PHONETIC AND LEXICAL PREDICTORS OF INTELLIGIBILITY

CHARLOTTE GOOSKENS, WILBERT HEERINGA AND KARIN BEIJERING

Abstract *In the present investigation, the intelligibility of 17 Scandinavian language varieties and standard Danish was assessed among young Danes from Copenhagen. In addition, distances between standard Danish and each of the 17 varieties were measured at the lexical level and at different phonetic levels. In order to determine how well these linguistic levels can predict intelligibility, we correlated the intelligibility scores with the linguistic distances and we carried out a number of regression analyses. The results show that for this particular set of closely related language varieties phonetic distance is a better predictor of intelligibility than lexical distance. Consonant substitutions, vowel insertions and vowel shortenings contribute significantly to the prediction of intelligibility.*

1. INTRODUCTION

Gooskens (2007) correlated lexical and phonetic distances with mutual intelligibility scores for the Mainland Scandinavian standard languages, Danish, Norwegian and Swedish. Subjects from different places in Denmark, Norway and Sweden listened to the two standard languages spoken in the neighbouring countries and linguistic distances were measured between the language varieties of the listeners and the test languages. In total there were 18 mean intelligibility scores and 18 corresponding linguistic distances. The distances were measured at the two linguistic levels that are generally taken to be most important for mutual intelligibility in Scandinavia, namely the lexical and the phonetic level (Delsing and Lundin Åkesson, 2005; Torp, 1998). The results showed a high

International Journal of Humanities and Arts Computing 2 (1–2) 2008, 63–81
DOI: 10.3366/E1753854809000317

correlation between intelligibility scores and phonetic distances ($r = -.80$, $p < .01$). The correlation with lexical distances was low and not significant ($r = -.42$, $p = 0.11$), probably due to the fact that the lexical differences between the Scandinavian language varieties are small.

The purpose of the present paper is similar to the goal of the study we mentioned above: investigating the relationship between the intelligibility of closely-related language varieties and linguistic distances among these varieties. In our study a larger set of languages varieties is tested including *dialect* varieties, while Gooskens (2007) focused on standard languages. The intelligibility of 17 Scandinavian language and dialect varieties was tested among speakers of standard Danish. Like in the previous study, we correlated intelligibility scores with lexical and phonetic distances. However, in the present study the phonetic distances were also split up into consonant and vowel distances and subclassified by type (substitutions, insertions, deletions, lengthening and shortening of consonants and vowels).

Since only aggregate phonetic distances were included in the investigation by Gooskens (2007), no conclusions could be drawn about the nature of the phonetic differences that contribute most to intelligibility. It is generally assumed that consonants are more important than vowels for the identification of a word and that vowels carry less information than consonants. Consonants function as reference points in words while vowels tend to be more variable and change over time more rapidly than consonants do (Ashby and Maidment, 2005). Therefore, the occurrence of deviant segments in the consonantal structure of a word is presumably more disturbing for the intelligibility than changes in the vowels of a word. Ashby and Maidment illustrate this observation by an example. They assume that if all vowels in the sentence *Mary has a little lamb* are replaced by for example [ɛ], most people will probably still understand the sentence. However, when replacing all consonants with [d] and keeping the correct vowel qualities, the sentence will be unintelligible. The relative contribution of consonants and vowels to intelligibility may be different across languages since the size of consonant and vowel inventories can vary considerably and so can the number of vowels and consonants used in running speech. In the present investigation we measured consonant and vowel distances separately in order to test the hypothesis that consonants are more important for intelligibility than vowels in a Scandinavian setting.

We also took a closer look at the role of vowel and consonant distances subclassified by formal operation. We made a distinction between the following operations: substitutions, insertions, deletions, lengthening and shortening of vowels and consonants. The effect of these operations on intelligibility may vary. For example, substitutions change the basic framework of a word. In case of insertion or deletion of a segment, there is either a sound segment too many or too few in comparison with the native variety of the listeners and this may change

the structure of the words. Also the lengthening and shortening of segments may make words more difficult to understand.[1]

The research questions posed in the study are as follows:

1. What is the relative contribution of lexical and phonetic distances to the intelligibility of Scandinavian language varieties to standard Danish listeners?
2. What is the relative contribution of aggregate consonant and vowel distances to the intelligibility of Scandinavian language varieties to standard Danish listeners?
3. What is the relative contribution of subclassified consonant and vowel distances (insertions, deletions, substitutions, lengthening, shortening) to the intelligibility of Scandinavian language varieties to standard Danish listeners?

2. MATERIAL

We included recordings and transcriptions of the fable *The North Wind and the Sun*[2] in 18 different language varieties in our investigation (see Figure 1). One of the varieties, standard Danish, was only included in order to be able to check that the test was feasible. From a selection of recordings in more than 50 different Norwegian dialects, we chose eight dialects that form a good representation of the dialectological and geographical diversity of Norway.[3] In addition, we made extra recordings of Faroese (Torshavn), standard Swedish (as spoken in Stockholm), four Swedish dialects representing the four major dialect groups (including Finland Swedish), standard Danish (as spoken in Lyngby, close to Copenhagen) and three Danish dialects spoken on the peninsula of Jutland. The standard varieties of Danish, Norwegian[4] and Swedish all belong to the Mainland Scandinavian branch of the North Germanic language family and are known to be mutually intelligible. Speakers of these varieties can communicate with each other in their own language, though sometimes with some effort, see e.g. Delsing and Lundin Åkesson (2005). So far, the intelligibility of non-standard language varieties in a Scandinavian context has not been investigated. Faroese belongs to the Insular Nordic branch of the North Germanic language family and without prior instruction it is almost unintelligible to speakers of Mainland Scandinavian (Torp 1998: 34).

