Abstract: The most straightforward way to explain why Danes understand spoken Swedish relatively better than Swedes understand spoken Danish would be that spoken Danish is intrinsically a more difficult language to understand than spoken Swedish. We present circumstantial evidence suggesting that Danish is intrinsically poorly intelligible, and the report on a formal experiment in which we tested the intelligibility of Danish and Swedish materials spoken by three representative male speakers per language (isolated cognate and non-cognate words, words in semantically unpredictable sentences, words in spontaneous interaction in map tasks) presented in descending levels of noise to native listeners of Danish (N = 18) and Swedish (N = 24), respectively. The results show that Danish is as intelligible to Danish listeners as Swedish is to Swedish listeners. In a separate task, the same listeners recognized the same materials (presented without noise) in the neighboring language. The asymmetry that has traditionally been claimed was indeed found. Possible reasons for the asymmetry are discussed.
Abstract

The most straightforward way to explain why Danes understand spoken Swedish relatively better than Swedes understand spoken Danish would be that spoken Danish is intrinsically a more difficult language to understand than spoken Swedish. We present circumstantial evidence suggesting that Danish is intrinsically poorly intelligible, and the report on a formal experiment in which we tested the intelligibility of Danish and Swedish materials spoken by three representative male speakers per language (isolated cognate and non-cognate words, words in semantically unpredictable sentences, words in spontaneous interaction in maptasks) presented in descending levels of noise to native listeners of Danish (N = 18) and Swedish (N = 24), respectively. The results show that Danish is as intelligible to Danish listeners as Swedish is to Swedish listeners. In a separate task, the same listeners recognized the same materials (presented without noise) in the neighboring language. The asymmetry that has traditionally been claimed was indeed found. Possible reasons for the asymmetry are discussed.
Is spoken Danish intrinsically unintelligible?

Charlotte Gooskens¹, Vincent J. van Heuven²,
Renée van Bezooijen¹ & Jos J.A. Pacilly²

¹University of Groningen, The Netherlands
²Leiden University Centre for Linguistics, The Netherlands

Submitted to Speech Communication, special issue on:
Non-native Speech Perception in Adverse Conditions: Imperfect Knowledge,
Imperfect Signal.

Corresponding author: Vincent J. van Heuven
Address: Phonetics Laboratory, Leiden University Centre for Linguistics
PO Box 9515
2300 RA Leiden
The Netherlands
Phone: +31 527 2319
Fax: +31 527 2615
E-mail: v.j.j.p.van.heuven@hum.leidenuniv.nl
Abstract

The most straightforward way to explain why Danes understand spoken Swedish relatively better than Swedes understand spoken Danish would be that spoken Danish is intrinsically a more difficult language to understand than spoken Swedish. We present circumstantial evidence suggesting that Danish is intrinsically poorly intelligible, and the report on a formal experiment in which we tested the intelligibility of Danish and Swedish materials spoken by three representative male speakers per language (isolated cognate and non-cognate words, words in semantically unpredictable sentences, words in spontaneous interaction in maptasks) presented in descending levels of noise to native listeners of Danish (N = 18) and Swedish (N = 24), respectively. The results show that Danish is as intelligible to Danish listeners as Swedish is to Swedish listeners. In a separate task, the same listeners recognized the same materials (presented without noise) in the neighboring language. The asymmetry that has traditionally been claimed was indeed found. Possible reasons for the asymmetry are discussed.

Keywords

Mutual intelligibility; Danish; Swedish; babble noise; Semantically Unpredictable Sentences; Map tasks; cognates
1. Introduction

Mutual intelligibility between related languages varieties, i.e. dialects of the same language or sibling languages within the same language family, is not always reciprocal. For instance, Portuguese (Brazilian) listeners have significantly fewer problems understanding spoken Spanish than (Latin-American) Spanish listeners have understanding Portuguese (Jensen, 1989). Similarly, listeners whose native language is a Southern Chinese ‘dialect’ understand Mandarin (Northern Chinese) varieties much better than the other way around (Tang and van Heuven, 2009).

An especially intriguing case of non-reciprocal intelligibility is found in Northern Europe. Here Danish has the unenviable reputation of being virtually unintelligible to non-native listeners, whilst the other Scandinavian languages, i.e. Swedish and Norwegian, are readily understood by Danish listeners. On the anecdotal level we find comments and jokes such as the following:

Danish is an interesting language although it is commonly considered to be a very ‘flat’ language due to its pronunciation, which tends to be fairly monotonous. The ‘flatness’ of the language is a cause for many quips as the country of Denmark itself is also fairly flat, and people seem to find it amusing that Denmark’s language matches its geography.

One of the most commonly noted features of Danish is the difficulty of its pronunciation. Foreigners are typically unable to pronounce certain things in Danish and the Danes find it an endless source of amusement. One of the running jokes about the Danish language is that ‘Danish is not so much a language as a throat disease.’

Is Danish a beautiful language, or is it, as some foreigners claim, a disease of the throat? Sometime last century, according to the Danish linguist Otto Jespersen, a little Dutch boy was sitting practising his Danish. His father interrupted him with the words: ‘Don’t hiccup like that, boy, it’s not good for your throat’, to which the boy replied: ‘I’m not hiccupping, father, I’m speaking Danish’.

More recently, we have seen a spoof (the comedy programme Uti Vår Hage) on Norwegian television where the point of the sketch was that Danes find it impossible to understand each other to the extent that the international community is asked to come to the rescue of Denmark. The following is the (Norwegian-accented) English commentary with the video, which begins with a supposed Dane complaining:

I know, the Danish language has always been impossible to understand for most Scandinavians, but in recent years, it has become impossible to understand for us in Denmark too. So, for me, the Danish language has collapsed into meaningless guttural sounds.

