores obtained for each of the seven linguistic distance types as well as the results for the extra-linguistic measurements. Attitude, exposure and years of learning are the means across participants in the spoken and written cloze test. The results will serve as input for the regression analyses presented in the next section. A closer look at Table 6 makes clear that all variables show a large variability. The linguistic distance between Swedish and Danish is generally the smallest. Exceptions are phonetic and affix distances that are smallest between Dutch and German. With the exception of affix distance that is largest between Swedish and German, the largest distance between language pairs

Language combination		Linguistic							Extra-linguistic		
Participant	Test	Lex	Stem	Affix	Phon	Move	Indel	Trigr	Att	Exp	Learn
Danish	Dutch	45.6	44.6	32.5	48.4	7.17	1.79	0.32	1.87	1.22	0.00
Danish	English	48.5	51.2	28.5	53.0	1.18	2.12	0.19	3.83	3.98	10.77
Danish	German	40.0	47.6	34.8	46.6	6.73	1.39	0.35	2.50	1.77	4.49
Danish	Swedish	4.6	24.1	22.2	46.3	0.27	0.61	0.07	3.71	2.05	0.00
Dutch	Danish	43.9	44.6	32.5	48.4	6.08	2.62	0.31	3.06	1.23	0.00
Dutch	English	38.1	53.3	20.4	47.6	8.36	2.61	0.33	3.76	3.86	7.97
Dutch	German	19.8	44.6	18.1	36.7	0.59	1.06	0.07	2.92	1.98	4.04
Dutch	Swedish	46.8	45.1	38.4	45.2	5.74	3.02	0.31	3.56	1.31	0.00
English	Danish	49.2	51.2	28.5	53.0	0.89	1.88	0.16	2.32	1.22	0.00
English	Dutch	39.0	53.3	20.4	47.6	9.79	1.73	0.32	2.68	1.36	0.00
English	German	46.4	59.6	29.7	50.4	8.45	1.77	0.44	2.79	1.64	1.84
English	Swedish	51.4	55.0	26.0	48.3	1.47	1.77	0.14	2.94	1.14	0.00
German	Danish	39.0	47.6	34.8	46.6	7.18	1.85	0.34	2.92	1.25	0.00
German	Dutch	21.5	44.6	18.1	36.7	3.27	1.92	0.17	2.86	1.49	0.00
German	English	44.5	59.6	29.7	50.4	9.55	2.86	0.40	4.14	3.24	9.37
German	Swedish	39.9	48.1	40.7	44.7	6.44	1.77	0.35	3.06	1.16	0.00
Swedish	Danish	5.8	24.1	22.2	46.3	0.14	0.68	0.04	2.42	1.90	0.00
Swedish	Dutch	47.8	45.1	38.4	45.2	7.39	2.76	0.36	2.13	1.24	0.00
Swedish	English	49.5	55.0	26.0	48.3	3.23	2.71	0.20	3.79	3.68	9.86
Swedish	German	40.4	48.1	40.7	44.7	8.21	1.95	0.37	3.10	1.65	2.18

Table 6. Seven linguistic distance variables: lexical (Lex), orthographic stem (Stem), orthographic affix (Affix), phonetic (Phon), movement (Move), indels (Indel) and trigrams (Trigr); and three extra-linguistic variables: attitude (Att) and exposure (Exp) and years of learning (Learn) for the 20 combinations of participant language and test language.

all involve English and one of the other languages. The participants are most positive towards English and the lowest mean attitude score is found among Danes evaluating Dutch. The participants have had most exposure to English and learned the language for the longest length of time, while they had never learned Danish and Dutch and had little exposure to these languages. We will discuss the relationship between the linguistic and extra-linguistic variables in Section 4.

4. THE RELATIONSHIP BETWEEN INTELLIGIBILITY AND PREDICTORS

In this section we will examine the relationship between the linguistic and extra-linguistic measurements presented in Table 6 and the intelligibility scores presented

in Figures 2 and 3 (see Section 2.2) by means of a stepwise regression analysis. However, to be able to interpret the results of this analysis we first present the correlations between the dependent and independent predictors.