The Norwegian version of *The North Wind and the Sun* was first translated into Danish, Swedish and Faroese. The texts were then presented to speakers of the 18 varieties in the standard language of their country. The 18 text versions comprised between 91 and 111 words, with a mean of 98 words. The 18 recordings were used for the intelligibility experiment (see Section 3). In addition, phonetic transcriptions were made of all recordings.[5] These transcriptions were used to calculate the Levenshtein distances (see Section 4).

Figure 1. Map of Scandinavia with the locations of the 18 Scandinavian language varieties.

3. INTELLIGIBILITY

Design

The fable *The North Wind and the Sun* consists of six sentences. Each listener heard these six sentences, each sentence being presented in another Scandinavian language variety. In total 18 Scandinavian language varieties (including standard Danish) had to be tested. In order to be able to test all varieties we divided them into three groups. Every group contained a standard variety of one of the Mainland Scandinavian languages (Danish, Norwegian or Swedish). Furthermore, the three groups contained an equal number of Danish, Norwegian and Swedish dialects. We distributed the varieties among the groups in such a way that all groups contained both varieties that were likely to be difficult to understand, as well as varieties that were expected to be easy to understand. Faroese is the most deviant language variety in the test and the dialects of Oppdal and Storliden also differ considerably from standard Danish. Hence, each test contained one Danish, Norwegian or Swedish standard variety, one 'deviant'

variety, and an equal number of other Danish, Norwegian and Swedish dialects. In (1) an overview of the three groups is given. The abbreviations behind the dialects represent the language areas, i.e. NO stands for Norwegian, SW for Swedish, DA for Danish, FA for Faroese, and 's' preceding an abbreviation stands for the standard variety.

(1)

group 1	**group 2**	**group 3**
Oslo (sNO)	Stockholm (sSW)	Lyngby (sDA)
Torshavn (sFA)	Oppdal (NO)	Storliden (SW)
Høgsted (DA)	Hjordkær (DA)	Katrad (DA)
Tromsø (NO)	Bjugn (NO)	Rana (NO)
Fyresdal (NO)	Gaular (NO)	Trysil (NO)
Lidköping (SW)	Gryttinge (SW)	Finland Swedish (SW)

In order to test all sentences in all varieties, 18 different versions of the listening experiment were needed (6 versions per group). For example, test 1A contained sentence 1 in the Tromsø dialect, test 1B contained sentence 2 in the Tromsø dialect and so on (crossed design). The order of the sentences was randomised for each version, except that the first sentence, which contains the title of the fable and therefore reveals the content of the story, always occurred as the last sentence in the test. Furthermore, we also made sure that the sentences did not follow each other in the original order (for example sentence 3 preceding sentence 4). In this way, the listeners could not derive the content of a sentence from the context.[6]

Listeners

The subjects were 351 native speakers of standard Danish between 15 and 20 years of age (average 17.6) from 18 high-school classes in Copenhagen. Since the listeners lived in Copenhagen, we assumed that they all spoke standard Danish or at least were familiar with this language variety. Some of the listeners may have been familiar with some of the language varieties presented in the test. However, people living in Copenhagen do not generally have much contact with the Danish dialects of Jutland or the other Scandinavian dialects.

Procedure

While listening to the six sentences from the fable *The North Wind and the Sun*, the subjects had to translate each word they heard into standard Danish. Each sentence was presented twice. First the whole sentence was presented and next it was presented once more in smaller pieces of four to eight words, depending on the position of prosodic breaks. In this way we made sure that saturation

of the listeners' short-term memory would not influence the results and that the listeners had enough time to write down their translations.[7] Between the sentences there was a pause of 3 seconds and every sentence was preceded by a beep. The listening experiment started with an example sentence in Swedish (not from *The North Wind and the Sun*) so that the listeners could get used to the task.

Intelligibility scores

The percentage of correctly translated words constituted the intelligibility score of a given language variety. A correctly translated word was awarded one point and partly correctly translated words half a point. For example, if only the last part of the word *nordenvinden* 'The North Wind' was correctly translated, half a point was given.[8] We excluded Lyngby, representing standard Danish, from the analysis. This recording was only included to check that the test was feasible. Since 99 per cent of the Lyngby words were translated correctly, we conclude that this was indeed the case. So the remaining analyses were based upon 17 varieties.

4. LINGUISTIC DISTANCES

4.1 Phonetic distances

Phonetic distances between standard Danish (Lyngby) and each of the other 17 Scandinavian language varieties were calculated by means of the Levenshtein algorithm (see detailed explanations in Heeringa 2004). The distance between two phonetic transcriptions of two cognate pronunciations is calculated as the number of operations needed to transform one transcription into the other. There are three types of operations: insertions, deletions and substitutions of phonetic segments. The power of the Levenshtein distance is that it chooses the least costly set of operations that transform one pronunciation into another.

