Origin of asymmetry
Mutual intelligibility of spoken Danish and Swedish
5.1.2. Previous Research of Tempo and Reduction in Speech

5.2. Method

5.2.1. Material and speakers

5.2.2. Measurements

5.3. Results

5.3.1 Phonological Syllables

5.3.2 Phonetic Syllables

5.3.3 Reduction

5.4. Discussion

5.4.1 Reduction in Danish, Norwegian and Swedish

5.4.2 Articulation Speed in Danish, Norwegian and Swedish

5.4.3 Comparison of our findings with previous studies

5.4.4 Implications of this research

5.5. Conclusion

6 The role of L1 orthography for spoken word recognition of a closely related L2

Abstract

6.1. Introduction

6.1.1. Mutual intelligibility between Danish and Swedish

6.1.2. On-line activation of orthography during spoken word recognition

6.2. Experimental Procedure

6.2.1. Participants

6.2.2. Material

6.2.3. Procedure

6.2.4. EEG Recordings

6.2.5. EEG data reduction and analyses

6.3. Results

6.3.1. Behavioural data

6.3.2. EEG data

6.4. Discussion

7 Conclusion

8 Semantically Unpredictable Sentences, Norwegian version

9 Semantically Unpredictable Sentences, Swedish version

10 Semantically Unpredictable Sentences, Danish version

11 The role of L1 orthography for spoken word recognition of a closely related L2

12 Abstract

13 6.1. Introduction

14 6.1.1. Mutual intelligibility between Danish and Swedish

15 6.1.2. On-line activation of orthography during spoken word recognition

16 6.2. Experimental Procedure

17 6.2.1. Participants

18 6.2.2. Material

19 6.2.3. Procedure

20 6.2.4. EEG Recordings

21 6.2.5. EEG data reduction and analyses

22 6.3. Results

23 6.3.1. Behavioural data

24 6.3.2. EEG data

25 6.4. Discussion

26 7 Conclusion

27
1 Introduction

Swedish nurse: hon ringde te maj i morges, hon kunde inte finna sina dörrnycklar när hon skulle gå ti skolan
Danish patient: hendes: ø:h?
Nurse: dörrnöjle
Patient: nå nøglen til døren ja
Nurse: nycklarna
Nurse: jahh

Ridell (2008:129), my translation in italics

Native speakers of Danish and Swedish can generally communicate with each other using their native languages. The two languages are closely related and generally mutually intelligible. However, as anecdotally illustrated in the conversation above, which was recorded at a Danish nursing home, this theoretical possibility is not always entirely successful. Given that Danish and Swedish have been regarded as separate languages for several hundred years, this is not entirely surprising. What is striking about the success and failure of Danish-Swedish oral communication, however, is the fact that the mutual intelligibility between spoken Danish and Swedish asymmetrical. Danes tend to encounter fewer problems when they hear a Swedish person speak Swedish, than vice versa. This is a surprising tendency, as from a purely phonetic point of view, one could state that native (L1) speakers of Danish encounter the same sort of problems when they are confronted with spoken Swedish as the other way around, because the phonetic distance between the items in the two languages must be symmetrical. This thesis investigates which factors cause the asymmetry in mutual intelligibility of spoken Danish and Swedish and thereby hopes to shed some light on which factors influence mutual intelligibility between closely related languages in particular and spoken language recognition in general.

The Nordic countries Denmark, Finland, Iceland, Norway and Sweden as well as their associated territories Åland, the Faroe Islands and Greenland share history events as well as contemporary cultural and political norms. Their political and economic systems are characterised by generous welfare criteria and among other things emphasise gender equality and egalitarian benefit levels. The Nordic countries have co-operated officially in the Nordic Council since 1952 and in the Nordic Council of Ministers since 1972. Both authorities have strongly promoted inter-Nordic collaboration, e.g. by implementing the Nordic passport union in 1954 which allows Nordic citizens to travel and reside in any of the Nordic countries without a valid passport, by maintaining an inter-Nordic job exchange platform (Nordjobb, founded...
in 1985) and by emphasising the importance of using Nordic languages in inter-Nordic communication situations rather than a *lingua franca* such as English. Among other incentives, this was secured by the *Språkkonvention* (‘language convention’) that ensured that citizens of the Nordic countries are entitled to use their native language in written communication with authorities. Also, pupils in Finland, Iceland, Greenland, and the Faroe Islands learn at least one of the mainland Scandinavian languages in school, namely Swedish in Finland and Danish in the three remaining territories.

Particularly within mainland Scandinavia, i.e. the countries Denmark, Norway and Sweden, communicating across linguistic borders using the language of the speaker is a habit strongly encouraged by the authorities. Danes, Norwegians and Swedes are likely to use their native language rather than a lingua franca when speaking to each other. This manner of communication has been called *receptive bilingualism* by Hockett (1958) and *semicommunication* by Haugen (1966).

The first researcher to investigate mutual intelligibility in mainland Scandinavia was Haugen (1953). He elicited data on inter-Scandinavian communication patterns by asking Danish, Norwegian and Swedish members of *Föreningen Norden*1 (‘The Nordic Society’) how much of the neighbouring language they thought they could understand. By this, he elicited self-reported intelligibility abilities. Haugen (1953) reports promising intelligibility scores for most of the six communication situations (Danish in Norway and Sweden, Norwegian in Denmark and Sweden, and Swedish in Denmark and Norway), as intelligibility is above 80% for four of the language pairs (see Figure 1). However, communication between Danes and Swedes seemed to be problematic. More specifically, the Danish participants in his study report understanding only 56% spoken Swedish, while the Swedish participants reported 54% spoken Danish. Haugen’s 1953 publication was written in Norwegian, but 13 years later, he published his findings in English as Haugen (1966).

**Figure 1.** Self-reported intelligibility scores reported by Haugen (1966).

Haugen’s (1953; 1966) study was pioneering work and documented communication patterns (or rather, self-reports hereover) in Scandinavia of the early 1950’s. In the 1970’s, several short articles were published dealing with linguistic influence between

---

1 *Föreningen Norden* is a non-governmental organisation which promotes cooperation between the five Nordic countries Denmark, Finland, Iceland, Norway and Sweden and their associated territories Åland, the Faroe Islands and Greenland.
the mainland Scandinavian languages, such as Gösta Bergman’s *Svenska lån från danskan, norskan och finskan* (‘Swedish loan words from Danish, Norwegian and Finnish’, 1971), Allan Karker’s *Om svensk og norsk indflydelse på moderne dansk* (‘Swedish and Norwegian influences on contemporary Danish, 1971) and Poul Lindegård Hjorth’s *Nabosprogene i den højere danske skole* (‘Neighbouring languages in Danish high school, 1972). However, while the 1970’s experienced a great scientific and popular interest in this topic, Denmark seemed to take a big step away from the Scandinavian community politically in 1973, when the country was the first of the Nordic countries to join the European Economic Community (EEC), the precursor of the European Union (EU).

After Haugen’s (1966) publication, it took a decade before the topic of mutual intelligibility was picked up in an experimental investigation. This was done by Maurud (1976). In contrast to Haugen (1966), who based his study on the participants’ self-reported comprehension abilities, Maurud (1976) conducted an experimental investigation to assess exactly how much of their neighbouring languages Danes, Norwegians and Swedes could understand. He did so by presenting the participants from the three capital cities Copenhagen, Oslo and Stockholm with the neighbouring languages in a translation task. The highest intelligibility scores for spoken language reported by Maurud (1976) were achieved by Norwegians confronted with Swedish (median = 78.8 from 90 points, i.e. 87.5%), while the lowest scores were obtained by Swedish listeners confronted with Danish (median = 20.5 from 90 points, i.e. 22.7%). The intelligibility scores as elicited by Maurud (1976) are given in Figure 2.

**Figure 2.** Intelligibility scores reported by Maurud (1976) transformed into percent.

![Intelligibility scores reported by Maurud (1976)](image)

Importantly, not only the manner of data elicitation differed from Haugen, but so did his results, particularly those on Danish-Swedish communication. According to Maurud (1976), Danes and Swedes still seemed to encounter the biggest problems when communicating with each other. However, while the Danish-speaking participants in his study could understand about 43% of spoken Swedish, the Swedish-speaking participants only understood 23% of spoken Danish. That means, while Haugen (1966) reported that Danes and Swedes (thought that they) could understand a similar amount of their neighbouring language, Maurud (1976) reported that Danes actually comprehend more spoken Swedish than vice versa.
Interestingly, the intelligibility asymmetries reported for written texts are much less pronounced than those for spoken texts.

One of the major criticisms of Maurud’s (1976) investigation (Gregersen 2004), however, has been the fact that he compared the intelligibility of Swedish among Danes in Copenhagen to intelligibility of Danish among Swedes in Stockholm. As Copenhagen is located only 30 kilometres from the Swedish border, while Stockholm is located about 570 kilometres from the Danish border, there is a substantial geographical asymmetry in the data (see Figure 3). This geographical asymmetry is likely to be linked to an asymmetry in patterns of contact with the neighbouring language, such as through travelling to the neighbouring country, talking to people visiting their own country, or even watching television in the neighbouring language. While Copenhageners can access Swedish broadcasting programmes and can visit the neighbouring country rather easily, people living in Stockholm neither can watch nor listen to Danish broadcasting programmes, nor can they cross the border to Denmark within a couple of hours. The fact that Swedish participants had a lower intelligibility score than the Danish participants might partly be linked to the geographical asymmetry. However, Maurud’s (1976) conclusion was that “Swedes’ low understanding of the neighbour languages is a sign that the habit of hearing them and the attitude towards the need for understanding them are of major importance for the Scandinavians’ ability to communicate with each other in their respective languages” (Maurud 1976: 71), thereby suggesting that attitudes towards a specific language held by the listener are linked to the listener’s intelligibility of that language.

Figure 3. Map of Scandinavia.

Two years after Maurud’s study was published, a third investigation of mutual intelligibility of Danish, Swedish and Norwegian (Bø 1978) saw the light of the day. Interestingly, this study picked up the factor that was considered to distort Maurud’s (1976) data, namely the amount of contact with the neighbouring languages. In all three Scandinavian countries, the subjects in Bø’s (1978) study were chosen to form
two groups, one living inside and one living outside the border regions. Indeed, the group of subjects living in the border regions not only had more opportunities to visit the neighbouring country, but also had access to television programmes in the neighbouring language. Overall intelligibility scores per group of L1 speakers as reported by Bø (1978) are given in Figure 4.

**Figure 4.** Intelligibility scores reported by Bø (1978).

Bø (1978) found that subjects living near the border had fewer difficulties decoding the neighbouring variety than subjects living outside the border region, thereby indicating that a high degree of contact enhances intelligibility abilities, and, at the same time, confirming that Maurud’s (1976) data has to be interpreted with caution.

After publication of Bø’s (1978) study, the interest in mutual intelligibility between the Scandinavian languages seemed to decrease. Not much research was conducted in this field in the 1980’s. On a political level, after a 22-year delay, Sweden followed Denmark into the European Union in 1995. One year before Sweden became an EU-member, Börestam (1994) published a monograph entitled *Skandinaver samtaler* ('Scandinavians communicate') and picked up the topic of inter-Nordic communication again. Her work focuses on communicative strategies such as linguistic accommodation by the speaker, i.e. repetitions, clarifications, confirmations or paraphrases used towards the listener. She analysed ten videotaped discourses between speakers of Danish, Norwegian and Swedish. In an earlier publication (Börestam 1991), she elicited self-reported intelligibility of the three Scandinavian languages in Danes, Norwegians and Swedes. The intelligibility results are given in Figure 5.

**Figure 5.** Intelligibility scores reported by Börestam (1991).

Börestam (1991) reported a Danish-Swedish asymmetry similar to the one found in the experimental studies by Maurud (1976) and Bø (1978). One of the main findings
from Börestam (1994) on accommodation towards the listeners was that Swedish speakers and listeners seemed to create the majority of misunderstandings in communication situations between two or three speakers of different languages.

In 2001, an investigation by Lundin & Zola Christensen was published, which investigated mutual intelligibility of written texts by Danish and Swedish high school students. Again, it was reported that Danish participants comprehended more Swedish (namely 79.4%) than vice versa (69.6%), this time in a translation task of a newspaper article which was used in its original Danish form, and a version that had been translated to Swedish. However, this asymmetry is not as pronounced as has been typically reported for spoken language recognition.

At the same time, a large-scale project was initiated outside of Scandinavia, namely at the University of Hamburg in Germany. In the context of the Sonderforschungsbereich Mehrsprachigkeit (‘Collaborative Research Centre Multilingualism’), funded by the German Research Foundation (DFG), a group of researchers started to investigate mutual intelligibility in Scandinavia. The focus in this group has been on discursive strategies such as accommodation of the speakers towards each other (Braunmüller 2002, Zeevaert 2004, Golinski 2007), rather than quantifying success and failure of semi-communication.