4.1 Predictors of overall intelligibility

In the top triangle of Table 7 we present the correlations between the linguistic and extra-linguistic measurements. The correlations with the extra-linguistic predictors are based on the data from all participants in the spoken and written cloze test. A majority of the participants (69%) had previous exposure to the test language or had learned the test language (acquired intelligibility). Inspection of the table reveals that all linguistic levels are significantly intercorrelated with the exception of the phonetic and the affix levels, which in general behave rather independently of the other levels. The three extra-linguistic predictors also have high intercorrelations, but there are no significant intercorrelations between linguistic and extra-linguistic predictors. The two columns on the right side of the table show that there are no significant correlations between the acquired intelligibility scores in the two close tests and linguistic distance measurements, with the exception of the affix level that correlates significantly at the .05 level ($r = -.49$). The extra-linguistic predictors on the other hand all correlate significantly with the intelligibility scores. Especially the amount of exposure and the number of years the test language has been learned are good predictors of intelligibility with correlations between .76 and .93. We will come back to the correlations presented in the bottom triangle of Table 7.

In order to be able to draw a conclusion about the contribution of the ten variables in Table 7 to the prediction of intelligibility we regressed the intelligibility scores against all linguistic and extra-linguistic predictors in a step-wise procedure. Table 8 specifies the results of the regression analysis for the spoken and the written test. The variability in the intelligibility scores obtained by each of the 20 language combinations can be predicted well with just three predictors in the case of the spoken close test and with four predictors in the case of the written cloze test. The most powerful predictor is the amount of exposure to the test language reported by the individual participant. This predictor by itself accounts for 86% and 74% of the variance. In addition, years of learning contributes to the model while the third extra-linguistic predictor, attitude, does not make a contribution. When it comes to linguistic predictors, lexical distance is the only predictor that makes a significant contribution to predicting intelligibility. Including this predictor in the regression model increments the R^2 value by seven points in the case of the spoken close test and with twenty points in the case of the written close test. Orthographic stem adds another point to the model of written intelligibility.

	Lex	Stem	Affix	Phon	Move	Indel	Trigr	Att	Exp	Learn	Spk – all	Wrt – all
Lex		.82**	.52*	.54*	.45*	.74**	.61**	.06	.03	.24	-.26	-.42
Stem			.15	.33	.51*	.61**	.59**	.20	.22	.43	.00	-.20
Affix				.22	.43	.38	.63**	-.14	-.34	-.18	-.44	-.49*
Phon					.13	.25	.26	.13	.27	.28	.03	-.05
Move						.50*	.94**	-.02	-.07	.04	-.21	-.33
Indel							.57**	.34	.25	.34	-.02	-.08
Trigr								-.01	-.08	.07	-.25	-.40
Att	-.15	-.09	.01	-.28	-.05	.13	-.06		.71**	.66**	.60**	.57**
Exp	-.64**	-.58**	-.35	-.05	-.42*	-.52**	-.43**	-.04		.94**	.93**	.86**
Learn	—	—	—	—	—	—	—	—	—		.86**	.76**
Spk – inh	-.95**	-.91**	-.49	-.28	-.61*	-.74**	-.72**	-.03	.78**	—	.92**	.97**
Wrt – inh	-.94**	-.95**	-.51	.31	-.69**	-.63	-.78**	.13	.92**	—	.82**	.97**

* $p < .05$; ** $p < .01$

Table 7. Correlation matrix for 20 language combinations computed for seven linguistic distance variables: lexical (Lex), orthographic stem (Stem), orthographic affix (Affix), phonetic (Phon), movement (Move), indels (Indel) and trigrams (Trigr); three extra-linguistic variables: attitude (Att), exposure (Exp) and years of learning (Learn); and two intelligibility tests: spoken (Spk) and written (Wrt) cloze tests. In the top triangle the correlations with all test data are shown and in the bottom triangle are the test scores of participants with a minimum of exposure (inherited (inh) intelligibility).