We will illustrate the algorithm with an example. The form *enige* (meaning 'in agreement') is pronounced as [ʔeːni] in Lyngby (standard Danish) and as [eːnɪgɑ] in Stockholm (standard Swedish). Ignoring suprasegmentals and diacritics (including length marks), the Levensthein algorithm will find the alignment as shown in (2).[9]

(2)

	1	2	3	4	5	6
Lyngby	ʔ	e	n	i		
Stockholm		e	n	ɪ	g	ɑ
	del			sub	ins	ins

The pronunciation of Lyngby is changed into the pronunciation of Stockholm by one deletion, one substitution and two insertions, four operations in total. Since we are using Levenshtein distance to model intelligibility, we want to calculate the extent to which a speaker of the Lyngby variety understands the pronunciation of Stockholm as a percentage. In our example we get six alignment slots, therefore the phonetic distance is $4/6 = 0.67$ or 67 per cent. This is the distance for one word. The distance for the whole text is the summed distances for all words divided by the number of words.

In order to obtain distances which are based on linguistically motivated alignments that respect the syllable structure of a word or the structure within a syllable, the algorithm was adapted so that a vowel may only correspond to a vowel and a consonant only to a consonant. The semi-vowels [j] and [w] may also correspond to a vowel or the other way around, their vowel counterparts [i] and [u] may correspond to a consonant or the reverse. The central vowel schwa [ə] may correspond to any sonorant. In this way, unlikely matches like [o] and [t] or [s] and [e] are prevented.

In the example above, all operations have the same cost. In the present study, we used graded operation weights calculated on the basis of sound samples from *The Sounds of the International Phonetic Alphabet* (1995). On the basis of the spectrograms, distances were measured between the IPA vowels and pulmonic consonants (see Heeringa 2004, pp. 97–107). We used a Barkfilter representation, which we consider a perceptually-oriented spectrogram since it has a more or less logarithmic frequency scale, a logarithmic power spectral density and the 24 first critical bands are modeled (see Heeringa 2004, pp. 87–88 for more details). The Barkfilter distances are used as operation weights. In this way the fact that for example [i] and [e] are phonetically closer to each other than [i] and [a] is taken into account. On average the Bark filter distances among vowels are smaller than among consonants. The vowel space is about one third of the consonant space (Heeringa 2004, p. 94). Gradual weights for insertions and deletions are obtained by measuring distances between the IPA sounds and silence. Using the Barkfilter representation, the glottal stop is closest to silence, and the [a] is most distant.

In validation work Heeringa (2004) found the tendency that Levenshtein distances based on logarithmic gradual segment distances approach perception better than Levenshtein distances based on linear gradual segment distances (see pp. 185–186). Although the Barkfilter represention already is logarithmic in itself since it has a logarithmic power spectral density, the use of logarithmic Barkfilter segment distances still gave some further improvement.

The length of different segment types is represented by changing the phonetic transcriptions as in (4).

(4) extra short sounds are retained as they are [ă] = a
normal sounds are doubled [a] = aa
half long sounds are trebled [aˑ] = aaa
long sounds are quadrupled [aː] = aaaa

Differences in length are formalised as insertions or deletions (indels), for example [a] versus [aː] is represented as *aa* versus *aaaa*, which results in two indels. Lengthening of a segment compared to standard Danish is processed as the insertion of a segment whereas shortening of a segment compared to standard Danish is processed as the deletion of a segment. Indels of this type are not regarded as 'real' indels, but as 'indels because of length difference' so that we are able to distinguish between insertions/deletions and lengthening/shortening. Diphthongs are processed as sequences of two monophthongs, or as sequences of a monophthong and a glide ([j] or [w]).

First Levenshtein distances on the basis of all segments (full phonetic distance) were calculated between each Scandinavian language variety and standard Danish. Together with the lexical distances (see Section 4.2) these distances were used to answer the first research question formulated in Section 1. Next, in order to be able to answer the second research question, the full phonetic distance was subdivided for consonants on the one hand and vowels on the other hand. When we calculated distances on the basis of consonants only, the full phonetic strings of the corresponding pronunciations were initially compared to each other using the Levenshtein algorithm. Once the optimal alignment was found, the distances were based on the alignment slots in which only consonants are involved, i.e. slots with either consonant insertion, deletion or substitution.[10] This distance was divided by the length of the full alignment. Vowel distance is calculated in a similar way, including only operations where vowels are involved.

We may illustrate this with example (2). The length of the alignment on the basis of all segments is 6. The slots 2, 4 and 6 concern vowels, involving a substitution and an insertion. Therefore the vowel distance is $2/6 = 0.33$. We find one vowel substitution, therefore the vowel substitution distance is $1/6 = 0.17$. Likewise the vowel insertion distance is $1/6 = 0.17$. We do not find vowel deletions, therefore the vowel deletion distance is $0/6 = 0$. For the consonant (sub)levels the distances may *mutatis mutandis* be calculated in the same way.

Above we mentioned that an [i] may match with a [j], and a [u] with a [w]. These matches are counted both as vowel (substitution) distances and consonant (substitution) distances.

4.2 Lexical distances

The lexical distance between standard Danish and the 17 other language varieties is measured in the simplest way like Séguy (1973) did. The lexical distance is expressed as the percentage of non-cognates (words that are not historically related to the corresponding words in standard Danish) in the 17 language varieties. For each word in the 17 varieties the corresponding cognate was aligned if existent. If no cognate exists the corresponding non-cognate was aligned. Some words may appear several times in the text. For example, *Nordenvinden* 'The North Wind' appears four times. In such cases four word pairs are considered.

Non-cognates are, in principle, unintelligible to listeners with no prior knowledge of the test variety. A large number of non-cognates should necessarily decrease the extent to which another language variety is intelligible. An example is *krangla* in the dialect of Fyresdal and *skændes* in standard Danish. Lexical differences hardly concern internationalisms in our material and therefore loan words are not excluded in the analyses.