---

2 http://nordicvoices.blogspot.com/2009_05_01_archive.html (last accessed 24 July 2009). We have not been able to trace the reference to Jespersen, which may well be apocryphal.
3 http://www.youtube.com/watch?v=s-mOy8VUEBkn (last accessed 26 July 2009)
Beyond the anecdotal, it has been recognized in both first and second-language acquisition studies that Spoken Danish is an extraordinarily challenging language. Bleses and Basbøll (2004), and more recently Bleses et al. (2008), have shown that the early language development of Danish children is slower than that of children with other mother tongues, such as English and Swedish. A comparison with 15 different languages revealed that Danish children score lowest on vocabulary comprehension (as reported by the parents) during the entire period they were followed, i.e. between 0;8 and 1;3 years. At the beginning of the period, the size of the passive vocabulary of the Swedish toddlers was about twice that of their Danish counterparts, at the end of the period the Swedish toddlers’ vocabulary was still 38 percent larger. Bleses et al. attribute this difference to the poor segmentability of Danish, which is caused by prosodic phenomena such as lack of specific juncture cues, absence of compulsory sentence accents and of local signals to utterance function. At the segmental level, lenition of consonants and other reduction phenomena, in particular schwa-assimilation and schwa deletion, tend to result in poor segmentability. In the context of second-language learning, Grønnum (2003) lists a number of characteristics of the phonology of Danish, which would make it difficult for foreign learners to understand. These are largely the same factors as those mentioned by Bleses et al. In addition, Grønnum mentions the large number of phonologically distinct vowels. Danish has more than forty vowel sounds (including long vowels with a stød, a kind of creaky voice) and a large number of diphthongs.

Various investigations on inter-Scandinavian intelligibility have shown that Danish is the most difficult neighbouring language in Scandinavia. For instance, Maurud (1976) reports that Danes are reasonably adept at understanding Swedish (a comprehension score of 43%), whilst Swedes do much worse when trying to understand Danish (23% on the listening comprehension test). Interestingly, reading comprehension scores were perfectly symmetrical (mean scores of 69%) in the same study by Maurud (1976). According to Teleman (1987: 76), changes in the Danish pronunciation might have made it more difficult for a Swede to ‘find the letters behind the sounds’ than vice versa. This assumption is based on the fact that mutual reading comprehension is symmetrical while oral comprehension is not. The same tendency towards asymmetrical speech intelligibility between Swedes and Danes was reported by Bø (1978), Börestam (1987), and by Delsing and Lundin Åkesson (2005). The latter authors also point out that immigrants of various origins have more difficulties understanding Danish than immigrants in Norway and Sweden have understanding Norwegian and Swedish, respectively.

The asymmetry in intelligibility between Swedish and Danish has often been explained as the result of a difference in attitudes. Danes would hold Sweden and its language, Swedish, in high esteem while Swedes tend to take a low view of Denmark and Danish. A second, related, account of the asymmetry rests on a difference in language contact. Denmark is a geographically small country so that most Danes live close to Sweden, which makes it possible to watch Swedish television and visit the country. The great majority of Swedes, however, live far away from the Danish
border and get little exposure to Danish. Nevertheless, correlations between intelligibility scores and attitudes, and between intelligibility and geographic remoteness, were found to be low; a direct relationship is difficult to prove (Gooskens, 2007). Moreover, if the asymmetry were the consequence of a difference in attitude, it is hard to see why mutual intelligibility is asymmetric for spoken but not for written language. For a further discussion of the role of extra-linguistic factors in the mutual intelligibility between Swedes and Danes see Schüppert and Gooskens (2009).

The mainstream view on speech communication holds that every language is an optimal solution to the problem of finding a workable compromise between speaker needs and hearer needs. On the one hand, the speaker should be allowed to talk quickly and with minimal effort. On the other hand, the speaker’s sounds must remain distinct enough to allow the listener to recognize words and reconstruct the message without having to ask the speaker to repeat himself. The evidence presented above seems to bear out that in Danish the speaker-hearer balance is unduly tipped in favor of the speaker. Indeed, the last century has seen a large number of changes (lenition processes) in the phonology of Danish, creating a ever larger gap between the orthography and the pronunciation. The letter-to-sound correspondences were rather straightforward around a century ago but the relationship has become much more opaque in the course of the twentieth century. This would also be the reason why there is no asymmetry in mutual understanding between written Danish and written Swedish and only non-reciprocity in mutual understanding of spoken language.

In the present paper, then, we assume that linguistic factors are likely to play a role. Specifically we will test the hypothesis that Danish is intrinsically more difficult to understand than Swedish. If Danish is indeed difficult for the Danes themselves, it will be difficult for non-natives as well. The idea that Danish would be difficult to understand even for native speakers has never been addressed experimentally. In the studies cited above, it was taken for granted that listeners would obtain perfect comprehension scores when exposed to speech samples of their own language variety. The first aim of our study was to compare the intelligibility of Danish for Danes with the intelligibility of Swedish for Swedes. An intelligibility test was developed both with sentences and isolated words read out in the mother tongue of the listeners. The test sentences and words were presented in a series of descending noise levels, i.e. from difficult to easy. If Danish is indeed an intrinsically more difficult language, the intelligibility threshold should be higher for Danish than for Swedish native listeners. A similar methodology was applied successfully earlier to examine possible subtle differences in intelligibility of native and Dutch-accented English and between native and English-accented Dutch (Van Wijngaarden, 2001).

Of course, we also needed to ascertain that the mutual intelligibility between Swedish and Danish is asymmetrical, as has traditionally been claimed (see above). Therefore, we also administered the test cross-linguistically, by having Swedes respond to the Danish version of the materials and Danes to the Swedish version. We
had established in pilot experiments that no added noise was required to obtain sufficient discrimination between native and non-native listeners. We point out here that our test is the first full-fledged functional test on the mutual intelligibility of Swedish and Danish. The test did not rely on opinion scores (judgments), nor on self reports, nor did it test overall comprehension (‘the gist’) of the message.