Model	Spoken cloze test	R^2
1	Exposure	.86
2	Exposure + Lexical	.93
3	Exposure + Lexical + Years of learning	.98
Model	Written cloze test	R^2
1	Exposure	.74
2	Exposure + Lexical	.94
3	Exposure + Lexical + Years of learning	.96
4	Exposure + Lexical + Years of learning + Orthographic stem	.97

Table 8. Stepwise regression model optimally accounting for intelligibility in 20 language combinations. R^2 values are cumulative.

4.2 Predictors of inherited intelligibility

It is obvious that exposure to the test language is a factor of overriding importance when it comes to intelligibility of a closely related language. Participants who are familiar with the non-native language will understand it better than those who have had no prior exposure to it. The intelligibility scores presented in black bars in Figures 2 and 3 might therefore be largely a matter of acquired intelligibility (see Section 1). We will now present a second analysis, in which we zoom in on inherited intelligibility (the grey bars in Figure 2 and 3). Since the participants have hardly had any exposure to the test language, the only source of information from which we may predict the differences in intelligibility is the linguistic distance between the languages.

The correlations of the results of the inherited intelligibility scores with the linguistic and extra-linguistic predictors are presented in the two bottom rows of Table 7 above. Most linguistic predictors now correlate significantly with the intelligibility scores. Exceptions are affix distance and phonetic distance. We introduced three ways of measuring syntactic distance as predictors in our statistical model of intelligibility. All three measures correlate significantly with spoken intelligibility, with indel distance having the highest correlation. Move and trigram measurements correlate significantly with written intelligibility, with trigram measurements showing the highest correlation.

Attitude scores do not correlate with overall intelligibility but exposure shows a high correlation even though none of the participants had much previous exposure to the test language. Therefore, even though the participants had little exposure to the test language, exposure is still a strong predictor of intelligibility. The high correlation between intelligibility and exposure may at least partly be explained by covariation with linguistic distances. In general, linguistically similar varieties are spoken in places that are in close geographical proximity. If a variety is spoken in an

Model	Spoken cloze test	R^2
1	Lexical	.89
2	Lexical + Orthographic stem	.93
Model	Written cloze test	R^2
1	Orthographic stem	.89
2	Orthographic stem + Orthographic affix	.96

Table 9. Stepwise regression model optimally accounting for intelligibility in 14 language combinations with a minimum of exposure. R^2 values are cumulative.

area close by, participants are likely to have had more exposure to them than if it is spoken in a geographically more remote place. In the lower part of Table 7 above we see significant correlations between exposure and all linguistic distances except for phonetic and affix distances. Still, exposure cannot be excluded as a factor that may have had some influence on the results, since some participants had some exposure to the test language.

We tested the predictive strength of the linguistic and extra-linguistic predictors in a stepwise multiple regression analysis. The results are summarized in Table 9. Prediction of inherited intelligibility is possible with considerable success: 93% of the variance in the spoken cloze test scores can be predicted by a combination of two linguistic predictors i.e. lexical and orthographic stem distance between the participant's native language and the non-native related test language, and 96% of the variance in the written cloze test is explained by orthographic stem and affix distances. None of the extra-linguistic predictors are included in the models. It is unexpected that the orthographic distances are included in a model of spoken intelligibility rather than phonetic distances. A closer inspection of the data revealed that the low correlation between intelligibility and phonetic data (see Table 7) can be explained by two outliers, Danish participants listening to Swedish and Swedish participants listening to Danish. The Swedish–Danish mutual intelligibility is much higher than could be expected on the basis of phonetic distances, presumably due to the very small lexical distances between the two languages (see Table 6). If we exclude the Danish–Swedish and the Swedish–Danish language combinations from our analysis we get a correlation of .87 between intelligibility scores and phonetic distances and a regression model that only includes phonetic distances and explains 76% of the variance.