The lexical and the phonetic measurements and the intelligibility scores are given in Appendix A, Tables A.1 and A.2.

5. THE RELATION BETWEEN LINGUISTIC DISTANCES AND INTELLIGIBILITY

We correlated the intelligibility scores of 17 Scandinavian language varieties (see Section 3) with lexical distance measures and different phonetic distance measures. In addition, we also correlated the major linguistic distances (lexical, phonetic, consonants and vowels) with all the linguistic distance measures. This provides information about colinearity. The results are shown in Table 1. In addition, we carried out regression analyses in order to investigate the relative contribution of the various linguistic levels to intelligibility. In Sections 5.1 to 5.3, the results of the correlations and the regression analyses that are relevant to the three research questions will be discussed.

5.1 Lexical and phonetic distances

We correlated the intelligibility scores of the 17 Scandinavian language varieties with the lexical and the phonetic distances, see Table 1. Both correlations are significant at the .01 level, and the correlation with phonetic distances is higher than with lexical distances, but not significantly higher ($r = -.86$ versus $r = -.64$, $p = .08$).[11] The corresponding scatterplots are presented in Figures 2 and 3. Most lexical distances are rather small (below five percent) and the

Table 1. Correlations between linguistic distances, intelligibility scores and major linguistic distances.

Linguistic levels	Intelligibility	Lexical	Phonetic	Consonants	Vowels
lexical	−.64**		.49*	.46	.50*
phonetic	−.86**	.49**		.88**	.36
consonants	−.74**	.46	.88**		.21
substitutions	−.57*	−.08	.66**	.56*	−.10
insertions	−.11	.47	.17	.43	.03
deletions	−.22	.13	.26	.25	.56
lengthening	.07	−.25	−.15	−.22	−.79**
shortening	−.22	.40	.42	.39	.11
vowels	−.29	.50*	.36	.21	
substitutions	.11	.16	−.18	−.34	.77*
insertions	−.39	.67**	.47	.53	.41
deletions	−.44	−.02	.53*	.35	.07
lengthening	−.42	.06	.51*	.37	−.45
shortening	−.49*	.21	.29	.07	.05
schwa vs. sonorant	.33	−.10	−.47	−.37	−.38

* $p \leq .05$; ** $p \leq .01$.

corresponding intelligibility scores vary greatly. As expected from the fact that Faroese belongs to another branch of the North Germanic language family, the Faroese variety from Torshavn is most deviant from standard Danish, both lexically and phonetically, and it was also most difficult to understand for the Danish listeners. Storliden is lexically and Oppdal phonetically almost as deviant as Faroese and both varieties are also hard to understand. The standard languages of Sweden and Norway (Oslo and Stockholm, plotted on top of each other in Figure 2) are most similar to standard Danish and also most easily understood, even more so than the Danish dialects of Hjordkær, Katrad and Høgsted.

The strong correlations between the linguistic distances and the intelligibility scores show that the intelligibility of closely related languages and language varieties can be predicted well from the linguistic distances between the target language and the language of the listener (the larger the distances, the more difficult it is to understand the varieties). The correlation between the lexical and phonetic distances is rather low ($r = .49$, $p = .05$), which shows that the two levels are to a large degree independent. Still, a stepwise regression analysis excludes lexical distances from the model. A regression analysis including both lexical and phonetic distances results in a non-significantly higher prediction of intelligibility ($r = .90$, $p < .0001$) than phonetic distances only ($r = .86$, $p < .0001$), see Table 2.

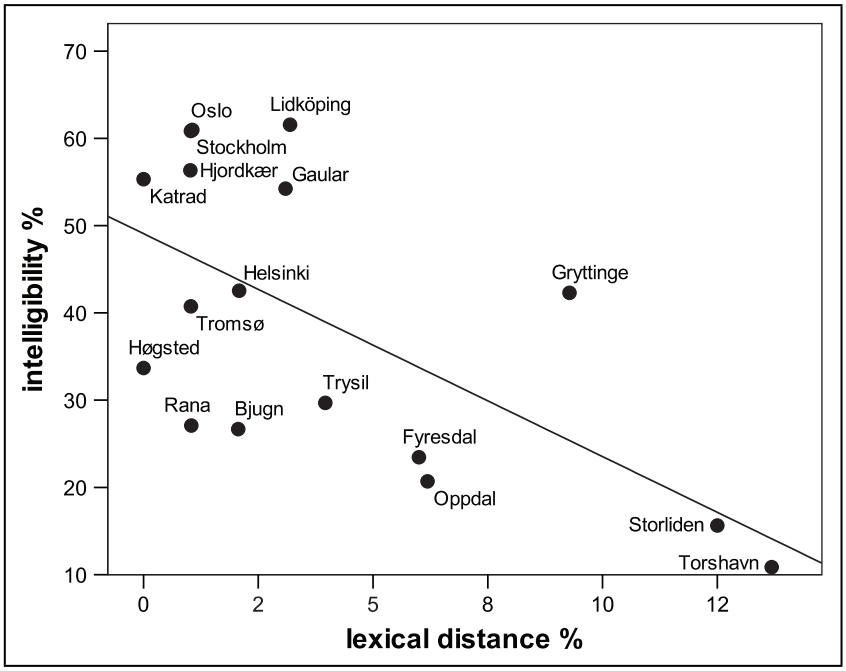

Figure 2. Scatterplot of the correlation between lexical distances and intelligibility scores ($r = -.64$, $p < .05$).