An additional aim was to investigate the role of prosodic phenomena and reduction phenomena across word boundaries. As mentioned above, according to Bleses et al. and Grønnum such phenomena would degrade the segmentability of Danish. Poor segmentability is claimed to compromise intelligibility. By comparing the intelligibility results for isolated words with those for sentences, conclusions can be drawn about the influence of these factors on intelligibility. If the isolated words are easier to understand (suffer less from noise) than read sentences in comparison to Swedish, this supports the hypothesis by Bleses and Grønnum that Danish is difficult to understand due to the low segmentability caused by prosodic phenomena and reduction phenomena across word boundaries.

2. Method

2.1 Stimulus materials

Three kinds of stimulus material were included in the test: read sentences, spontaneous sentences, and isolated cognate words. The use of read sentences allows us to test the same types of sentences with the same cognate words in both languages and for all speakers. Moreover, it is possible to control the input completely. The influence of semantic context can be minimalized and the words can be selected on the basis of characteristics such as frequency, lexical category and length.

The speaking style and the variation in spontaneous speech are quite different from constructed read sentences. It is uncertain to what degree this will influence intelligibility scores. On the one hand, spontaneous speech might be less carefully pronounced than read speech, which might result in more assimilation. On the other hand, the listeners might be helped by the contextual cues when interpreting the spontaneous sentences. Also, spontaneous speech is likely to be more varied as far as prosody is concerned.

In contrast with read sentences, reduction phenomena are less frequent in isolated words since they only occur within the words and not across word boundaries. Also, the prosodic features which are found at sentence level, are absent. In the next sections we will go into detail about the different kinds of stimulus material.

2.1.1 Semantically unpredictable sentences (SUS)

To construct sentences to be read aloud by the speakers, the so-called SUS-generator was used that was developed by Benoît et al. (1996). The SUS generator produces
Semantically Unpredictable Sentences, which can be used to measure overall intelligibility. The generator was originally developed for the evaluation of the intelligibility of text-to-speech systems at the sentence level, but it is also a useful method for testing the intelligibility of natural language. The sentences are syntactically correct but semantically anomalous. For example, in a semantically anomalous sentence such as *He drank the wall* the syntactic structure is correct. Listeners receive cues as to syntactic category only but other than that they will not be able to make any further predictions about word identity by means of semantic or syntactic contextual cues.

The SUS-sentences can be automatically generated using five basic syntactic structures and a number of lexicons containing the most frequently occurring short words in each language. The syntactic structures are simple and the sentence length does not exceed seven words (eight for English because of the auxiliary in questions) in order to avoid saturation of the listeners’ short-term memory. The following lexical categories are used to construct the sentences:

- nouns
- transitive verbs (trans. verb)
- intransitive verbs (intrans. verb)
- adjectives (adj)
- relative pronouns (rel pron)
- prepositions (prep)
- conjunctions (conj)
- question-words (quest)
- determiners (det).

These word classes are used to implement the following syntactic structures:

S1 Intransitive structure: det + noun + intrans. verb + prep + det + adj + noun
S2 Transitive structure: det + adj + noun + trans. verb + det + noun
S3 Imperative structure: trans. verb + det + noun + conj + det + noun
S4 Interrogative structure: quest + trans. verb + det + noun + det + adj + noun
S5 Relative structure: det + noun + trans. verb + det + noun + rel pron + intr. verb

All words are selected from the most frequent words in their syntactic category using published databases which list words in terms of their token frequency in written texts or spoken recordings. Only those words containing the smallest number of syllables within a word class are used. This means that most words are monosyllabic. All words are unambiguous in terms of their phonological shape and syntactic category, so that homophones are included and no words which can belong to more than one syntactic category. Furthermore for each syntactic category, there are special restrictions. For the Scandinavian languages the most important restrictions are the following:

- verbs: no auxiliaries and reflexives, only present tense is used (including the imperative in S3).
- nouns: only singular forms
adjectives: only forms which can be used attributively, no comparative and superlative forms
- prepositions: only single-word prepositions
- determiners: only indefinite forms

Swedish SUS sentences

The Swedish SUS-generator consists of words taken from a Swedish word frequency list based on 1,000,669 running words from five Swedish newspapers from 1965 (Allén, 1970). For each word category the following numbers of most frequent words from the frequency list were included:

- nouns 120
- adjectives 36
- transitive verbs 48
- intransitive verbs 24
- question words 3
- prepositions 11
- conjunctions 3
- relative pronouns 1
- determiners 2

All included words are monosyllabic, except for the present tense verbs, which are bisyllabic in 80% of the cases. Also half of the imperative verb forms in the word lists are bisyllabic.

Danish SUS sentences

No Danish SUS-generator was available, so we had to program one ourselves. For this purpose we used the frequency list that has been compiled by Bergenholtz (1992) between 1987 and 1990. This list differs from the Swedish list in that it is based on newspapers (25% of the words) as well as novels (50%) and magazines (25%) and on a larger number of words (4 million). For the sake of comparability with the Swedish materials, we used only the part of the list that is based on newspapers (750,000 words). The texts are more recent than the Swedish texts (a difference of 22 to 25 years). We do not expect these differences to be of great importance for our purpose. The words chosen for the SUS-generators can still be expected to be commonly used words in both languages (see comparisons of frequencies below). For the Danish SUS-generator we included the same number of words per word category as for the Swedish generator following the general principles for the SUS-generator as sketched above. The verbs in the present tense are bisyllabic in 80% of the cases. Furthermore, two question words are bisyllabic in Danish. Unlike Swedish, also the all imperative forms in the Danish SUS sentences are monosyllabic.
Generation of SUS sentences for the intelligibility experiment

In the intelligibility test twelve different SUS sentences were used. They consisted of three sentences from each of the four syntactic structures S2-S5 (see above). No sentence from S1 was included since this structure is rather similar to S2 and since we wanted to be able to distribute the twelve sentences equally over the structures. In order to counterbalance possible language-specific influences such as differences in word frequency, half of the twelve SUS sentences originated from the Swedish SUS-generator and the other half from the Danish SUS generator. The Swedish sentences were then translated into Danish and the Danish sentences into Swedish so that in total we had the same twelve sentences in Swedish and in Danish. The syntax of Swedish is so similar to the Danish syntax and the four syntactic structures are so simple and general that it provided no problems to translate the Swedish sentences word for word into Danish or visa versa.