Isochronically with the ending of the Hamburg project in 2005, a large-scale investigation of inter-Nordic communication patterns and abilities was published by Delsing and Lundin Åkesson (2005). To avoid repetition of the shortcomings of Haugen’s (1966) and Maurud’s (1976) studies, such as diverging attitudes or contact patterns across the groups of subjects, they elicited language attitudes held towards the fellow Nordic languages as well as contact patterns among different groups of participants from the Nordic countries, along with spoken and written text comprehension. The groups of participants hailed from at least two different sites per country, except for Finland, Åland, Greenland and the Faroe Islands. Danes were tested in Århus (340 km from Sweden via land route and 170 km via sea route) and Copenhagen, while Swedes were tested in Malmö (40 km from Danish mainland) and Stockholm. Thereby, the geographic asymmetry was somewhat neutralised, although Stockholm is still roughly two to three times as far from Denmark as Århus is from Sweden. Figure 6 gives intelligibility scores for the three Scandinavian languages.

Figure 6. Intelligibility scores reported by Delsing & Lundin Åkesson (2005).
Confirming Maurud’s (1976), Bø’s (1978), and Börestam’s (1991) findings, Delsing & Lundin Åkesson (2005) reported an asymmetry in mutual intelligibility between Danish and Swedish, where Danes have fewer difficulties in decoding Swedish than vice versa. In line with Gregersen’s (2004) comment, they found that Danes from Copenhagen have more contact with Swedish than Swedes from Stockholm have with Danish. However, Swedes living on the other side of the Öresund in Malmö (i.e. very closely to Denmark) have even more contact with Danish. It has to be born in mind, however, that the Öresund Bridge was opened in 2000, closely connecting Malmö and Copenhagen. That means that the contact indices, particularly those for Malmö and Copenhagen inhabitants, are likely to be higher in Delsing and Lundin Åkesson’s (2005) study compared to Maurud’s (1976) study due to the enhanced access to the neighbouring country. Delsing and Lundin Åkesson (2005) report a significant link between intelligibility and the amount of contact. However, two limitations to this result have to be pointed out. Firstly, they do not report correlation coefficients, but significance values only. The effect that the amount of contact has on intelligibility can therefore not be derived from their publication. Secondly, although they collected data from four sites which were shown to differ substantially with regard to the amount of contact to the neighbouring country (Århus and Copenhagen in Denmark and Malmö and Stockholm in Sweden), Delsing and Lundin Åkesson (2005) correlated intelligibility and contact in two subgroups (Danish participants and Swedish participants) only, thereby merging a high-contact and a low-contact group into the same analysis. It is possible that there is no causality in the correlation between contact and intelligibility they report, as there could be more factors involved. Evidence that attitude and intelligibility are not directly linked to each other but may be highly intercorrelated comes from Hilton and Gooskens (forthcoming). In their investigation of Danish intelligibility in Norwegian pupils from the northernmost province Finnmark (2000 km from Denmark) and the southernmost province Buskerud (300 km from Denmark) they report that geographical distance was not significantly correlated with intelligibility, but instead, with language attitude. Participants who lived closer to Denmark had been to Denmark more often than participants who lived further away from Denmark. Importantly, however, they also turned out to hold more positive attitudes towards Danish. Even if Delsing and Lundin Åkesson’s data (2005) report significant correlation coefficients, it can therefore not be concluded that there exists a causal relationship between intelligibility and contact.

A second focus of Delsing and Lundin Åkesson’s (2005) study was an empirical investigation of the proposed link between intelligibility and the attitudes held towards the neighbouring language and the neighbouring country. The participants were asked how much they liked the sound of the neighbouring language, and whether or not they were willing to move to the neighbouring country. They reported that Danes found the Swedish language more beautiful than vice versa, while Swedes were more positive towards moving to Denmark than Danes were towards moving to Sweden. They also report that Danes’ comprehension abilities correlate significantly
and positively with their judgment of how beautiful the Swedish language sounds and that Swedes' comprehension abilities correlate significantly and positively with their willingness to move to Denmark, while neither Danes' comprehension of Swedish and their willingness to move to Sweden correlated, nor Swedes' comprehension of Danish and their judgment of the beauty of the Danish language. Unfortunately, no correlation coefficients were reported in this subsection either, which makes it difficult to determine the amount of the variance explained by this factor. For the first time, however, empirical evidence was presented supporting the assumption that intelligibility and language attitudes are linked within the Scandinavian language area – although the nature of this link is still unclear. It is possible that listeners holding positive attitudes make a greater effort to understand the language in question than those holding negative attitudes, but it might also be the case that those participants who understand the language better, simply perceive the language as being more beautiful because their comprehension makes them feel as part of the speech community and facilitates a development of positive feelings towards a said variety. It is also possible that language attitude and intelligibility are not directly linked, but covary with the amount of contact, i.e. that listeners who live close to the border hold more positive attitudes towards their neighbouring country and therefore make a greater effort to understand its speakers. Intelligibility scores reported by Delsing and Lundin Åkesson (2005) are given in Figure 7. Again, it can clearly be seen that this study confirmed earlier findings that Danes understand more spoken Swedish than vice versa, while this asymmetry seem less clearly pronounced for written texts (Maurud 1976, Bø 1978).

Another large project on mutual intelligibility between the Scandinavian languages entitled *Linguistic determinants of mutual intelligibility in Scandinavia* was located at the University of Groningen from 2006 to 2011. Funded by the Netherlands Scientific Organization (NWO), the focus in this project has been on assessing how well speakers of Danish, Norwegian and Swedish understand each other and which linguistic factors influence their comprehension abilities (see below). Within this project, Gooskens (2006) reanalysed a subset of the data elicited by Delsing and Lundin Åkesson's (2005) which was chosen to keep the educational background of the participants similar across the language groups. In this subset, Gooskens (2006) could not confirm any significant correlation between contact and intelligibility. She reported only a limited correlation between language attitudes and intelligibility, so the question as to whether language attitudes and contact with the neighbouring language plays a role for the degree of mutual intelligibility still remains. Within the framework of the project, a series of experiments has been conducted. This thesis has been written as a part of this project.

To sum up, Norwegians tend to have the fewest problems understanding their fellow Scandinavian languages, while communication between Swedes and Danes is somewhat more difficult. In previous investigations, mutual intelligibility between spoken Danish and Swedish has been reported to be asymmetrical, in that Danes have fewer problems decoding spoken Swedish than Swedes have decoding spoken
Danish (see Figure 7). A number of factors have been suggested to cause the asymmetry in mutual intelligibility of spoken Danish and Swedish. Maurud (1976) suggested that the asymmetry in mutual intelligibility is due to extra-linguistic factors such as different attitudes towards the neighbouring language and/or an asymmetric amount of contact to the neighbouring language. Bø (1978) presented evidence in favour of the latter hypothesis and reported that access to broadcast programmes in the neighbouring language enhances intelligibility of that language, while Delsing & Lundin-Åkesson (2005) concluded that contact as well as language attitudes correlate with intelligibility. However, Gooskens (2006) could not confirm the link between intelligibility and the amount of contact.

Figure 7. Danish-Swedish intelligibility scores for reported by Maurud (1976), Bø (1978), and Delsing & Lundin Åkesson (2005).

Interestingly, the consistently reported asymmetry between Swedish and Danish-speaking listeners is not as strong in intelligibility of written texts (Maurud 1976, Bø 1978, Lundin & Zola Christensen 2001, Delsing & Lundin Åkesson 2005; see Figure 7). This seems to suggest that the asymmetry between the spoken forms is mainly caused by factors that are inherent in spoken language. Danish and Swedish differ in a number of linguistic features such as vowel space (Disner 1978, Vanhove et al. 2010) and some suprasegmental features such as stød and tone accents. These linguistic factors might also play a role in the asymmetry in mutual intelligibility. Some studies that have investigated the link between these factors and mutual intelligibility of Danish and Swedish are summarised below.

Gooskens & Kürschner (2010) investigated the role of different suprasegmental factors for the asymmetry in mutual intelligibility. In contrast to Danish, Swedish is a pitch-accent language and has two different tonal patterns, either of which is assigned to every word (accent 1 or accent 2). Gooskens & Kürschner (2010) report that Swedish listeners have more difficulties to decode Danish words which have accent 2 in Swedish, than those that have accent 1, while none of the Swedish accents has a detrimental effect on Danish listeners.
Furthermore, the phonemic vowel inventory of Standard Swedish generally consists of nine long and nine short distinctive vowels (e.g. Leinonen 2010), while the Danish phoneme inventory is mostly described as consisting of more contrastive vowels than that with 12 distinctive long vowels and 13 distinctive short vowels (e.g. in Basbøll, 2005:50). What is more, Vanhove et al. (2010) found that this asymmetry is not counterbalanced in realisation of the vowels by narrowing each Danish vowel space in comparison to Swedish vowel spaces, which generally leads to much more overlap between the specific vowels in Danish than in Swedish (see Figure 8).

**Figure 8.** Z-normalised vowel spaces of long and short vowels in Danish and Swedish (Vanhove et al. 2010). The vowels are labelled with XSAMPA characters (Wells 1995).

Conversion table XSAMPA – IPA.

<table>
<thead>
<tr>
<th>XSAMPA</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>E</td>
<td>æ</td>
</tr>
<tr>
<td>y</td>
<td>ø</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>2</td>
<td>ɔ</td>
</tr>
<tr>
<td>9</td>
<td>istringstream(9)</td>
</tr>
<tr>
<td>9_0</td>
<td>istringstream(9)</td>
</tr>
<tr>
<td>{</td>
<td>{</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>O</td>
<td>ɶ</td>
</tr>
<tr>
<td>Q</td>
<td>ɶ</td>
</tr>
<tr>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>1</td>
<td>istringstream(1)</td>
</tr>
<tr>
<td>8</td>
<td>istringstream(8)</td>
</tr>
</tbody>
</table>

Figure 8 shows z-normalised vowel spaces of long and short vowels in Danish and Swedish (Vanhove et al. 2010) based on formant measurements for every transcribed vowel type. It can be seen that the variance in realisation of the different vowels differed across vowel types, vowel lengths and language. For example, in Danish, the tokens for the vowel type transcribed with the IPA character /æ/ (XSAMPA character /E/) actually differed more in their formants than the tokens for the vowel /a/ (XSAMPA character /A/) did, i.e. the the realisation of /æ/ seems more variable than the realisation of /a/. Also, in Danish, the long vowel /ε:/ occupied a somewhat larger space than the short vowel /ε/ did, i.e. the realisation of /ε/ is more variable when the vowel is long. Finally, it can be seen that the vowel spaces occupied by Danish vowel types are roughly the same size as the ones for Swedish vowel types. As there are more vowel types in Danish than in Swedish, this leads to a significantly larger overlap of Danish vowels compared to Swedish vowels, which in turn makes it more difficult to recognise a vowel in Danish, than in Swedish.

In addition to differences in the vowel inventory, Danish has a supra-segmental feature that is not found in Swedish, namely the stød. Stød is generally described as a realisation of creaky voice or laryngealisation (Grønnum 1998: 179; Basbøll 2005: 83). There are monosyllabic and polysyllabic minimal pairs which differ only with regard to absence or presence of stød; however, in some Danish regiolects no stød is employed. Gooskens & Kürschner (2010) showed that the presence or absence of stød does not significantly impact intelligibility. This is the case for Danish subjects listening to Swedish, i.e. expecting the stød in certain words, as well as for Swedish subjects listening to Danish, i.e. confronted with an unfamiliar phenomenon.

On a phonological level, reduction processes such as schwa-assimilation and the vocalisation of consonants are well-documented phenomena in Danish (Basbøll, 2005; Grønnum, 1998; Grønnum, 2007). Doetjes (2010) pointed out that the word-final /d/ has been deleted in the Danish word land, which is pronounced /len'/ in contemporary speech, while it is still pronounced /land/ in Swedish. Deletion processes are still ongoing in contemporary Danish (Pharao 2010), where e.g. words such as helt /he:1d/ ‘completely’ are increasingly reduced to /he:1/. This is not the case in Swedish. Bleses et al. (2008) suggested that the high number of reduction and assimilation processes in Danish causes or boosts the delay in vocabulary development in Danish infants and children compared to that of their peers from ten European countries and from the US and Mexico. They point out that schwa-deletion and the vocalisation of consonants result in long vocalic stretches, making the Danish sound structure unclear with weak, or even no, cues for word and syllable boundaries. The number of reduction processes therefore has to be considered an important linguistic factor that differs across Danish and Swedish.