5.2 Consonants and vowels

We calculated the phonetic distances for vowels and consonants separately (see Section 4.1) and correlated these two distance measurements with the intelligibility scores. We expected consonants to play a more important role in intelligibility than vowels (see Section 1). The correlations for the consonants were indeed significantly stronger than for the vowels ($r = -.74$ versus $r = -.29$, $p = 0.04$), see Table 1. The correlations are lower than in the case of the full phonetic distances ($r = -.86$) but the consonant distances are still significant at the one percent level while the vowel distances are not. In Figure 4 a scatterplot is presented with the correlation of intelligibility with both the vowels and the consonants. Only consonant distances correlate significantly with the full phonetic distances ($r = .88$). The correlation between consonant distances and vowel distances is low and not significant ($r = -.21$, $p = .41$), which shows that the two levels are independent of each other. A regression analysis with vowels and consonants as independent variables and intelligibility as the dependent variable results in a correlation of .76. In a stepwise regression analysis, only consonants are included in the first step ($r = .74$).

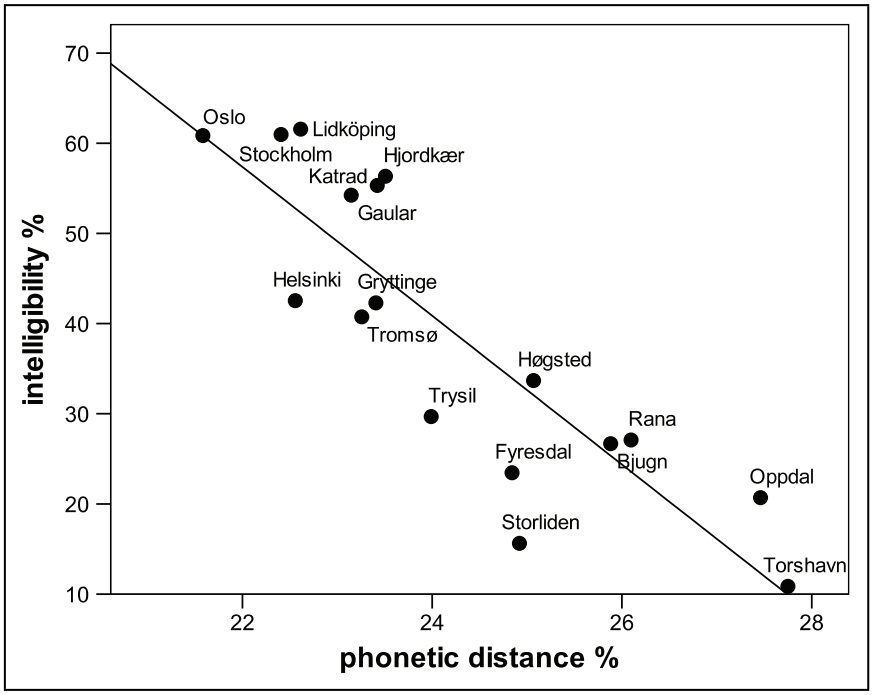

Figure 3. Scatterplot of the correlation between phonetic distances and intelligibility scores ($r = -.86$, $p < .01$).

Table 2. Regression analyses where intelligibility scores are the dependent variable and phonetic and lexical distances are the independent variables.

Method	Entered variables	Model	Result
Enter	Full phonetic distance, lexical distance	Full phonetic distance, lexical distance	$R = .90$ $R^2 = .81$ $p < .001$
Stepwise	Full phonetic distance, lexical distance	Full phonetic distance	$R = .86$ $R^2 = .74$ $p < .001$

5.3 Phonetic sublevels

Finally, we made separate analyses of the consonants and vowels for the different kinds of consonant and vowel operations (insertions, deletions, substitutions, lengthenings and shortenings). Only consonant substitutions ($r = -.57$, $p < .05$) and vowel shortenings ($r = -.49$, $p < .05$) correlate significantly with the intelligibility scores. A regression analysis with all phonetic sublevels gives