When generating the sentences in the two languages, the following precautions were observed:

- Words with different genders in the two languages were excluded. Listeners may get confused if a word is preceded by a determiner with a ‘misleading’ gender, and this is an effect that we are not interested in in the present investigation.

- Words that occurred more than once were substituted by other words. Exceptions are *en* (indefinite article common gender), *et* (indefinite article neuter gender), *og* ‘and’ and *som* (relative pronoun).

- Words for which no cognate (i.e. historically related word) exists in the other language were excluded. By excluding non-cognates we make sure that all errors made by the listeners when listening to the neighboring language could be attributed to phonetic differences between the two languages. However, since cognates can have different frequencies in the two languages, they might not always be equally easy to understand. In order to make sure that word frequency played no role in the intelligibility results, we decided to use only words that were among the 5000 most frequent words in both languages (according to the frequency dictionaries from which the words were taken, see above). For each word in the two languages, the relative frequency of the word form as used in the sentences was looked up. Only in the case of the imperative, was the frequency of the present tense looked up because the imperative form itself had mostly a rather low frequency. The Danish words had a mean relative frequency of 0.19 (0.03 when including only content words) with values between 0.008 and 0.037 and the mean relative frequency of the Swedish words was 0.17 (0.012 with only content words) with values between 0.002 and 0.031. The correlation between the Swedish word frequencies and the Danish word frequencies is $r = 0.97$ (0.50 when including only content words). This means that the Swedish and the Danish words
are rather similar as far as frequency is concerned and therefore there is no reason
to expect word frequencies to influence the results.

When generating a sentence, first ten sentences with the relevant syntactic structure
(S2-S5) in the relevant language (Danish or Swedish) were generated. The first
sentence was chosen if it fulfilled the conditions specified above. If this was not the
case, the part of the sentence which did not satisfy the conditions was substituted by
the same part in the next sentence which fulfilled the condition. These steps were
repeated until all twelve sentences had been generated. The total number of words
(tokens) is 106 per language (73 types). As mentioned above, a number of words in
the Swedish and the Danish generator were bisyllabic. In the sentences which were
generated for the experiment, ten of the 15 Danish verbs and eleven of the Swedish
verbs were bisyllabic. Furthermore, two of the three Danish question words were
bisyllabic. The rest of the words were monosyllabic. (1) and (2) are examples of SUS
sentences used in the experiment.

(1) Structure S2:
En rød nat tager et navn (Danish)
En röd natt tar ett namn (Swedish)
‘A read night takes a name’

(2) Structure S3
Vis en sol og en bog (Danish)
Visa en sol och en bok (Swedish)
‘Show a sun and a book’

2.1.2 Spontaneous sentences

We used two kinds of spontaneous sentences. Half of the sentences were collected by
means of so-called map tasks (cf. Anderson et al., 1991; Brown et al., 1984. Grønnum
(2009) used this method to collect materials for a Danish Spontaneous Speech corpus
(DanPASS). Part of these materials was used for our experiment and we made
Swedish recordings in the same manner. The map tasks result in speech which is
syntactically varied but the variation in content words in the material is rather limited.
For this reason we also had the speakers describe pictures showing common words
and activities. This method has been developed by Van Bezooijen and Van den Berg
(1999). The other half of the sentences were taken from this material.

Map task

The map task involved the cooperation of two participants who knew each other well
in order to make them feel as comfortable as possible. They were seated in separate
locations, one the phonetic department’s recording studio at the University of
Copenhagen, the other a recording facility established for the purpose in the main
control room, with curtains of very heavy material surrounding the speaker. The speakers communicated via headsets. They were recorded through professional headset microphones (Voice Technologies VT700), directly onto CD-ROM (HHB Professional Compact Disc Recorder CDR-850) to separate channels in a stereo recording. In this way recordings of good quality were achieved.

Each participant had a map. One, the instruction giver, had a route on his or her map; the other, the instruction follower, did not. Their goal was to collaborate so as to reproduce the giver’s route on the follower’s map. The maps were not exactly identical: landmarks were missing on one or the other map, a landmark may appear twice – in two different locations – on one map but not on the other; and the same landmark may have slightly different names on the two maps. This gave rise to a true negotiation, with questions and answers, backtracks, etc. Participants were explicitly informed about these irregularities in written instructions prior to the recording. It was left to them, however, to discover how and where the maps or the designations differed, and to supply the missing items and correct names on their respective maps. Each pair of speakers completed four different sets of maps.

Picture descriptions

The pictures to be described by the speakers showed five different everyday situations including common objects in and around the house: a woman peeling potatoes in a kitchen, a baby in a living room, a couple and their baby sleeping in a bedroom, a shopping street with two women, and a garden with laundry on a clothes-line. The subjects were asked to describe the pictures as detailed as possible, including the position of people and objects with respect to each other.

Selection of sentences

From the recordings, six sentences or fragments were selected per speaker from the map task materials and another six from the picture description task. The selection of the sentences was based on the following criteria:

- sentences should contain six or seven words only
- sentences should constitute single prosodic entities
- sentence types should vary
- Swedish and Danish sentences should be as similar as possible in terms of vocabulary and sentence type
- Words should be in both the Danish and Swedish top-5000 frequency lists
- Words should be cognates in Danish and Swedish

In (3) and (4) we give examples of Danish and Swedish spontaneous sentences, respectively.

(3)  
*Så den hedder altså den grønne sø* (Danish)  
‘So that is called the green lake’
(4) Nedanför sängen sitter det en liten råtta (Swedish)
   ‘Under the bed there is a little rat’

2.1.3 Isolated cognate words

Ideally, the same words should be tested as were used in the SUS sentences. This
would provide the best basis for a comparison of the intelligibility results of these to
kinds of stimulus materials. However, it would result in a design which would
demand the testing of too many groups of subjects or in too long test sessions. For this
reason we took care to select the isolated words on the basis of the same criteria as the
words in the SUS sentences. This still makes it possible to compare the results. We
decided to include only nouns so that the subjects had a cue to syntactic category as in
the case of the SUS sentences.