Gooskens & Doetjes (2009) suggested that literacy plays an important role for the asymmetry. Danish orthography is more conservative than Swedish orthography is, and generally reflects a pronunciation which is closer to its East Nordic root. As spoken Swedish has stayed closer to this root, it could be hypothesised that Danish listeners can use their orthographic system as an additional cue when they hear...
spoken Swedish. By calculating phoneme-grapheme consistencies for written Danish and spoken Swedish as well as written Swedish and spoken Danish, they confirm that Danes generally have more advantages from their native orthography than Swedes have. For example, it is likely that literate Danes confronted with the Swedish word /land/ can use their orthographic knowledge to match this word to their native correspondent land, while this is not the case for Swedish listeners confronted with Danish /leŋ/, as there is a phoneme missing which is present in Swedish pronunciation as well as orthography. However, it is not clear whether or not L1 orthography actually can be accessed during word recognition of a closely related L2, so the question remains whether Danish orthography serves as an additional cue during spoken language recognition of Swedish in Danish listeners.

Another linguistic factor linked to reduction is speaking rate. The faster an utterance is completed, the more segments have to be assimilated, lenited or deleted. Vice versa, the more reduction processes that are found in an utterance, the shorter the time must be to complete it. As previous research has shown that a higher speech rate impairs intelligibility (Vaughan & Letowski 1997, Gordon-Salant et al. 2007), the suggested higher amount of reduction in contemporary Danish might cause or increase the asymmetry in comprehension by impairing the intelligibility of spoken Danish for Swedish-speaking listeners.

This thesis focuses on the question why mutual intelligibility between Danish and Swedish is asymmetrical. As indicated in this chapter, previous research has suggested asymmetries in extra-linguistic factors such as languages attitudes and contact with the neighbouring language and country. These factors will be revisited, but linguistic factors that might play a role will also be investigated.

Chapter 2 investigates the general impact of extra-linguistic factors for the asymmetry in mutual intelligibility between Danish and Swedish. Subsequently, one of them, language attitudes is thoroughly investigated in chapters 3 and 4, and a second factor, namely literacy will be investigated in chapter 6. Chapter 5 focuses of the role of two linguistic factors that have hitherto not been investigated in the Danish-Swedish intelligibility context, namely the tempo of speech and the number of reduction processes in spoken language. In chapter 7, general conclusions will be drawn as to which factors cause or boost the asymmetry in mutual intelligibility between spoken Danish and Swedish.

References


Gregersen, F. 2004. Hvorfor kan vi københavnerne ikke forstå svensk når jyder kan forstå københavnsk? [Why can't we Copenhageners understand Swedish
while Jutlanders can understand Copenhagen Danish?] Sprogforum 29, 2004: 31-34.
Vanhove, J., T. Leinonen, V. J. van Heuven, C. Gooskens & A. Schüppert. 2010 Comparing the acoustic distinctiveness of Danish and Swedish vowels. Paper presented at Experimental Approaches to Perception and Production of Language Variation (ExAPP2010), Groningen.
Assessing the general role of extra-linguistic factors for mutual intelligibility between spoken Danish and Swedish

The final version of this paper will be published as Anja Schüppert & Charlotte Gooskens: "The role of extra-linguistic factors for receptive bilingualism: Evidence from Danish and Swedish preschoolers" in the International Journal of Bilingualism, Vol. 16, No. 3 (Sept. 2012) by SAGE Publications, Inc. All rights reserved.

Abstract

Danish and Swedish are closely related languages that are generally mutually intelligible. Previous research has shown, however, that Danes comprehend more spoken Swedish than vice versa. It has been suggested that this asymmetry is caused by extra-linguistic factors such as literacy, contact with, and attitudes held towards the test language. However, linguistic factors, such as supra-segmental features or differences in speech rate, could also cause or increase an asymmetry. The aim of the experiment reported in this paper was to exclude three extra-linguistic factors (attitude, contact, and literacy) in order to determine their role for mutual intelligibility. Participants were 19 Danish- and 26 Swedish-speaking illiterate preschoolers. Their task was to match 50 cognate nouns to corresponding pictures in a multiple choice task. Results revealed that word recognition scores in Danish children (63 percent) did not differ significantly from the Swedish scores (65 percent). That means that, in contrast to adult Danes, Danish children did not perform better on the word recognition task than their Swedish peers. This finding suggests that extra-linguistic factors play an important role for intelligibility of a closely related language, and that, as extra-linguistic factors develop, the intelligibility asymmetry emerges.

2.1. Introduction

Inter-Scandinavian communication, i.e. communication between speakers of Scandinavian languages, often takes place in the speakers’ native language rather than in English or another lingua franca. This custom is strongly encouraged by Nordic authorities and governments to support Scandinavian languages in a globalised world. The recommendations rely on the fact that mainland Scandinavian languages Danish, Norwegian and Swedish are closely related and have been proven to be mutually intelligible to a large extent in adults (Bø, 1978; Delsing & Lundin Åkesson, 2005; Maurud, 1976; Gooskens & Kürschner, 2010).

In previous investigations, mutual intelligibility of the Scandinavian languages was found to be asymmetrical. This asymmetry is especially large in Danish-Swedish communication: Danes have fewer problems decoding spoken Swedish than Swedes have decoding spoken Danish. This is illustrated in Figure 1, which shows word recognition scores reported by Maurud (1976), Bø (1978), Delsing & Lundin-Åkesson (2005), Gooskens & Kürschner (2010), and Lundin & Zola Christensen (2001). The first three studies report spoken as well as written intelligibility scores, whereas Gooskens & Kürschner (2010) only report spoken language intelligibility scores. The intelligibility scores for Swedish-speaking participants listening to spoken Danish
range from 50 to 70 percent, whereas the intelligibility scores for Danish-speaking participants listening to spoken Swedish range from 30 to 50 percent.

**Figure 1.** Intelligibility scores for spoken (left panel) and written items (right panel) reported by Maurud (1976), Bo (1978), Delsing & Lundin-Åkesson (2005), and Gooskens & Kürschner (2010). A number of factors have been suggested to cause the asymmetry in mutual intelligibility of spoken Danish and Swedish. Bo (1978), for example, investigated whether access to broadcast programmes in the neighbouring language enhances intelligibility of that language (see section 2.3.2.2). Delsing & Lundin-Åkesson (2005) elicited the attitudes towards the neighbouring country and language that were held by the participants of their study (see section 2.3.2.1). Gooskens & Kürschner (2010) investigated the role of different suprasegmental factors for the asymmetry in mutual intelligibility, (see section 2.3.1.1 and 2.3.1.2). Gooskens & Doetjes (2009) suggested that literacy plays an important role for the asymmetry (see section 2.3.2.3). These suggested factors can be divided into linguistic and extra-linguistic factors. Linguistic factors are language inherent features of the spoken forms of the languages, such as supra-segmental features, the average number of phonetic neighbours in the test language, the average word length in the target language, or language-specific speech rate. Danish and Swedish have common roots, but differ in many of these respects. However, in previous research, extra-linguistic factors such as the amount of contact with and attitudes held towards the neighbouring language have mostly been considered to cause the asymmetry in mutual intelligibility between two languages (see Zeevaert 2004 for a detailed overview).

Interestingly, the consistently reported asymmetry between Swedish and Danish-speaking listeners is not found in intelligibility of written texts. As Figure 1 illustrates, intelligibility scores for Danish-speaking participants confronted with written Swedish range from 40 to 63 percent (mean = 60 percent), whereas the scores range from 40 to 68 percent (mean = 55 percent) for the opposite direction. This suggests that the asymmetry between the spoken forms is mainly caused by factors that are inherent in spoken language, such as prosodic features or speech rate.

**2.2. Research question**

The aim of the experiment reported in this paper is to exclude the influence of extra-linguistic factors (attitude, contact, and literacy) as far as possible and thereby
investigate whether or not they play a role for the asymmetry in mutual intelligibility between Danish and Swedish-speaking participants. We hypothesise, here, that linguistic factors account for a large part of the asymmetry and therefore expect the asymmetry in mutual intelligibility to persist even when extra-linguistic factors are kept constant across the two groups of listeners (Danish and Swedish). In contrast to linguistic factors, which are usually based on the stimulus material, the extra-linguistic factors are found in the participants. Therefore, this study is conducted with a group of participants that is likely to be more neutral than participants in most other studies in these three respects: illiterate preschoolers from outside the border region. We assume that these participants have had a similar (low) amount of contact with the neighbouring language, cannot use their native orthography to decode the neighbouring language, and hold similar (neutral) attitudes towards it. These assumptions, however, will be tested individually (see section 2.4.1).

In the following section, we will first discuss several linguistic factors that have been found to be associated with intelligibility before turning to the discussion of extra-linguistic factors and their impact in intelligibility.

2.3. Linguistic and extra-linguistic factors

2.3.1. Linguistic factors

In this section, we introduce some linguistic factors that have been found to be associated with word recognition. Specifically, we discuss the role of the Swedish tonemes, the Danish *stød*, the number of phonetic neighbours for a specific word, and word length. We also suggest that a fifth factor, namely speaking rate, might partly account for the asymmetry in mutual intelligibility between Danes and Swedes.

2.3.1.1. Tonemes

Swedish is a pitch accent language, which means that every word has its specific prosodic contour. There are two forms of these word-specific prosodic contours, known as accent 1 and accent 2. The realization of the tonemes varies depending on regional variety. In standard Swedish, accent 1 is characterised by a single peak at the beginning of the first syllable, whereas the prosodic contour in disyllabic words with accent 2 peaks in the second syllable (see Figure 2). This means that monosyllabic words always carry accent 1. There are number of cases where the prosodic contour is the only distinction between segmentally similar, yet semantically completely different words (Elert, 1972: 163ff). These word pairs form minimal pairs that differ only with respect to their prosodic contour. Minimal pairs, however, that can be assumed to cause problems in actual communication are rare (Schüppert, 2003: 12), as many pairs consist of words belonging to different word classes, such as *rutten* /'rot/en/ ‘the route’ (definite noun) and *rutten* /'rot/en/ ‘rotten’ (past participle). In the vast majority of the cases, the linguistic context reveals the semantic content of the critical word, even if a wrong tonal contour is assigned to it. Furthermore, there are regiolects where no distinction between the tonal patterns is made (Meyer, 1954). Therefore, if an incorrect prosodic contour is assigned to a word, it will seldom cause
19

misunderstandings, but it might impair or delay word recognition. The two different pitch contours illustrated in Figure 2 show the pitch contour of two Swedish words with different tonemes produced by a native Swedish speaker of a Southern Swedish variety (Kronobergs län). The word bäbis ‘baby’ has accent 1 and the word äpple ‘apple’ has accent 2. It can be seen that the pitch contour in bäbis peaks at the beginning of the first syllable and falls continuously until the end of the word. In äpple, the first peak is at word onset, and in the middle of the second syllable, the pitch rises again and constitutes a second peak. It is commonly agreed that accent 1 is the unmarked accent (cf. Lahiri, Wetterlin & Jönsson-Steiner, 2005).

**Figure 2.** Pitch contours of the Swedish words äpple (‘apple’, solid line) and bäbis (‘baby’, dotted line).

In contrast to Swedish, Danish does not have any tonemes. All words stressed on the first syllable are essentially pronounced with the same prosodic contour. In standard Danish, stress is usually indicated by a lowering of the pitch, whereas in Jutland Danish, it is indicated by a higher pitch that falls on the next, unstressed syllable (Grønnum 1998:157, Lund 2001:79). In general, however, Danish is characterised by rather subtle movements in pitch contour. This can be seen in Figure 3, which displays the rather flat pitch contour of the Danish words æble ‘apple’ and baby ‘baby’ produced by a native speaker of Standard Danish from Fyn. In both words, the pitch is low at word onset and rises only subtly in the middle of the word before it falls again. In Figure 4, it can be seen that the Danish pitch contour of a disyllabic word stressed on the first syllable resembles neither the pitch contour of Swedish accent 1 nor the contour of accent 2.

**Figure 3.** Pitch contours of the Danish words æble (‘apple’, solid line) and baby (‘baby’, dotted line).

**Figure 4.** Prosodic contours of the Swedish words äpple ‘apple’ (solid line) and bäbis ‘baby’ (dotted line) and the Danish word baby ‘apple’, dashed line) from the stimulus material used in the experiment reported in the present paper.