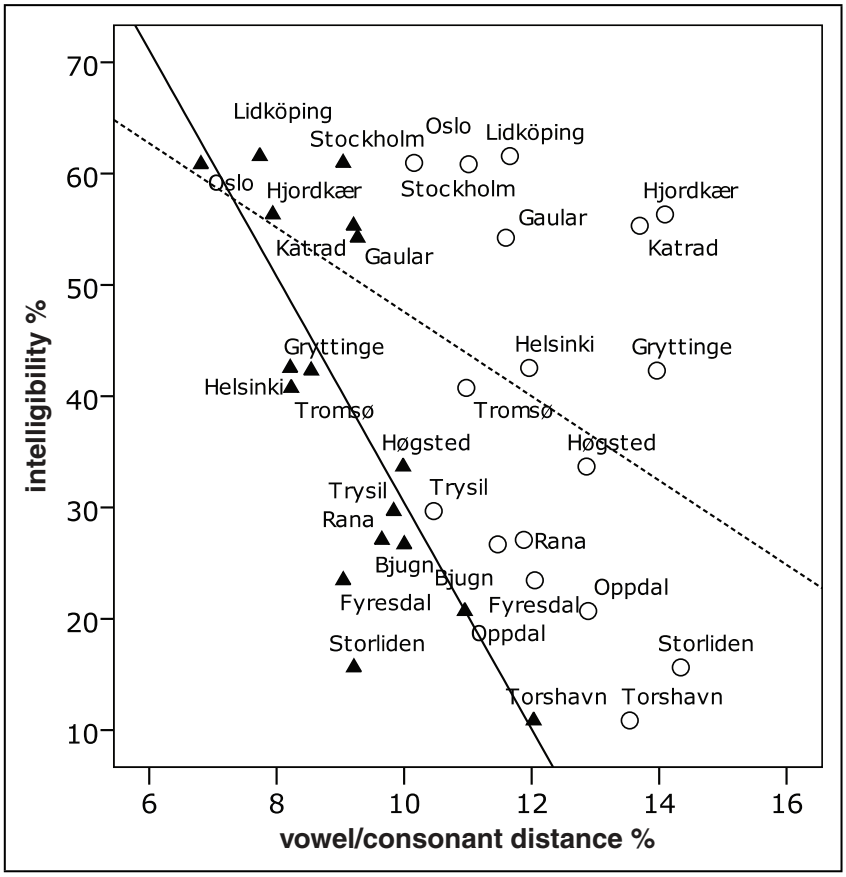

Figure 4. Scatterplot of the correlation between intelligibility and vowel distances (circles, $r = -.29$, $p = .25$) and between intelligibility and consonant distances (triangles, $r = -.74$, $p = .001$).

Table 3. Regression analyses where intelligibility scores are the dependent variable and consonant and vowel distances are the independent variables.

Method	Entered variables	Model	Result
Enter	Consonants, vowels	Consonants, vowels	$R = .76$ $p < .001$
Stepwise 1st step	Consonants, vowels	Consonants	$r = .74$ $p < .001$

Table 4. Regression analyses where intelligibility scores are the dependent variable and different phonetic sublevels are the independent variables.

Method	Entered variables	Model	Result
Enter	All phonetic sublevels	All phonetic sublevels	$R = .93$ $p < .001$
Stepwise 1st step	All phonetic sublevels	Consonant substitutions	$R = .57$ $p = .017$
Stepwise 2nd step	All phonetic sublevels	Consonant substitutions, vowel insertions	$R = .78$ $p < .001$
Stepwise 3rd step	All phonetic sublevels	Consonant substitutions, vowel insertions, vowel shortenings	$R = .87$ $p < .001$

a multiple correlation of $R = .93$ ($p = .05$). A stepwise regression analysis results in three models. As could be expected from the correlations, consonant substitutions are the most important factor. Even though vowel insertion does not correlate significantly with intelligibility ($r = -.39$, $p > .05$) it is included as the second factor and this results in a correlation of .78 ($p = .001$). Vowel shortening is included in the third model and results in a correlation of .87 ($p < .0001$).

6. DISCUSSION AND CONCLUSIONS

We investigated the role of different linguistic levels in the intelligibility of 17 Scandinavian language varieties among young Danes from Copenhagen. First, we wanted to investigate the relative contribution of lexical and phonetic distances (research question 1 in Section 1). We found significant correlations with both lexical and phonetic distances, but a multiple regression analysis included only phonetic distances, which suggests that phonetic distance is a significant predictor. We found the tendency that lexical distances correlate less strongly with intelligibility scores than phonetic distances, but the difference was not significant. The lower correlation with lexical distances is probably due to the fact that it is difficult to predict the effect of individual lexical differences. One single non-cognate word in a sentence or even a larger part of a text can lower intelligibility considerably if the non-cognate word is a central concept. On the other hand, if the non-cognate words in a text have little semantic content, it has less influence on intelligibility. Furthermore, it is possible that the listeners understand some non-cognate words because they are familiar with the words from previous experience with the test language. Finally, the lexical differences between the varieties in our investigation are rather small (see Appendix A).

Next, we investigated the relative contribution of consonant and vowel distances to intelligibility (research question 2). We found that there is a significantly stronger correlation between consonant distances and intelligibility than between vowel distances and intelligibility ($r = -.74$ versus $r = -.29$, $p = 0.04$). This suggests that consonants play a more important role in intelligibility than vowels and it confirms the assumption by Ashby and Maidment (2005) that vowels carry less information than consonant, see introduction. We measured the distances between the individual sound on the basis of spectograms. As mentioned in Section 3 distances between vowels are generally smaller than between consonants. However, our results show larger mean distances between vowels than between consonants (see Figure 3 and Appendices A and B), which means that there is much more vowel variation than consonant variation. The fact that consonants still show a higher correlation with intelligibility confirms the important role that consonants play for the intelligibility. Since our investigation is based on the overall intelligibility of a whole speech sample we are not able to make conclusions about the importance of the quantity and quality of the vowel and consonant differences. In future work, more controlled experiments would have to be set up in order to investigate the relative importance of these two factors in the intelligibility of different languages.

Finally, we took a closer look at the role of the different consonant and vowel operations (research question 3). A stepwise regression analysis showed that consonant substitutions, vowel insertions and vowel shortenings contribute significantly to the prediction of intelligibility. As far as the consonants are concerned, substitution is the only operation that correlates significantly with intelligibility. The first model in a stepwise regression analysis includes this level and excludes all other levels. It does not come as a surprise that consonant substitutions play an important role in intelligibility. When consonants in a word are substituted the framework of the word is changed (see Section 1). Insertions and deletions of consonants are likely to alter the framework of a word as well because the number of reference points change. However both operations play an insignificant role in our investigation. This can probably be explained by the fact that many insertions are found at the end of words which may be less important for the identity of a word than the beginning. Furthermore, Danish has many reduced spoken word forms and as a consequence consonant insertions often result in a word form that is known to Danish listeners from the written or underlying form of the Danish equivalent. Lengthening and shortening of consonants will not disturb the Danish listeners since in Danish consonant length is not phonemically distinct.