The 24 words were selected in the same way as the words in the SUS-generator. This
means that the words met the following criteria:

- monosyllabic
- cognates
- among the 5000 most frequent words
- singular
- unambiguous in terms of their phonological shape and syntactic category

The mean relative frequencies of the words were almost the same in the two
languages (0.013 for Danish ranging between 0.002 and 0.042, and 0.010 for Swedish
ranging between 0.003 and 0.030). The correlation between the Swedish and Danish
word frequencies is high ($r = 0.89$). Examples of cognate words are given in (5).

<table>
<thead>
<tr>
<th>(5)</th>
<th>Danish</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>hund</td>
<td>[hun’]</td>
<td>hund</td>
<td>[hønd]</td>
</tr>
<tr>
<td>hav</td>
<td>[hav]</td>
<td>hav</td>
<td>[hav]</td>
</tr>
<tr>
<td>luft</td>
<td>[loft]</td>
<td>luft</td>
<td>[loft]</td>
</tr>
<tr>
<td>bror</td>
<td>[bro:p]</td>
<td>bror</td>
<td>[brɔr]</td>
</tr>
</tbody>
</table>

2.1.4 Non-cognates

Listeners with prior experience with the L2 would be able to understand the language
on the basis of knowledge and not on the basis of linguistic similarity only. In order to
quantify the amount of experience of the listeners with the L2, they were asked to
translate twelve non-cognate L2 words. Since non-cognates are per definition
unintelligible to listeners with no prior experience with the L2, the number of
correctly translated non-cognates is a good measure of experience. The non-cognates were presented to the L2 listeners only, together with the isolated cognate words.

The non-cognates were selected from top-5000 in the frequency lists that were also used for the SUS-tests (see Section 2.1.1). It is important that the words are frequent because otherwise the listeners may never have heard them even though they had experience with the L2. The mean relative frequency of both the Swedish and the Danish non-cognates is 0.008, ranging from 0.002 to 0.039 for Danish and from 0.002 to 0.037 for Swedish. Examples of non-cognates are given in (6).

\begin{table}[h]
\centering
\begin{tabular}{llll}
\hline
  & Swedish & English & Danish & English \\
\hline
fråga & ‘question’ & evne & ‘ability’ \\
känsla & ‘feeling’ & avis & ‘newspaper’ \\
hot & ‘threatening’ & skuffelse & ‘disappointment’ \\
pojke & ‘boy’ & værelse & ‘room’ \\
\hline
\end{tabular}
\end{table}

2.1.5 Speakers

The three youngest male speakers from the DanPASS corpus (see Section 1.1.2) were selected. They were born in either 1976 or 1977. They were (former) students from the Department of General and Applied Linguistics at the University of Copenhagen, and they hailed from the greater Copenhagen area. The spontaneous speech which had already been recorded of these speakers was used. The same speakers were asked to come back for a recording of the SUS sentences and the isolated words.

The Swedish speakers were matched as well as possible with the Danish speakers. They had the same age, originated from Stockholm and were students of linguistics at the University of Stockholm. New recordings of all speaking styles were made in the same way as for the Danish speakers.

2.1.6 Speech manipulations

All sentences and words were downsampled to 16 KHz. Increasing levels of babble noise were added to yield five versions. Babble noise consists of the mixed recordings of a large number of speakers. By adding babble noise to the recording we imitated a situation where the speaker is in the midst of a crowd of people. Babble noise is held to be the most effective and least obnoxious masker of speech. The noise was amplitude modulated in such a way that its intensity rose and fell proportionally to that of the speech wave. This was done to ensure that each sound was made unintelligible to the same extent. We added four descending noise levels to the original recording, in steps of 3 dB from 0 to \(-9\) dB speech-to-noise ratio (SNR) (i.e. from easy to difficult). This resulted in five different versions. Version 1 is the downsampled copy of the original recording (‘clean’), version 2 was mixed with 0 dB noise (noise has the same level as the original recording), version 3 with \(-3\) dB (i.e.
noise is 3 dB stronger than target speech), version 4 with –6 dB and version 5 with –9 dB.

2.2 Experiment

2.2.1 Design

A total of twelve spontaneous sentences, twelve SUS sentences, 24 isolated cognate words and twelve non-cognate words were tested. The SUS sentences and the isolated cognate words were tested in a crossed design so that all sentences or words by each of the six speakers were tested. The spontaneous sentences are all different and for this reason a crossed design is not necessary for this type. First the isolated words (cognates and non-cognates) were tested, next the SUS sentences, and finally the spontaneous sentences.

Each listener listened to two spontaneous sentences, two SUS sentences and four isolated words for each of the three L1 and L2 speakers (see overview in Table 1). Half of the listeners listened to the neighboring language first (part A of table) and the other half listened to their own language first (part B of table). In order to be able to test all sentences by all speakers in a crossed design, twelve groups of listeners had to be tested in total.

Table 1 Overview of the number of test items per listener. A: Listeners who begin with speakers of the neighboring language. B: listeners who begin with speakers of their own language.

<table>
<thead>
<tr>
<th></th>
<th>L2 speakers</th>
<th>L1 speakers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>Cognates</td>
<td>4</td>
<td>4 (× 5 SNR)</td>
<td>4 (× 5 SNR)</td>
</tr>
<tr>
<td>Non-cognates</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>SUS</td>
<td>2</td>
<td>2</td>
<td>2 (× 5 SNR)</td>
</tr>
<tr>
<td>Spontaneous</td>
<td>2</td>
<td>2</td>
<td>2 (× 5 SNR)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>L2 speakers</th>
<th>L1 speakers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>Cognates</td>
<td>4 (× 5 SNR)</td>
<td>4 (× 5 SNR)</td>
<td>4 (× 5 SNR)</td>
</tr>
<tr>
<td>Non-cognates</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>SUS</td>
<td>2 (× 5 SNR)</td>
<td>2 (× 5 SNR)</td>
<td>2 (× 5 SNR)</td>
</tr>
<tr>
<td>Spontaneous</td>
<td>2 (× 5 SNR)</td>
<td>2 (× 5 SNR)</td>
<td>2 (× 5 SNR)</td>
</tr>
</tbody>
</table>

2.2.2 Listeners

Eighteen Danish and 24 Swedish listeners were tested. The Danish subjects were university students from Roskilde University, all living at the Island of Sealand with a
mean age of 24.2 years ranging from 20 to 27 years. These listeners spoke the same standard variety of Danish as the speakers although they hailed from parts of Sealand that were as remote from the Swedish border as was possible. The Swedish students were from Stockholm University with a mean age of 22.6 years ranging from 19 to 27 years. These listeners spoke the same variety of standard Swedish as the Swedish speakers did. None of the listeners reported hearing problems.