5. CONCLUSIONS

The first conclusion that can be drawn from the results presented in this paper is that there are large differences in intelligibility among the various Germanic language

combinations. High scores are found where the test language is a school language (English and German). Our results also confirm the finding from publications on inter-Scandinavian intelligibility that the mutual intelligibility between Danish and Swedish is reasonably high. Previous investigations have shown that Danes understand spoken Swedish better than Swedes understand Danish. In contrast to previous investigations (see [Section 1](#)) we did not find a significant asymmetry in the Danish–Swedish mutual intelligibility at the .01 level (Bonferroni test). We used a different intelligibility test than in previous investigations where mostly open questions about a text were used. It is possible that the cloze test for some reason has not been able to capture the asymmetric Danish–Swedish mutual intelligibility. However, Gooskens & Van Heuven (2017) have shown that the cloze test correlates highly with a word translation task ($r = .73$ for spoken and $.79$ for written intelligibility) and with the intelligibility as perceived by the listeners (correlations between $.94$ and $.99$). Therefore we have no reason to conclude that the cloze test might be less suitable for testing intelligibility than other tests. The other asymmetric relationship that is often mentioned in the literature is the mutual intelligibility between Dutch and German. This asymmetry was also found in our results ($p < .01$). German is a school language for many children in the other Germanic countries, which explains why German is so well understood even among listeners from the North Germanic group. The lowest scores are found with combinations of languages where listeners and test languages are from different main branches of the Germanic language tree (West or North), such as Swedish–Dutch, as well as intelligibility of Dutch for English listeners.

As expected, we found that a closely related language is easier to understand in the written form than in the spoken form. It is easier to grasp the concrete written form of the language than the volatile spoken form. In addition, the written form of a closely related language is likely to be easier to match with the corresponding form in the native language because written language reflects an earlier form of the language where the two languages had diverged less from the original common form than in the spoken form (Doetjes & Gooskens 2009, Schüppert 2011).

The main purpose of this paper was to determine the relative importance of linguistic and extra-linguistic predictors of intelligibility in the Germanic language area. A large part of the acquired intelligibility appears to be due to extra-linguistic factors, of which prior exposure to the test language is the most important one. This shows that to improve mutual understanding it is important to create opportunities for people to be exposed to the neighbouring language, for example through school exchange programs. Interestingly, positive or negative attitudes towards the test language do not make a significant contribution to the prediction of mutual intelligibility. Also in previous research the link between intelligibility and attitude has mostly been rather weak (see [Section 1](#)). It is possible that this is caused by the experimental settings where participants may be less strongly influenced by their attitudes towards the language and its speakers than in real life.

If we eliminate the influence of extra-linguistic factors as well as possible by selecting participants with little or no prior exposure to the test language only the inherent similarity between the related languages remains as a way to crack the code of the non-native test language. Our results show that we are able to predict inherent intelligibility to a high extent. All linguistic distance measures, except phonetic and affix distances, correlate significantly with spoken and written intelligibility, with lexical distances explaining most of the variance in a regression model of spoken intelligibility and orthographic distances explaining most of the variance in a model of written intelligibility. We noted that Danish–Swedish mutual intelligibility is much higher than expected on the basis of phonetic distances. If we remove the Danish–Swedish outliers from our data set, phonetic distances turn out to be the best predictor of spoken intelligibility.

In addition to socio-political and cultural status, linguistic distance has been used as one of the criteria for distinguishing between dialects and languages (Kloss 1967). It is a problem, however, that languages may differ to different extents in their lexicon, phonetics and phonology, morphology and syntax. It is difficult to decide how much weight should be given to each of these linguistic dimensions when determining overall distance. Maybe to circumvent this problem, Trudgill (1986) introduced the intelligibility criterion and this has become the primary criterion among many linguists. According to this criterion, dialects are mutually intelligible varieties, whereas languages are so linguistically different that their speakers are unable to understand each other. The high predictive power of linguistic distances confirm that intelligibility measurements can indeed be used as an alternative to linguistic distance measurements.