In contrast with consonant substitutions, vowel substitutions play a negligible role in intelligibility. Considering the large amount of vowel variation in Danish this may not come as a surprise. Van Ooijen (1994:108) suggested that ‘listeners

are both aware of and equipped to meet [vowel] variability, especially in the case of a vowel-rich language.' The correlation between vowel insertions and intelligibility scores is low, but vowel insertions are still added to consonant substitutions in the second model resulting from the stepwise regression analysis. Vowel insertions may result in additional syllables or diphthongs in comparison to the corresponding words in Danish. This may conceal the identity of the words for Danish listeners. In the third model resulting from the multiple regression analysis, vowel shortening is added. This is also the only vowel operation that correlates significantly (at the 5 per cent level) with intelligibility. Vowel length is phonemically distinct in Danish and deviant vowel lengths may therefore be rather disturbing.

Our analysis is based on a limited number of words (a mean of 98 words). This sample may be too limited for a lexical analysis, while a small number of words has been proven to be sufficient for reliable measurements of phonetic distance by means of the Levenshtein distance (Heeringa 2004). Furthermore, the results found in this study may not generalise to other situations. If the number of non-cognates varies around the critical breakdown threshold, as it may well be when varieties of other languages (or languages families) are compared, lexical distance will be more important than phonological/phonetic distance within the cognates. Similarly, if a language has more (or fewer) consonants relative to vowels in its phoneme inventory, the importance of vowel and consonant distance will be different. Typically, the relationship between number and magnitude of deviations for the stored representation and intelligibility of a linguistic unit is non-linear. Identification of a sound or recognition of a word remains very good for small discrepancies from the norm, and then abruptly breaks down. Future work with different language varieties and more controlled representations of the various linguistic units can hopefully give more insight into the relative contribution of linguistic phenomena to intelligibility and show when the limits of intelligibility are reached.

APPENDIX A

Table A.1. Distances between standard Danish and 17 Scandinavian language varieties. The table shows lexical and phonetic distances, and (sublevel) consonant distances. Distances are expressed as percentages, see Sections 4.1 and 4.2. The range, mean and standard deviation are given in the last three rows.

Dialect	Lexical	Phonetic	Phon. Cons.	Phon. Cons. Subst.	Phon. Cons. Ins.	Phon. Cons. Del.	Phon. Cons. Leng.	Phon. Cons. Short.
Høgsted	.00	25.07	9.98	5.97	1.46	2.56	.02	1.17
Lidköping	3.19	22.61	7.74	3.90	2.33	1.51	.72	.21
Oslo	1.03	21.58	6.81	4.41	1.39	1.01	1.65	.33
Torshavn	13.68	27.75	12.03	5.00	4.76	1.76	.32	.41
Fyresdal	6.00	24.84	9.05	4.14	2.83	2.07	1.28	.19
Tromsø	1.03	23.16	8.23	5.29	.83	2.01	1.19	.60
Oppdal	6.19	27.46	10.95	6.30	2.06	2.26	.79	.79
Gaular	3.09	23.15	9.27	4.99	2.31	1.80	1.12	.12
Stockholm	1.06	22.41	9.04	3.90	3.89	1.08	1.03	.13
Hjordkær	1.02	23.51	7.94	4.24	1.87	1.83	.11	.38
Gryttinge	9.28	23.41	8.54	3.82	2.46	1.75	.36	.07
Bjugn	2.06	25.88	10.00	7.13	1.74	.97	1.37	.76
Helsinki	2.08	22.56	8.21	3.91	2.71	1.60	.57	.20
Katrad	.00	23.42	9.21	4.57	2.29	2.18	.06	1.12
Rana	1.04	26.09	9.65	6.98	1.22	1.19	1.21	.85
Storliden	12.50	24.92	9.21	5.25	1.68	1.81	.29	.20
Trysil	3.96	23.99	9.83	5.81	2.46	1.30	1.04	.57
Range	13.68	6.17	5.22	3.31	3.93	1.59	1.63	1.10
Mean	3.95	24.22	9.16	5.04	2.25	1.69	.78	.48
Std. dev.	4.24	1.78	1.24	1.08	.96	.46	.51	.35

Table A.2. Distances between standard Danish and 17 Scandinavian language varieties. The table shows (sublevel) vowel distances, expressed as percentages (see Section 4.1) and intelligibility scores representing the percentage of correctly translated words (see Section 3). The range, mean and standard deviation are given in the last three rows.