In addition to translation of non-cognates (see Section 2.1.4) the listeners were asked what percentage of words from the neighboring language they expected to be able to understand. This was expressed on a 7-point scale as follows: ‘1’ 0-5%, ‘2’ 6-10%, ‘3’ 11-20%, ‘4’ 21-50%, ‘5’ 51-80%, ‘6’ 81-90% and ‘7’ 91-100%. This is a simpler way of asking about previous experience with the neighboring language than by means of more direct questions about the amount of contact such as ‘How often do you watch Swedish television?’.

2.2.3  Test

The test was carried out on the computer. All subjects listened to the stimuli via headphones, individually or in small groups. While the isolated words were presented only once, the read and spontaneous sentences were presented twice, with a 1-s pause in between the first presentation and the repetition, in order to exclude effects of memory limitations.

3.  Results

We will first present the results of the intrinsic intelligibility tests, i.e. when speakers and listeners share the same native language. Here we test the hypothesis that Swedish listeners will understand the Swedish speakers under more adverse speech to noise ratios while Danish listeners can only understand the Danish speakers when the SNR is more benign. In the second part of the analysis we will deal with the cross-linguistic part of the experiment, where Danish listeners respond to Swedish speakers and Swedish listeners decode Danish speakers. Here we expect to replicate the usual asymmetry between Danish and Swedish, even if the expected difference in intrinsic intelligibility should not be obtained. In the third part of the data analysis we will run post-hoc checks on the concurrent validity of the test components by examining the correlation between the word and sentence intelligibility parts. We will also consider the question whether listeners who do well on the intrinsic intelligibility tasks tend to be also skilled listeners in the neighbouring language. Part of the post-hoc checks will be devoted to possible confounds due to differences in familiarity with the neighboring language between our Danish and Swedish listeners.
3.1 Part 1: Intrinsic intelligibility

Figures 1A-B-C present percentages of correctly recognized words in each of the three parts of the experiment, i.e. for cognate words, words in SUS sentences and words in spontaneous sentences, respectively. In the panels the scores are broken down by Speech-to-Noise Ratio (SNR) plotted from left to right from adverse to benign, with separate lines for Danish and Swedish speakers/listeners.

The results were analysed by a repeated-measures Analysis of Variance (RM-ANOVA) with SNR and test component as within-listener factors, and language of speaker/listener as a between-subjects factor. Within each test component, the data were averaged over speakers (three speakers per language) and over items (24 for cognate words, and twelve items for the two sentence components each).

The results of the ANOVA are summarized in table 2. The ANOVA shows a main effect of test component. This effect is caused by the lower overall mean scores on the cognates. Since the effect is found in both the Danish and the Swedish half of the data, the component × language interaction is insignificant.

All three panels of figure 1 show the same effect of SNR: the percentage of words correctly recognized increases monotonically from about 50 for the −9-dB SNR to roughly 90 for the most benign SNR (‘clean’). The main effect of SNR is significant (see table 2). There is no interaction between SNR and language. There is, however, a small interaction between SNR and test component. No other interactions reach significance. Crucially, the results also show that there is no overall difference in performance level between the Danish and the Swedish speaker/listener combinations. If anything, the Danish listeners withstand a more adverse SNR than the Swedes but the difference fails to reach significance.

Table 2. Effects and interactions in within-language intelligibility test.

<table>
<thead>
<tr>
<th>Effect/interaction</th>
<th>df1</th>
<th>df2</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language (L)</td>
<td>1</td>
<td>40</td>
<td>&lt;1</td>
<td>ins</td>
</tr>
<tr>
<td>SNR (S)</td>
<td>4</td>
<td>160</td>
<td>151.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Test component (T)</td>
<td>2</td>
<td>80</td>
<td>14.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>L×S</td>
<td>4</td>
<td>160</td>
<td>1.2</td>
<td>ins</td>
</tr>
<tr>
<td>L×T</td>
<td>2</td>
<td>80</td>
<td>&lt;1</td>
<td>ins</td>
</tr>
<tr>
<td>S×T</td>
<td>4</td>
<td>320</td>
<td>3.8</td>
<td>=.002</td>
</tr>
<tr>
<td>L×S×T</td>
<td>8</td>
<td>320</td>
<td>&lt;1</td>
<td>ins</td>
</tr>
</tbody>
</table>
Figure 1. Intrinsic intelligibility of Danish and Swedish. Percentage of correctly recognized target words as a function of speech-to-noise ratio and broken down by speaker/listener group. The test battery comprised three components: Cognate words (panel A), words in SUS sentences (panel B) and words in spontaneous sentences (panel C). The error bars are +/- 1 standard error of the mean.
Part 2: Cross-language intelligibility

Now that Danish is not intrinsically less intelligible than Swedish, it is important that we replicate the usual finding that Danes understand Swedish better than Swedes understand Danish. For this part of the data analysis we only use the versions of the materials that were presented in quiet (‘clean’). Figure 2 shows the results. It plots percent correctly recognized words separately for the three test components (cognate words in panel A, words in SUS sentences in panel B and words in spontaneous sentences in panel C). In each panel we present the means for four combinations of speaker and hearer native language.