Gooskens & Kürschner (2010) conducted a word recognition experiment with high-school students in Denmark and Sweden. The participants were auditorily presented
with 96 highly frequent words in a translation task. Participants living near the Danish-Swedish border were excluded from the analysis. A translation task designed to filter out Danish/Swedish participants that knew some Swedish/Danish showed that none of the remaining participants had substantial knowledge of neighbouring language. The authors found that Swedish intelligibility scores of Danish stimuli were significantly related to the accent the corresponding word had in Swedish. Danish cognates, i.e. words that share form and meaning across languages, of words that had accent 1 in Swedish were easier to decode for Swedish-speaking subjects than Danish cognates of words that had accent 2 in Swedish. The Danish-speaking participants, however, performed equally well in decoding stimuli from both categories. This suggests that Swedish subjects are confused by the missing tonemes in Danish to a larger degree than Danish listeners are by the unexpected presence of tonemes in Swedish.

2.3.1.2. Stød
The stød is a typical supra-segmental feature of the Danish language and is not found in Swedish. It is sometimes misinterpreted as a glottal stop (Lundskær-Nielsen, Barnes & Lindskog 2005; also, the phonetic symbol is /ʔ/), but generally described as a realisation of creaky voice or laryngealisation (Grønnum 1998: 179; Basbøll 2005: 83). Historically, the Swedish tonemes and the Danish stød are related. Generally, cognates to words that have accent 1 in Swedish have stød in Danish, and cognates to words that have accent 2 do not have stød. However, since all monosyllabic Swedish words have accent 1, this rule would imply that all Danish cognates of monosyllabic Swedish words, which are mostly monosyllabic in Danish as well, would have stød. This is not the case. Monosyllabic words can have stød, as in mand /mæn/ ‘man’, or not have stød, as in man /mæn/ ‘generic you’. Also, as Riad (2000: 18) points out, many Danish cognates to words that have accent 2 in Swedish lack stød because of low sonority. There are monosyllabic and polysyllabic minimal pairs which differ only with regard to absence or presence of stød. As is the case for Swedish tonemes, however, there are regions in Denmark where no stød is employed.

Bannert (1981) found that Swedish listeners interpret words with stød as having one syllable more than they actually have. He argues that, because listeners have difficulties recognising words with a different number of syllables from the corresponding word in their native language has, stød deteriorates intelligibility of spoken Danish for Swedish listeners. Gooskens & Kürschner (2010), however, reported that the presence or absence of the stød did not significantly impact intelligibility, neither for Danish subjects listening to Swedish expecting the stød in certain words, nor for Swedish subjects listening to Danish confronted with an unfamiliar phenomenon.

2.3.1.3. Number of phonetic neighbours
Linguistic neighbours are words that are very similar to a specific target word, usually differing in one sound only. The three Danish neighbours syng /søŋ/ ‘sing’, hæng /hen/ ‘hang’, and stæng /sdəŋ/ ‘close’, for example, are neighbours to the Danish
target word *seng* /scəŋ/ 'bed'. It can be assumed that words that are less ambiguous, i.e. less likely to be mixed up with a segmentally similar, but semantically unrelated word, are easier to understand. Luce & Pisoni (1998) computed neighbourhood density by considering three factors: the number of words that fall within a neighbourhood range (i.e. the number of words that differ in only one phoneme from the target word), the degree of similarity to the target word (i.e. the degree of similarity of the two differing phonemes), and the frequency of the neighbours. They found that word recognition correlates negatively with neighbourhood density, i.e. words in a dense neighbourhood are more ambiguous and therefore more difficult to understand than words with a low neighbourhood density. For example, the probability that the Danish word *syng* is misinterpreted by subjects, is likely to be higher than the probability that the word *motorcycle* is misinterpreted. Indeed, Kürschner, Gooskens, & Van Bezooijen (2008) found a low, but significant correlation (*r* = -.13, *p* < .001) between neighbourhood density and intelligibility scores in Danish participants confronted with spoken Swedish in a translation task. Neighbourhood density, therefore, has to be considered a linguistic feature which might be able to cause or increase an asymmetry in mutual intelligibility between Danish and Swedish.

2.3.1.4. Word length

Kürschner et al. (2008) showed that intelligibility scores correlate positively with word length (*r* = .21, *p* < .001). These results are consistent with findings from previous research, showing that longer words are easier to decode than shorter words (Scharpff & Van Heuven, 1988; Wiener & Miller, 1946). It is assumed that this is caused by the fact that short words are more ambiguous, because they have more 'neighbours'.

2.3.1.5. Speaking rate

It has been shown that higher speech rate deteriorates the intelligibility of speech (Vaughan & Letowski 1997, Gordon-Salant, Fitzgibbons & Friedman 2007). It is still controversial, however, whether speech rate (i.e. number of syllables produced per second, pauses included) and articulation rate (number of syllables per second omitting pauses) generally differ cross-linguistically (cf. Roach, 1998). Osser and Peng (1964) compared American English to Japanese, and did not find a significant difference in speech rate between these two languages. Neither did Kowal, Wiese, and O’Connell (1983), who re-evaluated findings from earlier studies based on spontaneous speech in English, German, French, Spanish, and Finnish. In her comparative study of Italian and Dutch, Den Os (1988) analysed speech and articulation rate in six native speakers per language and did not find any significant differences, when syllables per second were compared across languages. When phonemes per second were compared, however, Italian articulation and speech rate were both significantly slower than Dutch articulation and speech rate. Verhoeven, de Pauw & Kloots (2004) compared speech tempo in Belgian and Netherlandic Dutch
and found that Netherlandic Dutch is spoken at a significantly higher speech rate than Belgian Dutch.

To our knowledge, hitherto, there has been no comparative study of Danish and Swedish speech or articulation rate. As we are interested in comparing the intelligibility of a concept and the tempo of communicating this concept, we consider that a third measure should be employed for investigating the asymmetry in mutual intelligibility between speakers of Danish and Swedish, namely the number of concepts communicated per second (‘communication rate’). Alternatively, if isolated words with a similar underlying phonological structure are compared, the duration of these items could be analysed. For example, the monosyllabic Danish word *vand* /vɔn/ and the disyllabic Swedish cognate *vatten* /vatːɔn/ may be produced at a similar speech rate in the standard languages, but as they differ in number of syllables, the Danish word would have a communication rate that is twice the Swedish one. We assume, therefore, that communication rate might capture the difference between the recognition problems that a Danish and a Swedish speaker are confronted with in a more accurate way than the traditional measures articulation and speech rate do for our population, as the vast majority of words in Danish and Swedish are cognate words.

2.3.2. Extra-linguistic factors

2.3.2.1. Attitude

Wolff (1959) suggested that attitude towards a closely related variety may influence the effort made to decode it. Having investigated mutual intelligibility between speakers of two Nigerian languages (Kalabari and Nembe), he concluded that negative attitude held towards a language variety might result in less effort to decode it, while a positive attitude might encourage listeners to do their best in decoding the other variety. In Scandinavia, Sweden has been the country with the biggest population and the most prosperous industry for a long time. Therefore, within Scandinavia, Sweden is often called *storebror* (Engl. lit. 'big brother'), indicating that Norwegians and Danes generally regard Sweden as more influential and more dominant than their own country. Following Wolff’s (1959) assumption, the fact that Sweden has had a higher status in Scandinavia than Denmark might cause a bias in the willingness to understand the neighbouring language, and thereby a bias in actual intelligibility. In other words, Danish adults may be more willing to understand Swedish than the other way around, and therefore perform better in perception tasks. However, the causal relationship between intelligibility and attitude is not known. It might also be the case that intelligibility influences attitude, instead of vice versa. This would mean that Danish adults have fewer difficulties decoding Swedish than the other way around, and therefore have a more positive attitude towards the Swedish language.

To investigate the relationship between attitude and comprehension, Delsing & Lundin Åkesson (2005) conducted an intelligibility experiment, in which every participant completed a detailed questionnaire. Among other things (see section...
2.3.2.2), subjects were asked two questions to elicit their attitude towards the neighbouring country and language: (a) Do you think Danish/Swedish sounds beautiful? (b) Would you like to live in Denmark/Sweden? They showed that Danish-speaking subjects considered the Swedish language more beautiful than Swedish-speaking subjects considered the Danish language. However, in the same investigation, the asymmetry of the attitude scores was reversed when the same subjects were asked if they would like to live in the neighbouring country. Danes were less willing to move to Sweden than vice versa. In a correlation analysis, Danish comprehension scores were better predicted by the subjects’ answer to question (a), whereas Swedish comprehension scores were better predicted by the answer to question (b). Unfortunately, only $p$-values, but no correlation coefficients are reported by Delsing & Lundin Åkesson (2005), which makes it difficult to interpret these results. It seems that intelligibility and attitude are somehow associated with each other, but Delsing & Lundin Åkesson’s (2005) data provides evidence neither about the direction, nor of the strength of the causal relationship between these two factors. It is possible that a positive attitude enhances intelligibility, but it might also be the case that higher intelligibility of a given language results in a more positive attitude towards that language.

2.3.2.2. Contact

Several investigations have been concerned with the relationship between the amount of contact to a closely related variety and intelligibility of that language. It has been assumed that a higher amount of previous exposure in form of watching television, reading newspapers, visiting the neighbouring country and other forms of personal contact enhance one’s ability to decode a closely related variety.

Maurud (1976) tested mutual intelligibility in young recruits coming from the three Scandinavian capitals Copenhagen, Oslo and Stockholm. Copenhagen is located about 40 km from the Swedish border, whereas Stockholm is located about 570 km from the Danish border. This means that Copenhageners have access to Swedish broadcasting programmes and can visit the neighbouring country rather easily, while people living in Stockholm neither can receive Danish broadcasting programmes nor cross the border to Denmark within a couple of hours. If contact plays a role for intelligibility of the neighbouring language, one would expect the asymmetry between Swedish and Danish intelligibility scores found by Maurud (1976) to be larger than in other investigations. This tendency is indeed confirmed by Figure 1 above. In Maurud’s study, Swedish participants had a comprehension score of 21 percent, whereas Danish participants comprehended 40 percent. The Swedish-speaking subjects thus reached only slightly more than half of the Danish score (53 percent), whereas this figure is higher in all other investigations (Bø, 1978: 59 percent; Delsing & Lundin Åkesson, 2005: 74 percent; Gooskens & Kürschner, 2010: 76 percent).

Bø (1978) investigated mutual intelligibility of Danish, Swedish and Norwegian. The subjects were chosen in such a way that they formed two groups, one living inside and one living outside the border regions. The group of subjects living within the border regions not only had more opportunities to visit the neighbouring country,
but also had access to television programmes in the neighbouring variety. Bø (1978) found that subjects living near the border had fewer difficulties decoding the neighbouring variety than subjects living outside the border region, thereby indicating that a high degree of contact enhances intelligibility abilities.

Delsing & Lundin Åkesson (2005), in addition to eliciting individual language attitudes (see section 2.3.2.1), also collected data on four different types of contact for every participant in their intelligibility experiment. Specifically, they asked the participants whether or not they had watched television in the neighbouring language, read newspapers in the neighbouring language, had personal contact to the neighbouring language, and visited the neighbouring country. In line with Bø (1978), they found a significant correlation between contact and comprehension scores. As for the attitude data, however, the authors chose not to present any correlation coefficients, which makes the interpretation of the results difficult.

2.3.2.3. Literacy
The fact that Danish and Swedish are very closely related languages leads to their having a similar morphology, syntax, and orthography. They also share a large part of their vocabularies, i.e. there is a great number of Swedish-Danish cognates. However, contemporary Danish and Swedish differ considerably in pronunciation of these cognates. Spoken Swedish has stayed closer to its East Nordic root than spoken Danish has (Elbro 2006), resulting in a more “opaque phoneme to grapheme relation in Danish” (Bleses & Thomsen 2004) compared to Swedish. This is illustrated in Table 1 below.

In the first example given in Table 1, the Swedish word *huvud* and the Danish word *hoved* ‘head’ are both spelt CVCVC. However, the Swedish pronunciation is CVVCVC, whereas the Danish pronunciation comes close to CVV(V), if we interpret the approximant /ð/ as semivowel. Note that the phonetic sound which is transcribed by the IPA character ð is an approximant articulatorily, and not a fricative.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Three words (‘head’, ‘hand’ and ‘star’) written in Swedish and Danish contemporary orthography and IPA symbols to indicate standard pronunciation.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Orthography</strong></td>
<td><strong>Swedish Pronunciation (IPA)</strong></td>
</tr>
<tr>
<td>huvud</td>
<td>huv:od</td>
</tr>
<tr>
<td>hand</td>
<td>hand</td>
</tr>
<tr>
<td>stjärna</td>
<td>stjärna</td>
</tr>
</tbody>
</table>

In the second example *hand* vs. *hånd* ‘hand’, the final (written) consonant is pronounced in Swedish, but not in Danish. As a consequence, it can be assumed that a Dane has advantages when decoding the Swedish word because the final consonant is still written in Danish. On the other hand, a Swede has less support from the Swedish orthography when he or she hears the Danish pronunciation without the final consonant, as this form is found neither in spoken nor in written Swedish.