Dialect	Phon. Vow.	Phon. Vow. Subst.	Phon. Vow. Ins.	Phon. Vow. Del.	Phon. Vow. Leng.	Phon. Vow. Short.	Schwa versus Sonor.	Intelligibility
Høgsted	12.86	10.51	.79	.96	1.10	.31	.21	33.68
Lidköping	11.66	10.21	1.30	.14	1.47	.36	.46	61.57
Oslo	11.01	9.21	.89	.84	1.28	.40	.18	60.85
Torshavn	13.54	9.05	4.16	.27	1.60	.35	.06	10.86
Fyresdal	12.05	9.51	2.11	.43	1.39	.36	.53	23.46
Tromsø	10.97	9.48	.70	.80	1.27	.65	.33	40.74
Oppdal	12.89	9.15	1.49	2.25	1.64	.62	.12	20.70
Gaular	11.59	9.36	1.52	.71	.71	.32	.17	54.24
Stockholm	10.16	8.42	1.21	.52	1.08	.32	.81	60.97
Hjordkær	14.09	11.46	1.37	1.00	.50	.47	.29	56.34
Gryttinge	13.96	10.57	1.85	.61	.82	.44	.66	42.30
Bjugn	11.47	8.86	1.07	1.54	1.59	.57	.28	26.68
Helsinki	11.96	10.24	1.57	.15	.64	.50	.48	42.54
Katrad	13.70	10.34	1.54	.68	.26	.31	.06	55.32
Rana	11.88	8.84	1.18	1.86	1.98	.57	.20	27.10
Storliden	14.34	11.59	1.20	1.48	.54	.71	.15	15.63
Trysil	10.46	8.30	.84	1.33	1.41	.56	.37	29.68
Range	4.18	3.29	3.46	2.10	1.72	.40	.75	50.71
Mean	12.27	9.71	1.46	.92	1.13	.46	.32	38.98
Std. dev.	1.31	.98	.79	.60	.48	.13	.21	16.99

REFERENCES

M. Ashby and J. Maidment (2005), *Introducing phonetic science* (Cambridge).

K. Beijering, C. Gooskens, and W. Heeringa (2008), 'Predicting intelligibility and perceived linguistic distance by means of the Levenshtein algorithm' in M. van Koppen and B. Botma, eds, *Linguistics in the Netherlands 2008* (Amsterdam), 13–24.

L. O. Delsing and K. Lundin Åkesson (2005), *Håller språket ihop Norden?: en forskningsrapport om ungdomars förståelse av danska, svenska og norska* (Copenhagen).

C. Gooskens (2007), 'The contribution of linguistic factors to the intelligibility of closely related languages', *Journal of Multilingual and multicultural development*, 28(6), 445–467.

C. Gooskens and W. Heeringa (2004), 'The position of Frisian in the Germanic language area'. In D. Gilberts, M. Schreuder, and N. Knevel, eds, *On the Boundaries of Phonology and Phonetics* (Groningen), 61–87.

C. Gooskens and W. Heeringa (2006), 'The relative contribution of pronunciation, lexical and prosodic differences to the perceived distances between Norwegian dialects', *Literary and Linguistic Computing, special issue on Progress in Dialectometry: Toward Explanation* 21 (4), 477–492.

N. Grønnum (2003), 'Why are the Danes so hard to understand?' In H. Galberg Jacobsen et al., eds, *Take Danish – for instance: linguistic studies in honour of Hans Basbøll presented on the occasion of his 60th birthday, 12 July 2003* (Odense), 119–130.

W. Heeringa (2004), 'Measuring dialect pronunciation differences using Levenshtein distance' (Ph.D. thesis, University of Groningen).

B. van Ooijen (1994), 'The processing of vowels and consonants' (Ph.D. thesis, University of Leiden).

J. Séguy (1973), 'La dialectométrie dans l'atlas linguistique de la Gascogne', *Revue de linguistique romane*, 37, 1–24.

P. J. Scharpff and V. J. van Heuven (1988), 'Effects of pause insertion on the intelligibility of low quality speech', in W. A. Ainsworth and J. N. Holmes, eds, *Proceedings of the 7th FASE/Speech-88 Symposium, The Institute of Acoustics* (Edinburgh), 261–268.

C. Tang and V. J. van Heuven (2007), 'Mutual intelligibility and similarity of Chinese dialects. Predicting judgments from objective measures' in B. Los and M. van Koppen, eds, *Linguistics in the Netherlands 2007* (Amsterdam), 223–234.

A. Torp (1998), *Nordiske språk i nordisk og germansk perspektiv* (Oslo).

END NOTES

[1] At first glance it may look strange to consider lengthenings and shortenings as separate classes rather than particular kinds of substitutions. In Section 4.1 we explain the way in which lengthenings and shortenings are treated and come back to this.

[2] *The North Wind and the Sun* is a well-known text in phonetic research. In *The principles of the International Phonetic Association* (1949) the text is transcribed in 51 different languages.

[3] The recordings were made by Jørn Almberg and Kristian Skarbø. They are available via http://www.ling.ntnu.no. We are grateful for their permission to use the material.

[4] In Norway there is no spoken standard language. The Oslo variety represented the standard variety in this investigation. This variety functions as a spoken standard to some extent, even though it does not have a very strong position compared with spoken standards in many other European countries.

[5] The phonetic transcriptions of the Norwegian varieties were made by Jørn Almberg. The rest of the transcriptions were made by Andreas Vikran and corrected by Jørn Almberg to ensure consistency.

[6] The listeners were asked if they were familiar with the fable. Listeners who reported that they knew the story were excluded from the investigation.

[7] It also makes the task easier, but still not so easy that a ceiling effect is reached as becomes clear from the results in Section 5.1.

[8] Alternatively, we could have adopted the stem morpheme as scoring unit, so that getting the compound right would have counted as two units (cf. Scharpff & Van Heuven 1988).

[9] [ʔ] is the symbol used for the Danish phoneme called *stød* in Danish.

[10] Lengthenings and shortenings are excluded since they do not affect the structure of words, but in future work we want to include them as well.

[11] Significance of correlations and significance of differences between correlations are calculated with the Mantel test.