The data were again analysed by an RM-ANOVA, with stimulus language (Danish, Swedish), test component (cognates, SUS, Spontaneous) as within-listener factors, and native language of the listener as a between-subjects factor. As before, the data were accumulated over speakers and items. The results are summarized in table 3.

Table 3. Effects and interactions in cross-language intelligibility test.

<table>
<thead>
<tr>
<th>Effect/interaction</th>
<th>df1</th>
<th>df2</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listener type (L)</td>
<td>1</td>
<td>45</td>
<td>170.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Stimulus language (S)</td>
<td>1</td>
<td>45</td>
<td>23.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Test component (T)</td>
<td>2</td>
<td>45</td>
<td>1.7</td>
<td>ins</td>
</tr>
<tr>
<td>L×S</td>
<td>2</td>
<td>45</td>
<td>42.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>L×T</td>
<td>2</td>
<td>45</td>
<td>6.2</td>
<td>= .004</td>
</tr>
<tr>
<td>S×T</td>
<td>2</td>
<td>45</td>
<td>&lt;1</td>
<td>ins</td>
</tr>
<tr>
<td>L×S×T</td>
<td>2</td>
<td>45</td>
<td>&lt;1</td>
<td>ins</td>
</tr>
</tbody>
</table>

The results show that the three test components are roughly equally difficult so that there is no effect of test component. Overall, the Danish materials are more difficult to understand than the Swedish counterparts, which yields a significant effect of stimulus language. As we have seen in the previous section, however, the two languages are equally difficult for native listeners (recognition scores around 90%); therefore the effect of stimulus language is entirely due to the asymmetry in the cross-language conditions. Indeed, Danish listeners obtain word recognition scores with the Swedish materials at around 70% correct, while Swedish listeners have much poorer recognition scores with the Danish materials, i.e. between 40 and 50% correct. As a result, the stimulus language × listener language interaction is significant. Moreover, we find this configuration of results for each of the three test components, so that there is no third-order interaction.
Figure 2. Intelligibility scores (percent correctly recognized words) for Swedish and Danish materials as perceived by L1 listeners (decoding their own language) and by L2 listeners (decoding the neighboring language). Further see figure 1.
Finally, there is a small interaction between listener type and test component. In the SUS sentence part of the test the discrepancy between listening to the own language (almost perfect scores) and to the neighboring language (between 40 and 60% correct word recognition) is larger than in the other two test components. Apparently, the SUS test polarizes relative to the other tests.

3.3 Part 3: Supplementary analyses

In total 18 Danish and 24 Swedish subjects participated in the experiment as listeners. We intended to make sure that listeners would have little or no experience with the neighboring language so we would be able to attribute any asymmetry in mutual intelligibility between Danish and Swedish to intrinsic linguistic factors rather than to familiarity with the target language. One precaution we took was to sample our listeners from an area that we as remote from the Danish-Swedish border as possible. In the case of the Danish listeners, however, this was not really possible as the listeners had speak the same language variety that our Danish speakers used, i.e. Standard Danish. Even though we recruited our Danish listener from the Western part of the isle of Sealand, in absolute terms this part is still rather close to the border, so that familiarity with Swedish cannot be excluded. Therefore we computed the percentage of correctly translated non-cognates in the neighbor language for the Danish and Swedish listeners, as an index of their familiarity with the neighbor language. Ideally, listeners should not be able to translate even a single non-cognate word from the neighboring language. We also asked our listeners, before running the listening test, to rate their understanding of the neighboring language on a scale from 1 (no comprehension at all) to 7 (near-perfect comprehension). Ideally, our Danish and Swedish listeners should be perfectly matched on this measure.

Figure 3 plots the percentage of correctly translated non-cognates as a function of the self-estimation of comprehension in the neighbor language. Swedish and Danish listeners are represented by different symbols. The assumption of no familiarity with the neighbor language is largely met by the Swedish listeners. Twenty-two out of 24 Swedish listeners do not recognize a single Danish non-cognate, one Swede recognizes two words and one even recognizes three non-cognates, even though all non-cognates were chosen to be highly frequent words in the language. This functional measure is in stark contrast with the self-reported ability on the parts of the Swedes to understand spoken Danish; five Swedish listeners believe they would understand more than 50% (but no more than 80%) of spoken Danish. As a result, there is no correlation between the functional measure and the self estimation, $r = 0.145$ (N = 24, ins.). The assumptions are not met in the case of our Danish listeners. The Danes always recognize at least 25% of the Swedish non-cognates, and five Danes get more than 50% of non-cognates. This functional performance is matched by their self-estimated performance, as is also shown by the significant correlation
between the two measures, \( r = 0.564 \) (\( N = 18 \), \( p = 0.015 \), two-tailed). The Danish and Swedish listener groups therefore differ significantly in terms of both variables capturing experience with the neighbor language, \( t(40) = 3.4 \) (\( p = 0.002 \), two-tailed) for self-estimated comprehension and \( t(40) = 10.4 \) (\( p < 0.001 \), two-tailed) for percentage of correctly translated non-cognates. Moreover, since the Danes and Swedes are in virtually non-overlapping parts of the score range, we have no possibility to trim the data and run a re-analysis on the part of the range where the two listener groups are comparable.

Figure 3. Percentage of non-cognates correctly translated as a function self-estimated comprehension of neighbor language for 18 Danish (dark squares) and 24 Swedish (light circles) listeners. Markers may represent multiple listeners. Overall \( r = 0.569 \) (\( N = 42 \), \( p < .001 \)), for Swedish listeners \( r = 0.149 \) (\( N = 24 \), ins.), for Danish listeners \( r = 0.564 \) (\( N = 18 \), \( p = 0.015 \)).
Table 4 presents a correlation matrix that contains all relevant dependent variables measured in the experiment. The coefficients have been computed separately for the 24 Swedish (upper half of matrix) and the 18 Danish (lower half of matrix) listeners. Table 4 shows that the amount of listeners’ experience of the Danish listeners with the neighbor language (i.e. Swedish) correlates positively with their performance on the recognition tests, both overall and on the three components of the battery. The components themselves are only weakly intercorrelated, so that, also in hindsight, it makes sense to include separate tests for each of the components.