Examples of advantages in the opposite direction can also be given, however. The Danish word *stjerne* ‘star’, pronounced with the word-initial consonant cluster
/sdʃj/, is presumably easier to understand for a Swede than the corresponding Swedish word stjärna, word-initially pronounced with the fricative /ʃ/, is for a Dane. For a Dane it is unexpected that the written consonant cluster stj can be pronounced as /ʃ/. It can be assumed, however, that it is not equally unexpected for a Swede that stj is pronounced /sdʃ/ because this approximates an accumulated pronunciation of the three isolated phonemes /s/ /t/ /j/. Moreover, a phonological rule generally changes the consonant cluster /stʃ/ into /st/ in Swedish (Engstrand 1990), which is similar to the pronunciation in Danish /sdʃ/.

The fact that the Danish and Swedish orthographies are rather closely related makes the written forms of the languages highly mutually intelligible. This has been confirmed by many investigations (see Figure 1). However, as the Danish pronunciation has changed considerably, the distance between the spoken and written language forms is larger for contemporary Danish than for contemporary Swedish. Orthography reflects pronunciation more accurately in Swedish than in Danish. But what is more, as spoken Danish has developed further away from the common East Nordic root than Swedish has, there are instances when Danish orthography reflects Swedish pronunciation more accurately than Danish pronunciation.

Gooskens & Doetjes (2009) calculated the phonetic and orthographic distance for every pair of Danish and Swedish cognate words using the Levenshtein algorithm (for a detailed account of the application of the Levenshtein algorithm in dialectology see Nerbonne & Heeringa 2010). They then corrected the phonetic distance for the advantage that Danes and Swedes have from their native orthography. By comparing corrected and uncorrected phonetic distances, they found that both Danish and Swedish literate listeners have advantage from their reading and writing skills when confronted with the neighbouring language, since for both languages the averaged corrected distances were smaller than the averaged uncorrected distances. Interestingly, however, their data suggested that the advantage of native orthography in spoken word recognition of a closely related language is larger for Danish listeners than for Swedish listeners. In other words, the hand - hånd example given in Table 1 is more representative for the Swedish-Danish spelling-pronunciation situation than the stjärna - stjerne example. The findings by Gooskens & Doetjes (2009) therefore suggest that the asymmetry in mutual intelligibility could be caused or booster by an asymmetric advantage from native orthography. It has not been shown, however, whether Danish or Swedish listeners actually can use their native orthography in a word recognition task.

2.4. Method

2.4.1. Participants

Initially, 19 Danish-speaking and 30 Swedish-speaking children aged three to six years old were tested. To make sure that the children had not been exposed to the neighbouring language prior to the experiment, we chose participants living 200 km...
from the Swedish-Danish border. The Danish participants hailed from Odense and the Swedish participants from Växjö.

In order to be able to control for the three extra-linguistic factors attitude, contact and literacy, a questionnaire was completed for every participant. In addition to general information such as place and date of birth, the children’s parents were asked if their children were monolingual. All parents were asked if their children had visited the neighbouring country, if they had heard the neighbouring language, and if they had watched undubbed TV in the neighbouring language. The parents were also asked if the child had learnt to read and write and, if so, to indicate how many words the child could write.

After questionnaire evaluation, one Swedish child was excluded due to extensive exposure to the Danish language through his Danish father. Three Swedish children were excluded because their parents indicated that they could write “many” words or “almost everything”. No children were excluded on the basis of their attitude. Forty-five participants (19 Danish and 26 Swedish-speaking) remained for the analysis. The Danish children ranged in age from 4;6 to 6;7 and the Swedish children were aged between 4;0 to 6;8. There was no significant age difference between the language groups.

2.4.2. Stimulus material

The stimulus material was derived in the following manner: First, in a pre-experiment, 112 pictures, selected from the picture database developed at the Max-Planck-Institute for Psycholinguistics, were shown to five four-year-old Danish and five four-year-old Swedish-speaking children from Odense (Denmark) and Växjö (Sweden). These children were not used as participants in the word recognition experiment. The children were asked to label the depicted objects as spontaneously as possible. Then, a labelling consistency per picture was calculated for the most frequent label. To be selected for the experiment, a picture had to fulfil two criteria: It had to have a labelling consistency of at least 80 percent, i.e. it had to be given the same label by at least four of five children (intra-language criterion), and it had to be labelled with a cognate word (inter-language criterion). 53 pictures met these two criteria. Three of them were used in a demo version, leaving 50 pictures for the experiment. Their labels were used as auditory stimuli. In this way, it was ensured that the target pictures (i.e. the pictures that corresponded to the auditory stimuli in the actual experiment) were recognised and produced by children even younger than the age group tested in the experiment and labelled unambiguously by these children. The remaining 59 pictures, supplemented with further 91 pictures from the same database, served as distracter pictures for the word recognition experiment reported below. Figure 5 gives some examples of the pictures employed.
Figure 5. Example set of pictures, taken from the picture database of the Max-Planck-Institute for Psycholinguistics in Nijmegen, the Netherlands. These four pictures were presented simultaneously with the stimulus *gaffel* ‘fork’.

To ensure that linguistic features could be interpreted in a reliable manner between the employed varieties, we had to make sure that the children would be presented with a regional standard that corresponded to the variety that was spoken in their environment. The speakers were therefore primarily chosen on the basis of their regional standard language. The 53 labels per language (henceforth ‘stimuli’) were therefore produced by two female native speakers from Odense (Denmark) and Växjö (Sweden), respectively. The recordings took place in sound-attenuated rooms and the sound files were digitised at 44.1 kHz.

The stimulus material was originally chosen on the basis of frequency, cognateness, and ‘picturebility’, and not primarily on the basis of linguistic features. This is true for basically all other studies cited in section 2.1., which reported an asymmetry in mutual intelligibility between Danish and Swedish. To make sure, however, that the material exhibits the linguistic features discussed in section 2.3.2, we conducted an analysis on toneme distribution, distribution of stød, word length, and communication rate. The neighbourhood density was not calculated for our material. We assume that the influence of the neighbourhood effect is excluded, as the experiment employed a multiple-choice task. The results of the evaluation are displayed in Table 2. It shows the linguistic features of the stimulus material, and gives the corresponding values for the stimulus material used by Gooskens & Kürschner (2010), who, in line with all other investigations, reported a significant asymmetry, but did not control for literacy, attitude, and only partly for contact (see section 2.3.1.1).

Table 2. Linguistic features of the stimulus material used in Gooskens & Kürschner (2010) and in the present study.

<table>
<thead>
<tr>
<th></th>
<th>Present experiment</th>
<th>Gooskens &amp; Kürschner (2010)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Danish</td>
<td>Swedish</td>
</tr>
<tr>
<td>Words with toneme 1 (%)</td>
<td>-</td>
<td>64.0</td>
</tr>
<tr>
<td>Words with stød (%)</td>
<td>54.0</td>
<td>-</td>
</tr>
<tr>
<td>No. of phonemes</td>
<td>4.42</td>
<td>4.60</td>
</tr>
<tr>
<td>No. of syllables</td>
<td>1.70</td>
<td>1.68</td>
</tr>
<tr>
<td>No. of phonemes per syllable</td>
<td>2.86</td>
<td>2.92</td>
</tr>
<tr>
<td>Duration (ms)</td>
<td>478</td>
<td>719</td>
</tr>
</tbody>
</table>
As can be seen in Table 2, some of the investigated linguistic features are distributed differently across the two languages. Both in the present study and in Gooskens & Kürschner's (2010) study, a slight majority of the Danish stimuli had stoð, and a clear majority of Swedish stimuli had accent 1. The latter is due to the fact that highly frequent words are often monosyllabic, and monosyllabic words always have accent 1. These figures are therefore not necessarily representative for other Swedish samples, but they should be roughly comparable across the two experiments. Furthermore, in both experiments, the Swedish words consist of slightly more phonemes than the Danish words do. Interestingly, however, in both studies, the number of syllables is slightly larger for the Danish stimuli than for the Swedish. This means that, at least in these two samples, Danish syllables consist of fewer phonemes than Swedish syllables do. In both sets of stimuli, word duration was shorter for Danish stimuli (478 and 543 ms) than for Swedish stimuli (719 and 820 ms).

We therefore assume that the stimulus material we used in the experiment reported in this paper has linguistic features similar to the materials used in Gooskens & Kürschner's (2010) experiment, where a clear asymmetry in mutual intelligibility was found. As three crucial extra-linguistic factors are held constant here, an asymmetry in mutual intelligibility could indicate that it linguistic features play an important role for the asymmetry in mutual intelligibility.

2.4.3. Procedure

The testing session consisted of a picture-pointing task followed by a short interview with every child. Before the experiment started, the children were familiarised with the task by being presented a short demo version of the experiment. The child sat in front of a touch screen (LG L1510SF) wearing a head set. The demo version of the experiment consisted of two trials with stimuli in the children's native language, followed by one trial with a stimulus in the test language. Four pictures per stimulus were presented on the touch screen. The children were instructed to point to the picture that corresponded to the stimulus. Before the word-recognition experiment started, it was ensured that the children had understood the task.

The experimental design and the task of the word-recognition experiment were the same as in the demo version. 50 trials were presented in random order. In every trial, one auditory stimulus was presented in the neighbouring language. Simultaneously to word onset, four pictures appeared on a touch screen and remained on the screen until the participants touched the screen or for 10 000 ms. Every auditory stimulus was presented together with the same set of four pictures across sessions and languages. The inter-stimulus interval was 1000 ms and no feedback was given. After the experimental part, attitudes towards the neighbouring language held by the participants were assessed by asking the children to rate the language in question on a three-point scale. The scale was designed in such a way that also the youngest children were able to make a choice. The question was posed to the children after the experiment had been conducted. Specifically, the children were asked the following question: “The language you just heard was Danish/Swedish. Did
you like it more or less than the language you speak yourself?” The children could provide answers by choosing one of the following categories: (a) nicer than my native language, (b) as nice as my native language, or (c) less nice than my native language.

2.5. Results

Both groups of children held rather neutral attitudes: 89.4 percent of the Danish and 57.8 percent of the Swedish children either had no opinion about the neighbouring language or indicated that they found the neighbouring language as nice as their native language, i.e. answer (b). From the remaining Danish children, 50.0 percent indicated that they found Swedish nicer than their native language and from the remaining Swedish children, 45.4 percent indicated that they found Danish nicer than their native language (answer (a)). The two groups of children (Danish and Swedish) did not hold significantly different attitude towards the neighbouring language ($t(43) = 0.23, p = .82$).

The Danish children decoded 63 percent ($S = 0.12$) and the Swedish children decoded 65 percent ($S = 0.12$) of the presented stimuli. This difference was not significant $t(43) = -.50, p = .60, r = .08$). That means that, in contrast to adult Swedes, Swedish children did not encounter more problems decoding the neighbouring language than their peers from Denmark. The results are illustrated in Figure 6, which displays our results next to the results from previous research that were presented in section 2.1. Note that absolute intelligibility scores cannot be compared between adults and children, because the experimental design differed substantially across the studies. It can therefore not be concluded that children outperform adults from earlier investigations. The asymmetry, however, should be comparable, as the same experimental designs were used across the two languages in all investigations.

Figure 6. Child and adult intelligibility scores in Danish listeners confronted with Swedish stimuli, and Swedish listeners confronted with Danish stimuli found in the present experiment, and those reported by Maurud 1976, Bo 1978, and Delsing & Lundin Åkesson 2005, and Gooskens & Kürschner 2010.
2.6. Discussion and conclusion

Generally, the intelligibility scores obtained in this experiment were slightly higher than the scores reported in previous investigations, so we judged the task to be appropriate for this age group. As our aim was to investigate the role of extra-linguistic factors for mutual intelligibility between Danish and Swedish, we conducted a word recognition experiment with stimuli that had similar linguistic features as the stimuli employed in Gooskens & Kürschner’s (2010) study. At the same time, we controlled for extra-linguistic factors by choosing participants that were shown to be illiterate, to hold similar (neutral) attitudes towards the test language, and had a similar (very low) amount of contact with the test language. We hypothesised that linguistic factors account for a large part of the asymmetry and expected the asymmetry in mutual intelligibility to persist in our subjects even if the influence of three extra-linguistic factors suggested by previous researchers (attitudes, contact, and literacy) were excluded. Our results, however, revealed that mutual intelligibility of Danish and Swedish in preschoolers’ did not differ significantly across the two language groups. This finding suggests that extra-linguistic factors play an important role for the asymmetry in mutual intelligibility between adult Danes and Swedes. A comparison of our data with results from previous research suggests that, when one or several of the three extra-linguistic factors change with age, they give rise to an asymmetry in mutual intelligibility.