Traditionally, Scandinavians communicate by means of receptive multilingualism and Scandinavian authorities therefore strongly encourage cross-border communication in the native languages of the inhabitants. To the individual citizen there are many advantages of being able to express themselves in their own languages. Previous research (e.g. Hedquist 1985) has shown that in the case of less closely related languages such as for example Swedish and Dutch a short course that makes speakers conscious of the most important differences and similarities between their native language and the language of the speaker can improve mutual intelligibility considerably. It is therefore of great importance to explore the possibilities for communication with receptive multilingualism outside Scandinavia and to carry out more research into what specific linguistic differences between the languages and what strategies the speakers should be taught to communicate effectively in the receptive multilingualism mode.

ACKNOWLEDGEMENTS

This research was supported by the Netherlands Organisation for Scientific Research (NWO) grant number 360-70-430 (principal investigators Charlotte Gooskens and

Vincent van Heuven). We thank Aleksandar Mancic for programming the web application, Wilbert Heeringa for calculating the linguistic distances and three anonymous reviewers for valuable comments.

NOTES

1. The investigation reported here was part of a larger project including 20 Germanic, 20 Romance and 30 Slavic language combinations and several online intelligibility tests (see Golubović 2016, Swarte 2016, Gooskens et al. 2017).
2. At the B1-level the learner is able to independently use the language. The learner can understand texts that are related to work, school and free time, and is able to produce simple texts relating to these contexts.

REFERENCES

- Abraham, Roberta G. & Carol A. Chapelle. 1992. The meaning of cloze test scores: An item difficulty perspective. *The Modern Language Journal* 7, 468–479.
- Bø, Inge. 1978. *Ungdom og naboland. En undersøkelse av skolens og fjernsynets betydning for nabospråksforståelsen* [Youth and neighboring country: An investigation of the influence of school and TV on inter-Scandinavian comprehension]. Stavanger: Rogalandforskning.
- Börestam Uhlmann, Ulla. 1991. *Språkmöten och mötesspråk i Norden* [Language meetings and meeting languages in the Nordic countries] (Nordisk språksekretariats rapporter 16). Oslo: Nordisk språksekretariat.
- Bormuth, John. 1969. Factor validity of cloze tests as measures of reading comprehension ability. *Reading Research Quarterly* 4, 358–365.
- Braunmüller, Kurt. 2007. Receptive multilingualism in Northern Europe in the Middle Ages: A description of a scenario. In Jan D. ten Thije & Ludger Zeevaert (eds.), *Receptive Multilingualism*, 25–47. Amsterdam: John Benjamins.
- Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Delsing, Lars-Olof & Katarina Lundin Åkesson. 2005. *Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska* [Does the language keep together the Nordic countries? A research report of mutual comprehension between young speakers of Danish, Swedish and Norwegian]. Copenhagen: Nordiska ministerrådet.
- Doetjes, Gerard & Charlotte Gooskens. 2009. Skriftsprogets rolle i den dansk-svenske talesprogsforståelse [The role of written language in spoken Danish–Swedish intelligibility]. *Språk och Stil* 19, 105–123.
- Giesbers, Charlotte. 2008. *Dialecten op de grens van twee talen; een dialectologisch en sociolinguïstisch onderzoek in het Kleverlands dialectgebied* [Dialects on the border of two languages: A dialectological and sociolinguistic investigation in the Kleverlandic dialect area]. Ph.D. dissertation, Radboud University Nijmegen.
- Goebel, Hans. 1993. Probleme und Methoden der Dialektometrie: Geo-linguistik in Globaler Perspektive [Problems and methods of dialecto-metry: Geolinguistics in a global