Van Heuven (2008) suggested that understanding speech in a closely related language does not require the acquisition of new special skills. His hypothesis was that the listeners marshal up the decoding strategies that they have learnt to use when having to understand speech in their native language under adverse (noise) circumstances. Following up on this hypothesis we would now like to know if listeners who are adept at recognizing words in their own, native language under adverse signal conditions, are naturally better equipped to also understand words in a neighboring language, i.e. in a language variety that resembles the native L1 to some extent but is far from identical to it. If a correlation between the two skills exists, it would indicate that the population can be divided into naturally more and less gifted decoders.

Table 4 contains the information that allows us to check the viability of the naturally gifted decoder hypothesis. Since the group of Danish listeners is potentially contaminated by previous exposure to the neighbor language, we test the hypothesis on the Swedish listeners only, who by and large have had no useful experience with Danish. The correlation between the performance of the Swedish listeners on their native Swedish materials in the most adverse listening condition (i.e. with an SNR of −9 dB) and their performance on the Danish materials is significant only for the component that tests word recognition in spontaneous sentences, \( r = 0.441 \) (\( N = 24, p = 0.016, \text{ one-tailed} \)). The correlation is lower or even essentially 0 and insignificant for the other two components, and for the mean of the three components of the battery taken together. In sum, our results do not lend convincing support for the naturally gifted decoder hypothesis.

<table>
<thead>
<tr>
<th></th>
<th>Estimated understand.</th>
<th>L2 cognates correct</th>
<th>L2 non-cognates correct</th>
<th>Own language</th>
<th>Neighbour language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Estimated understanding</td>
<td>.569</td>
<td>.149</td>
<td>–.374</td>
<td>.151</td>
<td>.560</td>
</tr>
<tr>
<td>2. L2 cognates correct</td>
<td>.332</td>
<td>.358</td>
<td>.032</td>
<td>.172</td>
<td>.495</td>
</tr>
<tr>
<td>3. L2 non-cognates correct</td>
<td>.564</td>
<td>.470</td>
<td>–.045</td>
<td>–.110</td>
<td>.156</td>
</tr>
<tr>
<td>4. Cognates own</td>
<td>–.160</td>
<td>.494</td>
<td>–.059</td>
<td>.524</td>
<td>.442</td>
</tr>
<tr>
<td>5. SUS own</td>
<td>–.294</td>
<td>.243</td>
<td>.028</td>
<td>.610</td>
<td>.391</td>
</tr>
<tr>
<td>7. Cognates neighbour</td>
<td>.313</td>
<td>.989</td>
<td>.474</td>
<td>.450</td>
<td>.224</td>
</tr>
<tr>
<td>8. SUS neighbour</td>
<td>.034</td>
<td>.283</td>
<td>.392</td>
<td>.074</td>
<td>.353</td>
</tr>
<tr>
<td>9. Spont. neighbour</td>
<td>.179</td>
<td>.638</td>
<td>.582</td>
<td>.096</td>
<td>.430</td>
</tr>
<tr>
<td>10. Mean own</td>
<td>–.260</td>
<td>.395</td>
<td>.080</td>
<td>.849</td>
<td>.861</td>
</tr>
<tr>
<td>11. Mean neighbour</td>
<td>.225</td>
<td>.815</td>
<td>.599</td>
<td>.280</td>
<td>.410</td>
</tr>
</tbody>
</table>

Coefficients in bold are significant at p < .01.
4. Conclusions and discussion

In this study we aimed to find an explanation for the often observed asymmetry in mutual intelligibility between Swedish and Danish. The simplest possible explanation for the fact that Danes understand Swedish more readily than Swedes understand Danish would be that spoken Danish is intrinsically difficult. And indeed, there is abundant circumstantial evidence supporting the view that spoken Danish is more difficult than other Scandinavian languages. We also expected to find that Danish listeners would have greater problems decoding Danish speech in sentences rather than in isolated words, since more assimilation and reduction phenomena are reported across word boundaries than word internally. The results of our intelligibility test, however, show that Danish is as easy (or as difficult) to understand for Danish native listeners as is Swedish for Swedish native listeners. Moreover, we did not find a larger discrepancy in performance between recognition at the word and sentence level (whether in read-out semantically unpredictable sentences or in sentences taken from spontaneous interaction in map task).

At the same time, our study replicated the asymmetry between spoken Danish and Swedish. Indeed, our Danish listeners were clearly better at understanding the Swedish version of the materials than vice versa. In spite of our precautions, however, we cannot rule out the possibility than the asymmetry can be at least partly accounted for by a difference in experience with the neighbor language. It seems impossible, in retrospect, to find adult listeners of Standard Danish who do not have some familiarity with Swedish. The only way to rule out experience with the neighbor language as a confounding factor in the experiment, would be to recruit our listeners from the Western part of the country, e.g. in Århus or Esbjerg, where contact with Swedish and Swedish-spoken media is still minimal. This would be a conservative test of our hypothesis since listeners from the Western part of Denmark may have (minor) problems understanding Standard (Copenhagen) Danish and because their variety of Danish has a greater linguistic distance from Standard Swedish than Standard Danish has– but we still predict that they will understand Swedish relatively better than Swedes will understand Standard Danish.

A factor that should be mentioned as an explanation for the assumed difficulty of Danish is the fact that the Danish pronunciation has undergone an exceptionally fast development during the last century (Brink and Lund, 1975; Grønnum, 1998, see also the discussion of Teleman, 1987 in section 1). This may explain the fact that in Denmark especially elderly people complain having difficulties understanding people from the younger generation and it could also be part of the explanation for the asymmetry between Danish and Swedish. In order to test this hypothesis we will repeat our investigation with listeners between 60 and 70 years. If the elderly Danish subjects have more difficulties understanding the young speakers from their own countries than elderly Swedish subjects we may conclude that there is a larger generation gap in Denmark than in Sweden.
References