We hypothesised that the fact that mutual intelligibility of spoken language is asymmetric, while mutual intelligibility of written language is not, is due to linguistic factors that are found in spoken language only. Our results suggest, however, that the asymmetry is due to extra-linguistic factors that are relevant in spoken communication only, such as attitude towards a spoken variety or activation of native orthography during word recognition of the neighbouring language. These topics should be investigated more in depth in future research.

We cannot make assumptions about which of the three extra-linguistic factors are the most influential. As all three factors were close to being excluded, it is not possible to run a regression analysis with the three extra-linguistic factors per participants as independent variables and the individual intelligibility score as dependent variable. Therefore, we plan to conduct further studies investigating the roles of attitude and literacy for receptive bilingualism in children, as well as elicit intelligibility scores from Danish and Swedish adults.

Finally, we have to consider the possibility that the asymmetry in adult mutual intelligibility is caused by the neighbourhood effect (see above). This factor is almost completely excluded in a multiple-choice task as employed in our experiment, since our subjects did not have the possibility to translate freely, but had to select from only four different answers. The influence of this linguistic factor needs therefore to be examined in more detail as well.
Acknowledgements

This work was supported by the The Netherlands Organisation for Scientific Research (NWO) [grant number 276-75-005].
We thank Anne Cutler for generously providing us with access to the picture database developed at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands. Furthermore, we are grateful for valuable comments by Renée van Bezooijen, Nanna Haug Hilton, Sebastian Kürschner and two anonymous reviewers on an earlier version of this paper.

References


Investigating the role of language attitudes for mutual intelligibility using reaction time

This paper will appear as

Abstract

Danish and Swedish are mutually intelligible to a certain extent, but it has been shown that adult Danes confronted with spoken Swedish recognise more items than adult Swedes who are confronted with spoken Danish. However, this asymmetry was not confirmed for illiterate Danish and Swedish pre-schoolers, which suggests that the factors that were controlled for in the study with preschoolers, namely literacy, previous exposure and attitude, cause the asymmetry in mutual intelligibility in adults. In this paper, we investigate what attitudes adults and pre-schoolers hold towards the neighbouring language, and whether there is a relationship between attitudes held towards the neighbouring language and abilities to decode it. Attitude elicitation from 45 Danish-speaking and 39 Swedish-speaking participants revealed that attitudes change with age, but individual reaction time measurements towards 50 auditorily presented cognate nouns in a multiple-choice picture-pointing task showed no significant correlation with the individual attitudes we elicited.

3.1. Introduction

Within Scandinavia, communicating across linguistic borders using the language of the speaker is a habit strongly encouraged by the authorities. Danes, Norwegians and Swedes are likely to use their native language rather than a lingua franca when speaking to each other. This manner of communication has been called semicommunication by Haugen (1966) and receptive bilingualism by Hockett (1958). Haugen defined semicommunication as “the trickle of messages through a rather high level of ‘code noise’” (Haugen 1966: 281). As Braunmüller (2002) pointed out, Börestam (1997) tracked back the term “code noise” to Hockett (1958), who defined code noise as “divergence between the codes of two people who communicate with each other via speech” (Hockett 1958:331f.). Haugen thus suggests that semicommunication is characterised by the differences between the two varieties employed. It was also Hockett (1958), who defined semi-bilingualism as “receptive bilingualism accompanying productive monolingualism” (Hockett 1958:327), thereby introducing the term receptive bilingualism. While the term semicommunication focuses on the problems of this communication mode, receptive bilingualism emphasises the benefits rather than the shortcomings.

A lot of research on the mutual intelligibility of closely related language varieties has been conducted in Scandinavia over the past decades. Among many other things, it has been shown that mutual intelligibility between the closely related languages Danish and Swedish is asymmetrical in such a way that Danes have fewer
difficulties recognizing spoken Swedish words than Swedes have recognizing spoken
Danish. This asymmetry has been reported in several studies (Maurud, 1976; Bø,
1978; Delsing & Lundin Åkesson, 2005; Gooskens & Kürschner, 2010). In a recent
study, however, Schüppert & Gooskens (2010) showed that this asymmetry is not
found in pre-schoolers. In contrast to previous investigations of mutual intelligibility
between Danish and Swedish (Maurud, 1976; Bø, 1978; Delsing & Lundin Åkesson;
2005, Gooskens & Kürschner, 2010), the Swedish-speaking participants scored as
high as the Danish speaking participants in a word recognition task, indicating that
one or several of the factors that were controlled for by Schüppert & Gooskens (2010)
cause the asymmetry in mutual intelligibility that is consistently found among adults.
The variables that were controlled for were literacy, attitude held towards the
neighbouring language, and previous exposure to it. These factors have been shown
to have an impact on intelligibility in earlier studies. Maurud (1976) and Delsing &
Lundin Åkesson (2005) report that a larger amount of contact to the neighbouring
language is linked to better word recognition, and Delsing & Lundin Åkesson (2005)
and Gooskens (2006) found that a positive attitude towards the neighbouring
language is associated with higher word recognition. Gooskens & Doetjes (2009)
suggested that orthographic knowledge serves as an additional cue for Danish
listeners, as Danish orthography is more conservative than Swedish orthography and
therefore closer to Swedish pronunciation than vice versa.

In the present paper, we investigate the relation between the participants’
attitudes and their word recognition abilities. Evidence that these two factors
correlate was provided in earlier studies by Wolff (1959), Delsing & Lundin Åkesson
the two closely related Nigerian languages Kalabari and Nembe and found that
speakers of Nembe, which is commonly regarded as having a lower status than
Kalabari, have fewer difficulties understanding Kalabari, than vice versa. However,
Wolff (1959) did not elicit attitudes empirically from his participants and based his
conclusion on anecdotal evidence about the two languages.

In contrast to Wolff (1959), Delsing & Lundin Åkesson (2005) elicited their
Danish and Swedish-speaking subjects’ attitudes by posing them the following two
questions: (a) Would you like to live in Sweden/Denmark? and (b) Do you think
Swedish/Danish sounds nice or ugly? The answers to the first question were ternary
(yes, perhaps, no), and the answers to the second question were given on a five-point
Semantic Differential scale, ranging from nice to ugly. The individual attitude scores
were correlated with the individual comprehension scores obtained in three different
tasks: (i) watching a video sequence in the neighbouring language and answering five
open questions about the sequence, (ii) listening to news in the neighbouring
language and answering five open questions about its content, and (iii) reading a
newspaper article and answering ten multiple-choice questions about it. Delsing &
Lundin Åkesson’s (2005) results showed that attitudes towards the country (i.e.
answers to question (a)) correlated significantly with comprehension scores for the
Swedish but not for the Danish participants, while attitudes towards the language
(i.e. answers to question (b)) correlated significantly with comprehension scores for the Swedish but not for the Danish participants. In other words, Swedish comprehension of Danish is partially predicted by the participants’ attitudes towards the country of Denmark, whereas Danish comprehension of Swedish is partially predicted by the participants’ attitude towards the Swedish language. In a re-analysis of Delsing & Lundin Åkesson’s (2005) data, Gooskens (2006) correlated intelligibility scores and language attitudes per site and per test language. She reports a significant correlation between answers to question (b) and comprehension scores, but not between answers to question (a) and comprehension scores. The fact that the effect reported by Delsing & Lundin Åkesson (2005) disappeared under aggregation of the data confirms their finding that the link between language attitudes and intelligibility is not very strong and could only be established for a specific group of participants.

On the basis of Delsing & Lundin Åkesson’s (2005) and Gooskens’ (2006) results, we hypothesise that there is a link between individual word recognition and attitude towards the sound of the neighbouring language. More specifically, we assume that positive attitudes correlate with good word recognition abilities, but any causal relationship between these two variables will remain unknown in this experiment. It is possible that participants holding a positive attitude towards the neighbouring language make a greater effort understanding it; but it might also be the case that higher comprehension of a language variety leads to a more positive attitude. A third possibility is that attitude and comprehension are not linked directly, but that both variables are interrelated. For example, a positive attitude might not cause higher comprehension, but a higher amount of contact, which, in turn, might lead to higher comprehension. This study investigates attitudes held towards the neighbouring language in Danish and Swedish children and adults, as well as the link between attitudes and intelligibility of the neighbouring language.

3.2. Method

3.2.1. Participants

Participants were 19 Danish-speaking and 27 Swedish-speaking three to six-year-old preschoolers and 21 Danish-speaking and 19 Swedish-speaking 17 to 20-year-old adults. All subjects lived 200 km from the Swedish-Danish border, the Danish participants in Odense and the Swedish participants in Växjö. The children were tested at their day-care. Several day-care institutions in Växjö and Odense had been approached and two at both sites (i.e. four in total) were eventually picked for participation. As the experiment was conducted individually, sessions lasted about 30 minutes per child. To ensure that all children were tested at roughly the same time of the day, testing took place before noon and during several days. All children were part of the study reported in Schüppert & Gooskens (2010).

A questionnaire was used to exclude participants with previous high exposure to the neighbouring language. The children’s parents were asked if their children were monolingual, and the adult participants were asked the same question and, in
addition, which foreign languages they had acquired. In the same way, all children’s parents and all adult participants were asked if they had been to the neighbouring country, how often they had heard the neighbouring language, and if they had watched undubbed TV in the neighbouring language. After questionnaire evaluation, one Swedish child was excluded due to extensive contact with the Danish language through his Danish father, and one Danish adult was excluded due to L2 acquisition of Swedish. Eighty-four participants remained for the analysis. No participants were excluded on the basis of previous exposure due to occasional TV watching or short visits to the neighbouring country. An independent t-test revealed that the amount of previous exposure to the neighbouring language differed across the two groups of L1 speakers Danish and Swedish) neither in the children nor in the adult participants.

Table 1 gives an overview of how age and sex were distributed in the four groups. While the two groups of children did not differ significantly in any of these respects, the adult Danes were predominantly female (80 percent) and the adult Swedes were predominantly male (74 percent).

<table>
<thead>
<tr>
<th></th>
<th>Danish</th>
<th>Swedish</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Children</td>
<td>Adults</td>
</tr>
<tr>
<td>N</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>Age (Error! Bookmark not defined. $\bar{x} \pm S$)</td>
<td>5.7 ± 0.6</td>
<td>18.6 ± 1.2</td>
</tr>
<tr>
<td>Age range (years)</td>
<td>4.5 - 6.6</td>
<td>17 - 20</td>
</tr>
<tr>
<td>Males (%)</td>
<td>47</td>
<td>20</td>
</tr>
</tbody>
</table>

3.2.2 Stimulus material

The auditory stimulus material consisted of 50 Swedish-Danish cognate nouns. The visual material consisted of 200 pictures, of which 50 were target pictures and 150 were distracters. In order to find appropriate labels for the target pictures, and in order to make sure that target pictures would be clearly labelled with a cognate word even by the youngest children in the experiment, 112 pictures from the picture database developed by the Max-Planck-Institute for Psycholinguistics were shown to five Danish and five Swedish four-year-old children in a pre-test. The children participating in the pre-test were asked to label these pictures spontaneously, i.e. to name the object on the picture with one single word. To be included in the stimulus material, a picture had to meet two criteria. The intra-group criterion was a labelling consistency of at least 80 percent: only pictures that were labelled the same by at least four out of five children per language group were included in the stimulus material. If several labels were given by a child, only the first label was used for the calculation of the labelling consistency of every picture. The inter-group criterion was a cognate label. For example, the picture of a girl was consistently labelled pige by the Danish children and flicka by the Swedish children. These two words are not
cognates; therefore, this criterion was not met and the picture was excluded. 53 pictures met both the intra-group and the inter-group criterion.

The 106 labellings (53 Danish and 53 Swedish) were used as auditory stimulus material. They formed pairs of cognates with different degrees of phonetic distances, as indicated in Table 2. To calculate phonetic distances, the Levenshtein algorithm was employed, which identifies the ‘cheapest’ way to transform one string into another and counts the number of ‘costs’, i.e. operations (substitutions, insertions and deletions) needed for this transformation. Neither the Danish suprasegmental feature of laryngealisation (‘stød’) nor differences in phoneme quantity were counted as a deviating phoneme when calculating phonetic distances between Danish and Swedish. For a detailed discussion of the application of the Levenshtein algorithm for measuring phonetic distances, see Nerbonne & Heeringa (2010). Figure 1 illustrates the distribution of phonetic distances across the stimulus material.

Table 2. Examples of phonetic distances between Danish and Swedish stimuli.