- perspective]. In Wolfgang Viereck (ed.), *Proceedings of the International Congress of Dialectologists 1*, 37–81. Stuttgart: Franz Steiner Verlag.
- Golubović, Jelena. 2016. *Mutual Intelligibility in the Slavic Language Area* (Groningen Dissertations in Linguistics). Ph.D. dissertation, University of Groningen.
- Gooskens, Charlotte. 2006a. Linguistic and extra-linguistic predictors of inter-Scandinavian intelligibility. In Jeroen van de Weijer & Bettelou Los (eds.), *Linguistics in the Netherlands 23*, 101–113. Amsterdam: John Benjamins.
- Gooskens, Charlotte. 2006b. Hvad forstår svenskere og nordmænd bedst – engelsk eller dansk? [What do Swedes and Norwegians understand best: English or Danish?]. *Tidskrift for Sprogforskning* 4(1), 221–224.
- Gooskens, Charlotte & Renée van Bezooijen. 2006. Mutual comprehensibility of written Afrikans and Dutch: Symmetrical or asymmetrical. *Literary and Linguistic Computing* 21, 543–557.
- Gooskens, Charlotte, Renée van Bezooijen & Vincent J. van Heuven. 2015. Mutual intelligibility of Dutch–German cognates by children: The devil is in the detail. *Linguistics* 53, 255–283.
- Gooskens, Charlotte & Vincent J. van Heuven. 2017. Measuring cross-linguistic intelligibility in the Germanic, Romance and Slavic language groups. *Speech Communication* 89, 25–36.
- Gooskens, Charlotte, Vincent J. van Heuven, Jelena Golubović, Anja Schüppert, Femke Swarte & Stefanie Voigt. 2017. Mutual intelligibility between closely related languages in Europe. Ms., University of Groningen.
- Haugen, Einar. 1966. Semicommunication: The language gap in Scandinavia. *Sociological Inquiry* 36, 280–297.
- Ház, Éva. 2005. *Deutsche und Niederländer. Untersuchungen zur Möglichkeit einer unmittelbaren Verständigung* (Philologia 68). Hamburg: Dr. Kovač.
- Hedquist, Rolf. 1985. *Nederländares förståelse av danska och svenska. En språkpedagogisk undersökning med utnyttjande av likheterna mellan språken* [Dutchmen's understanding of Danish and Swedish: A linguistic pedagogical investigation with reference to similarities between the languages]. Umeå: Umeå universitet.
- Hedquist, Rolf & Eva Strangert. 1989. *Hur svenskar uppfattar och förstår nederländska ord* [How Swedes perceive and understand Dutch words]. Umeå: Umeå universitet, Institutionerna för fonetik och nordiska språk.
- Heeringa, Wilbert. 2004. *Measuring Dialect Pronunciation Differences Using Levenshtein Distance* (Groningen Dissertations in Linguistics). Ph.D. dissertation, University of Groningen.
- Heeringa, Wilbert & John Nerbonne. 2006. De analyse van taalvariatie in het Nederlandse dialectgebied: Methoden en resultaten op basis van lexicon en uitspraak [An analysis of language variation in the Netherlandic dialect area: Methods and results based on lexicon and pronunciation]. *Nederlandse Taalkunde* 11, 218–257.
- Heeringa, Wilbert, Femke Swarte, Anja Schüppert & Charlotte Gooskens. 2014. Modeling intelligibility of written Germanic languages: Do we need to distinguish between orthographic stem and affix variation? *Journal of Germanic Linguistics* 26(4), 361–394.
- Heeringa, Wilbert, Femke Swarte, Anja Schüppert & Charlotte Gooskens. 2017. Modeling intelligibility of written Germanic languages: How do we represent syntactical variation? *Digital Scholarship in the Humanities*, <https://doi.org/10.1093/llc/fqx029>. Published online by Oxford University Press, 19 June 2017.