<table>
<thead>
<tr>
<th>Danish Spelling</th>
<th>Danish Pronunciation</th>
<th>Swedish Spelling</th>
<th>Swedish Pronunciation</th>
<th>Phonetic distance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>måne</td>
<td>mo:n</td>
<td>måne</td>
<td>mo:n</td>
<td>0</td>
</tr>
<tr>
<td>båd</td>
<td>bo:ð</td>
<td>båt</td>
<td>bo:t</td>
<td>33</td>
</tr>
<tr>
<td>æble</td>
<td>ebla</td>
<td>äpple</td>
<td>epla</td>
<td>50</td>
</tr>
<tr>
<td>hoved</td>
<td>ho:ɔð</td>
<td>huvud</td>
<td>huvod</td>
<td>80</td>
</tr>
<tr>
<td>abe</td>
<td>e:ða</td>
<td>apa</td>
<td>a:pa</td>
<td>100</td>
</tr>
</tbody>
</table>

The stimuli were produced by two female native speakers who had grown up and still lived in Odense and Växjö, respectively. Three of the 53 stimuli were used for a training session that was shown in advance to every participant, leaving 50 stimuli for the experiment.

Figure 1. Distribution of phonetic distances between Danish and Swedish in the stimulus material.

3.2.3 Procedure

The testing session consisted of a stimulus-response experiment followed by a short interview with the participant. Both parts were conducted by a native speaker of the language spoken by the participant. The participant sat in front of the touch screen.
wearing ear-phones and was told that he or she would be presented with words from the neighbouring language, i.e. Danish participants were informed that they would hear Swedish and vice versa. The 50 stimuli were presented auditorily and, simultaneously with stimulus onset, four pictures per stimulus (i.e. 200 pictures in total) appeared on a touch screen (LG L1510SF). The pictures remained on the screen from the onset of the first presentation of the spoken word until the participant touched the screen, or until a time-out occurred after 10 sec. The subjects’ task was to match every auditory stimulus to a picture by touching it on the screen. The stimulus material was presented in random order, but the same four pictures were assigned to every stimulus across participants and across languages. An example is given in Figure 2. One session lasted between one and four minutes, depending on how quickly the participants responded. Prior to the experiment, every participant was familiarised with the task by being presented with a training session: two native language stimuli were presented first, followed by a stimulus in the neighbouring language. When the training session was completed, the participants were asked whether their task was clear, and, if necessary, further instructions were given before the experiment started.

Figure 2. Four pictures presented simultaneously with the Danish stimulus *hoved* /hoːd/ ‘head’ to Swedish participants and the Swedish cognate *huvud* /huːvʊd/ to Danish participants.

After the experiment, the participants’ attitudes towards the neighbouring language were assessed by asking whether the language they had heard during the experiment sounded (1) less nice than, (2) as nice as, or (3) nicer than their native language. The participants could also refrain from making a decision by choosing a fourth option, ‘no opinion’.

3.3. Results and discussion

3.3.1. Word recognition experiment

Figure 3 shows word recognition frequencies for all participants. The word recognition scores are non-normally distributed, which is confirmed by a Kolmogorov-Smirnov test with Lilliefors Significance Correction ($D(84) = 0.19, p < .001$). The scores are negatively skewed ($z_{skewness} = -2.86, p = .01$).
To investigate whether this is a tendency that is found to an equal extent in all four groups of participants or whether it is restricted to one or several of them, we split up the word recognition scores into the four sub-groups. Figure 4 displays word recognition frequencies for all participants per sub-group.

It can clearly be seen in Figure 4 that the distribution of word recognition scores differs across the two age groups (children vs. adults). Whereas the distribution of scores is near-normal or normal in the two groups of children, the scores of the adult group of participants are at ceiling, as all Danish and 89 percent of the Swedish participants recognise more than 90 percent of the stimuli correctly. This effect is presumably due to the fact that the experiment was designed in such a way that even four-year-old children should be able to complete it. The ceiling effect suggests that
adults use more cues than their native language to recognise the stimuli and therefore generally are more successful in word recognition than children are. These cues could be foreign-language knowledge, dialect knowledge, or their native language orthography.

In addition to word recognition scores, we measured reaction times (RTs) per participant and per stimulus. Word recognition scores were obtained by multiplying the number of correct identifications per participant by 2 as 50 stimuli were presented. Mean RT was calculated on basis of correct identifications only. RTs were originally measured from word onset. RT distribution was normal, as indicated by a Kolmogorov-Smirnov test with Lilliefors Significance Correction ($D(84) = 0.09, p = .07$). It was crucial, however, to normalise for the fact that the stimuli differed with respect to the number of phonemes and speech rate, across stimuli as well as across trials. Specifically, the Danish stimuli generally had a shorter duration ($\bar{x} = 478 \text{ ms}$) than the Swedish stimuli had ($\bar{x} = 719 \text{ ms}$), which is likely to yield shorter RTs for the Swedish participants listening to Danish than for Danish participants listening to Swedish, if RTs are measured from word onset. Therefore, word duration was subtracted from every obtained RT per word before individual RT scores were calculated on the basis of correctly decoded stimuli. This means that RTs were measured from word offset. Figure 5 shows a box plot of reaction times per age group. An independent t-test revealed that the children’s RTs differed significantly from the adult RTs ($t(82) = 5.44, p < .001$). Therefore, in order to be able to correlate children’s and adults’ RTs with attitudes in one analysis, we normalised for the differences in age by calculating $z$-scores for both groups of participants (children and adults). These normalised RTs are displayed in Figure 6.

**Figure 5.** Distribution of RTs in both age groups.

![Box plot of reaction times](image)

**Figure 6.** Distribution of age-group specific $z$-scores of RTs in both age groups.

![Box plot of z-scores](image)

It is generally assumed that the time it takes a participant to make a decision reflects the processing time and thereby the degree of complexity of the task (Gass & Mackey, 2007: 22ff). Because of the ceiling effect among the adult participants, we neglect accuracy scores in our further analysis and analyse RT as the dependent measure instead. The mean RTs are shown in Table 3.
Table 3. Mean RTs (ms) for all four groups of participants.

<table>
<thead>
<tr>
<th>Age group</th>
<th>L1</th>
<th>( \bar{x} )</th>
<th>S</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children</td>
<td>Danish</td>
<td>2809</td>
<td>592.19</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Swedish</td>
<td>2699</td>
<td>637.47</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>2746</td>
<td>614.30</td>
<td>45</td>
</tr>
<tr>
<td>Adult</td>
<td>Danish</td>
<td>1898</td>
<td>370.51</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Swedish</td>
<td>2297</td>
<td>469.70</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>2092</td>
<td>462.60</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Danish</td>
<td>2342</td>
<td>669.01</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Swedish</td>
<td>2529</td>
<td>601.26</td>
<td>45</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>2442</td>
<td>636.73</td>
<td>84</td>
</tr>
</tbody>
</table>

An independent t-test revealed that RT differed significantly between Danish and Swedish participants in the adult group (\( t(37) = -2.96, p = .005 \)), which confirms the asymmetry reported in earlier studies (Maurud, 1976; Bø, 1978; Delsing & Lundin Åkesson, 2005; Gooskens & Kirschner, 2010). The difference in reaction times between Danish and Swedish-speaking children, on the other hand, was not statistically significant (\( t(43) = 0.59, p = .56 \)). This result extends the one reported in Schüppert & Gosskens (2010), and is based on an expansion of their data (a superset).

It was pointed out in section 3.2.1 that the group of adult Danes consisted mainly of females (80%), while the group of adult Swedes mainly consisted of male participants (74%). That means that females tend to dominate in the cells with shorter reaction times and males tend to dominate in the cells with longer reaction times. There has been a substantial amount of research on differences in reaction times between men and women, but, importantly, men have consistently been shown to have shorter reaction times than females (Noble et al., 1964; Welford, 1980; Adam et al., 1999; Dane and Erzurumlugoglu, 2003; Der and Deary, 2006). Bellis (1933) reported that mean time to press a key in response to a light was 220 ms for males and 260 ms for females. The mean RTs to a sound were 190 ms (males) to 200 ms (females). Engel (1972) reported a RT to sound of 227 ms (male) to 242 ms (female). Interestingly, however, Barral and Debu (2004) found that while men were faster than women at aiming at a target, women were more accurate. Jevas and Yan (2001) reported that age-related deterioration in reaction time was the same in men and women. We assume, therefore, that the sex-related bias does not cause the asymmetry in RTs, but, rather, that the differences in RT between Danish and Swedish-speaking participants might be somewhat larger if sex had been controlled for more effectively.

In the same way, it could be argued that measuring reaction times from word offset is an overcorrection since Danish subjects will have a head start in their stimulus processing, as they have been exposed to the first part of the word for a longer time. To test whether the asymmetry remains significant if reaction times are...
measured from word onset instead of word offset, we conducted an independent \( t \)-test, which revealed that even RT measured from word onset differed significantly between Danish and Swedish participants in the adult group (\( t(37) = -2.88, p = .007 \)), but not in the child group (\( t(43) = 0.24, p = .82 \)). This means that the asymmetry in RTs measured from word offset are not due to the fact that Danes had an advantage because they had been exposed longer to the stimuli.

### 3.3.2. Attitudes

Figure 7 shows language attitudes held towards the neighbouring language, i.e. attitudes towards Danish held by Swedish participants and vice versa. It can be seen that children generally hold more neutral attitudes, while a rather negative attitude prevails among adults. More specifically, 89.4 percent of the Danish and 57.8 percent of the Swedish children either had no opinion about whether the neighbouring language sounds nicer than their native language, or judged it as equally nice as their meeting language.

![Figure 7](image_url)

In contrast to the children, the adults have clearer attitudes towards the neighbouring language, as the majority (61 percent) chose either ‘less nice than my native language’ (40.0 percent of the Danish adults and 84.2 percent of the Swedish adults) or ‘nicer than my native language’ (25.0 percent of the Danish adults and 5.3 percent of the Swedish adults).

If participants that had no opinion on this question were excluded, a Mann-Whitney-test with attitude as test variable and native language as a grouping variable
confirmed that the Swedish participants in the adult group are significantly more negative towards their neighbouring language than the Danish participants ($U = 76.5$, $p = .017$). The difference in attitude scores between Danish and Swedish-speaking children, however, was not significant. This finding suggests that the asymmetry in word recognition as indicated by RTs might be associated with an asymmetry in attitude held towards the neighbouring language. Furthermore, as Danish attitudes change from rather neutral to either positive or negative while Swedish attitudes change from rather neutral to negative, our data suggests that a change of attitude towards the neighbouring language takes place mainly in the Swedish-speaking group. This is illustrated in Figure 8, which displays means attitude score per age group and L1.

**Figure 8.** Mean attitude scores per age group for both language groups.

In other words, the difference between adult Danish and adult Swedish attitudes held towards the neighbouring language is not due to the Danish-speaking participants developing a more positive attitude towards Swedish with age, but to Swedish-speaking subjects developing a more negative attitude towards Danish. However, as the oldest pre-schooler in this study was 6.7 years old and the youngest adult participant was 17.0 years old, there is a large gap in our data. Although our data indicates that a significant shift in language attitudes takes place among the Swedish participants between the ages of 7 and 18, we cannot pinpoint the exact age period of this shift.

### 3.3.3. Correlation between attitude and word recognition

In section 3.3.1, it was shown that there was a significant difference in speed of word recognition between Danish and Swedish adult participants, and in section 3.3.2, we found a significant difference in attitudes held towards the neighbouring language. These results confirm the findings by Wolff (1959), Delsing & Lundin Åkesson (2005) and Gooskens (2006), who reported that that the group of L1 speakers that had a more negative attitude towards the neighbouring language was also the group encountering more difficulties decoding this language. No differences were observed.
between Danish and Swedish children, neither in attitude towards their neighbouring language nor in speed of word recognition.

We hypothesised that participants with a positive attitude towards the neighbouring language would perform better than participants with a negative attitude. To test this hypothesis, we conducted a Pearson correlation analysis between age-normalised RTs (see section 3.3.1.) and attitude scores, which resulted in a nonsignificant correlation coefficient of $r = -.09$ ($p = .20$, one-tailed). This is illustrated by a scatter plot in Figure 9, which shows the participants’ RT $z$-score broken down by their attitudes.