- Hilton, Nanna Haug, Charlotte Gooskens & Anja Schüppert. 2013. The influence of non-native morphosyntax on the intelligibility of a closely related language. *Lingua* 137, 1–18.
- Hirst, Graeme & Ol'ga Feiguina. 2007. Bigrams of syntactic labels for authorship discrimination of short texts. *Literary & Linguistic Computing* 22, 405–419.
- Impe, Leen. 2010. *Mutual Intelligibility of National and Regional Varieties of Dutch in the Low Countries*. Ph.D. dissertation, University of Leuven.
- Jørgensen, Nils & Eva Kärrlander. 2001. *Grannspråksförståelse i Öresundsregionen år 2000. Gymnasisters hörförståelse* [Mutual intelligibility in the Öresundsregion year 2000: The listening comprehension of secondary school students]. Lund: University of Lund.
- Keshavarz, Mohammad Hossein & Hossein Salimi. 2007. Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics* 17, 81–92.
- Kloss, Heinz. 1967. 'Abstand languages' and 'ausbau languages'. *Anthropological Linguistics* 9(7), 29–41.
- Kuhlemeier, Hans, Huub van denBergh & Leijn Melse. 1996. Attitudes and achievements in the first year of German language instruction in Dutch secondary education. *The Modern Language Journal* 80, 494–508.
- Kürschner, Sebastian, Renée van Bezooijen & Charlotte Gooskens. 2008. Linguistic determinants of the intelligibility of Swedish words among Danes. *International Journal of Humanities and Arts Computing* 2, 83–100.
- Lundin, Katarina & Robert Zola Christensen. 2001. *Grannspråksförståelse i Öresundsregionen år 2000. Gymnasisters läsförståelse* [Mutual intelligibility in the Öresund region year 2000: The reading comprehension of secondary school students]. Lund: University of Lund.
- Maurud, Øivind. 1976. Reciprocal comprehension of neighbour languages in Scandinavia: An investigation of how well people in Denmark, Norway and Sweden understand each other's languages. *Scandinavian Journal of Educational Research* 20, 49–71.
- Nerbonne, John & Wilbert Heeringa. 2010. Measuring dialect differences. In Jürgen Erich Schmidt & Peter Auer (eds.), *Language and Space: Theories and Methods* (Handbooks of Linguistics and Communication Science), 550–567. Berlin: Mouton De Gruyter.
- Nerbonne, John & Wybo Wiersma. 2006. A measure of aggregate syntactic distance. In John Nerbonne & Erhard Hinrichs (eds.), *Linguistic Distances Workshop at the Joint Conference of International Committee on Computational Linguistics and the Association for Computational Linguistics, Sydney, July 2006*, 82–90. Stroudsburg, PA: Association for Computational Linguistics.
- Schüppert, Anja. 2011. *Origin of Asymmetry: Mutual Intelligibility of Spoken Danish and Swedish* (Groningen Dissertations in Linguistics). Ph.D. dissertation, University of Groningen.
- Swarte, Femke. 2016. *Predicting the (Mutual) Intelligibility of Germanic Languages from Linguistic and Extra-linguistic Factors* (Groningen Dissertations in Linguistics). Ph.D. dissertation. University of Groningen.
- Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Blackwell.
- van Bezooijen, Renée & Charlotte Gooskens. 2005. How easy is it for speakers of Dutch to understand spoken and written Frisian and Afrikaans, and why? *Linguistics in the Netherlands* 22, 13–24.

- van Bezooijen, Renée, Charlotte Gooskens & Sebastian Kürschner. 2012. Deens is makkelijker voor Friezen dan voor Nederlanders – feit of fabel? [Danish is easier to understand for Frisians than for Dutchmen – fact or fairytale?]. In Piter Boersma, Goffe Th. Jensma & Reinier Salverda (eds.), *Philologia Frisica*, 286–298. Leeuwarden: Fryske Akademy.
- Vanhove, Jan. 2014. *Receptive Multilingualism across the Lifespan: Cognitive and Linguistic Factors in Cognate Guessing*. Ph.D. dissertation, University of Fribourg.