**Figure 9.** $Z$-normalised RT per attitude score.

Importantly, however, the correlation coefficients are significant neither for the groups as a whole ($r = .08$, $p = .26$, one-tailed), nor for any of the two age groups apart (preschoolers: $r = -.03$, $p = .44$, one-tailed, adults: $r = -.08$, $p = .15$) or for any of the two language groups (Danes: $r = .03$, $p = .44$, one-tailed, adults: $r = .19$, $p = .12$). This means that participants with a positive attitude perform equally well as participants with a negative attitude, indicating that, in contrast to our hypothesis, there is no link between a participant’s attitude and his or her word recognition abilities. This is in conflict with results reported by Gooskens (2006) and, partly, with those reported by Delsing & Lundin Åkesson (2005). Gooskens (2006), however, correlated attitudes and comprehension scores aggregated per test site rather than individually. This suppresses variance and increases correlation coefficients compared to calculations based on individual attitude and comprehension scores (as in the present study). Delsing & Lundin Åkesson (2005) correlated individual attitude (willingness to move to the neighbouring country and perceived beauty of the neighbouring language) with individual scores in the comprehension test for the 288 Danish-speaking participants listening to Swedish and for the 222 Swedish-speaking participants listening to Danish. They employed two different test series, and every participant completed one of them. They thus report significance values for eight different correlation analyses (2 language groups x 2 test series x attitude questions), each of which was based on approximately 100-150 participants. Five out of these eight correlation analyses yielded significant results. However, their sample sizes are approximately twice as big as ours, which inevitably yields larger significance values,
but does not distort correlation coefficients. It might be the case that the difference in significance values between our and their study is linked to the difference in sample sizes, and thereby due to the lack of statistical power in our study. This supposition, however, cannot be evaluated as Delsing & Lundin Åkesson (2005) did not report their correlation coefficients, but solely the significance values. Therefore, a comparison of correlation coefficients across the studies is not possible.

Another possible explanation for the deviation between Delsing & Lundin Åkesson’s (2005) and our findings might be found in the participants, as the preschoolers in our study held more neutral attitudes, which results in much less variance than in the adult group. Correlation coefficients in our adult group are higher and closer to significance than for the group as a whole, despite the fact that this correlation analysis is based on half as many participants, namely 32.

In fact, analysing preschoolers and adults separately might do data such as ours more justice than analysing them together as it is likely that adults and children base their judgments of the neighbouring language on different factors. It can be assumed that adults’ attitude is more likely to be contaminated by stereotypical ideas about the neighbouring country and its inhabitants than children’s attitudes are. Our data, however, does not provide evidence for this conclusion to be drawn, and further research needs to be conducted to test this hypothesis.

Another difference between our study and the study by Delsing & Lundin Åkesson (2005) is the fact that attitude ratings in our study were explicitly based on the speech sample employed, as participants were asked how they liked the language they had heard during the word recognition experiment. In contrast, Delsing & Lundin Åkesson (2005) overtly asked they participants how they liked the neighbouring language. It is likely that participants in their study referred to more than just the speech sample that was used in the comprehension tests, but also evaluated speakers of the neighbouring language that they had heard in completely different circumstances. By restricting our attitude elicitation to one speech sample per language, differences in voice quality between our two native speakers such as lively or monotonous intonation, the absence or presence of creaky voice, or differences in speech rate are more likely to bias attitude data.

Future research should therefore focus more in-depth on the development of language attitudes, i.e. how language attitudes might be influenced by stereotypes about a language variety and speakers, as well as the age-factor in the development of this influence. Also, the role of speaker or language-specific traits that influence language attitudes, such as differences in intonation or speech rate, should be investigated more thoroughly.

3.4. Conclusion

In this paper, we tested the hypothesis that attitude and word recognition of a closely related language are linked. We found that our two adult L1 groups (Danish and Swedish-speaking, respectively) differed significantly in their RTs when they were
presented with stimuli from the neighbouring language. They also had significantly different attitudes, suggesting that there might be a link between these two variables.

In section 3.1, it was pointed out that, even if a negative correlation between attitude and speed of word recognition is found in our data, it will not be possible to answer the question of the causal relationship between these variables, asking whether a positive attitude causes high recognition scores, or whether high recognition scores cause a positive attitude towards the language. As the variables attitude and speed of word recognition do not correlate in our data, however, this suggests that (a) a participant’s attitude does not have an influence on his or her word recognition performance, and (b) the difficulties that a participant encounters when confronted with a closely related language do not have an influence on his or her attitude towards that language.

However, a restriction in our study was the overt elicitation of consciously held attitudes towards a language by asking the participants directly if they think that the neighbouring language sounded less nice than, as nice as, or nicer than their native language. As it has been shown that that a person’s way of talking can elicit stereotypical ideas of the speaker (Giles & Coupland 1991), it is possible that participants who recognised the neighbouring language did not solely judge the sound of the language, but unconsciously incorporate their stereotypical ideas about speakers from the neighbouring country into their judgment. Future research should therefore focus more in-depth on the relationship of attitudes and word recognition by eliciting consciously and subconsciously held attitudes. This could be done by using the well-established affective priming task (Fazio et al. 1995) which has been employed successfully with auditory stimuli by Impe (2010) in an investigation of language attitudes towards Dutch varieties in Belgium and in the Netherlands.

Acknowledgements

We are indebted to Anne Cutler, who placed the MPI picture database at our disposal. Renée van Bezooijen and Nanna Haug Hilton provided valuable advice on the analysis, and three anonymous reviewers gave insightful comments on an earlier version of this paper. Their advice helped to improve this paper considerably.

References


Der, G. & Ian J. Deary. 2006. Age and sex differences in reaction time in adulthood: Results from the United Kingdom health and lifestyle survey, Psychology and Aging, 21, 1, 62-73.


Gooskens, C. & G. Doetjes. 2009. Skriftprogets rolle i den dansk-svenske talespråksforståelse [The role of orthography in spoken language comprehension between Danish and Swedish], Språk och stil.


4 Investigating the link between intelligibility and language attitudes using the matched-guise technique

This paper has been submitted as Anja Schüppert, Nanna Haug Hilton & Charlotte Gooskens: “Swedish is beautiful, Danish is ugly. Investigating the role of intelligibility for language attitudes”.

Abstract

This paper aims at investigating the hypothesis that attitudes towards a linguistic variety and intelligibility of that variety are linked. This is done by eliciting language attitudes and word recognition scores in 154 Danish and Swedish school children between 7 and 16 years. Language attitudes towards the neighbouring language are elicited by means of a matched-guise experiment while word recognition is tested by auditorily presenting the children to 50 spoken items in their neighbouring language (Danish for Swedish children and vice versa) in a picture-pointing task. Results revealed that while Danish children held more positive attitudes towards Swedish than vice versa and their word recognition scores generally are higher than those of their Swedish peers, the correlation between these two variables is very low, indicating that other factors need to be taken into account in order to explain the well-established asymmetry in mutual intelligibility between spoken Danish and spoken Swedish.

4.1. Introduction

Most people have strong ideas about whether a language variety sounds pleasant or not. Attempts to explain these language attitudes have been done either with a focus on the intrinsic value of linguistic features of the target language (inherent value hypothesis, Giles et al., 1974a; Giles et al., 1974b), or with a focus on cultural norms (imposed norm hypothesis, Giles et al., 1974a; Giles et al., 1974b). The inherent value hypothesis argues that some languages are evaluated more positively than others because they are inherently more correct, more logical or more aesthetically pleasing than others, while the imposed norm hypothesis argues that no language variety is inherently beautiful or ugly, and that such judgments are based solely on non-linguistic factors such as stereotypical ideas that are adopted by someone without critical evaluation, e.g. the notion that French is a romantic language or Dutch sounds harsh. As pointed out by Van Bezooijen (2002), Trudgill and Giles (1978) extended the imposed norm hypothesis to include social norms as well (social connotations hypothesis). In their view, language attitudes can also be based upon individual experiences which form individual social connotations, e.g. if a person has had a negative encounter with a speaker of Italian, this person’s attitude towards the Italian language might be more negative than after encountering a neutral or positive speaker of Italian. Yet another hypothesis about the formation of language attitudes is put forward by Boets and De Schutter (1977), who presented Belgian Dutch regional dialects to listeners from a geographically centrally situated village in
Belgium. They report that there is a link between intelligibility of a dialect and its aesthetic pleasantness in such a way that dialects that are judged as beautiful are more intelligible to the judges, and vice versa. Deprez and De Schutter (1980) and Van Bezooijen (1994) report findings that are in line with these results.

Many investigations testing the validity of these hypotheses report results supporting the social connotations hypothesis. Giles et al. (1974a) report no systematical differences in attitudes in their study of English informants judging several (both standard and non-standard) varieties of Greek. Giles et al. (1974b) report similar results for varieties of French among listeners from Wales. These findings, which indicate that listeners with no previous exposure to regional varieties of a language do not make meaningful evaluative distinctions between them, were interpreted as evidence against the inherent value hypothesis. Investigating classification of US, British and New Zealand English accents, Milroy and McClenaghan (1977) and Ladegaard (1998), on the other hand, failed to confirm conclusions from the above studies. In his experiment, Ladegaard (1998) presented speech samples from five male native speakers of different regional varieties of English (US American, Scottish, Australian, Cockney and Received Pronunciation) to 96 Danish judges, who were instructed to identify the speakers’ geographic background and to judge the speakers with regard to status, competence, personal integrity, social attractiveness and linguistic attractiveness. Ladegaard (1998) reports that a large number of participants judged the accents heard in the same manner as native speakers did (e.g. Australian English as laid-back or Cockney as having a low social status) despite the fact that the majority was not able to identify the variety’s geographical origin. Importantly, their classification reflected common stereotypes of the accents and its speakers as held in the English speaking community. Ladegaard (1998) interprets this finding as negative evidence for the social connotations hypothesis. Van Bezooijen (1994) investigated language attitudes towards four different varieties of Dutch (standard, urban, and two rural dialects from the Netherlands and Belgium) among Dutch school-children and adults. She reports evidence in conflict with the social connotations hypothesis, as she found that listeners have a tendency to evaluate the rural dialects as less attractive than the urban variety, while the social connotations hypothesis formulated to account for the aesthetic evaluation of English varieties predicts the reverse – namely the most negative attitudes towards the urban variety.

To summarise, within sociolinguistics, it has been proposed that language attitudes are based on (a) linguistic features of the target language, (b) stereotypical ideas about the target language held by a group of people, (c) personal experience, or (d) the intelligibility of the target language to the speaker. One of these hypotheses, the intelligibility hypothesis, states that the intelligibility of a linguistic variety is linked significantly to the listener’s attitude. Interestingly, within intelligibility research, a similar hypothesis has been put forward. It is similar in the way that it supposes a link between the attitudes held towards a linguistic variety and how intelligible the variety is to the listener. The hypothesis is different, however, in the
causal relationship it assumes. While the intelligibility hypothesis put forward by sociolinguistic researchers assumes that specific linguistic features of a language variety influence the attitude towards that variety, researchers within intelligibility research propose that attitudes held towards a specific language variety influence the effort the listeners make for decoding that variety.

Wolff (1959) investigated mutual intelligibility between the closely related Nigerian Ijo languages Kalabari and Nembe, and reports that Nembe speakers claim to understand Kalabari, while speakers of Kalabari judge Nembe to be unintelligible to them. Wolff (1959) suggests that this asymmetry in mutual intelligibility is linked to an asymmetry in language attitudes. He states that when his study was conducted, the Kalabari were the most prosperous group in the Eastern Niger Delta and that they regarded other Ijo speaking groups as inferior to them. Whether it is the case that the Kalabari actually have difficulties understanding their neighbouring languages, or they merely claim to have so for prestige reasons, is irrelevant according to Wolff (1959). He suggests that the intelligibility evidence simply underscores the Kalabari ascendancy and thereby links together the two factors language attitude and intelligibility. His assumptions, however, are not based on empirical data, but on anecdotal evidence solely. Recently, in their literature review of language attitudes written for the scientifically interested layman, Giles and Niedzielski (1998:87) pick up Boets and de Schutter’s (1977) hypothesis and argue that an impaired intelligibility of a specific variety can cause a negative attitude towards the variety in question. Boets and De Schutter’s (1977), Deprez and De Schutter’s (1980), Wolff’s (1959) and Giles and Niedzielski’s (1998) hypotheses are similar in that they assume a relationship between the intelligibility of a language and the attitudes that are held towards that language, but as noted above they are different in the causal relationship they assume, however.

This paper aims at re-investigating the intelligibility hypothesis. Specifically, we are interested in finding out whether ‘asymmetric’ attitudes held by Scandinavians towards neighbouring, closely related Scandinavian languages are linked to ‘asymmetric’ intelligibility of these neighbouring languages. This link has been suggested by several researchers, whose studies are summarised in the following section.

4.2. Inter-Nordic communication

The Nordic countries Denmark, Finland, Iceland, Norway and Sweden as well as their associated territories Åland, the Faroe Islands and Greenland share some important historic events and contemporary cultural and political norms. Their political and economic systems are characterised by generous welfare criteria and among other things emphasise gender equality, egalitarian benefit levels and economic systems based on keynesianism. The Nordic countries co-operate in the Nordic Council since 1952 and the Nordic Council of Ministers since 1972. Both authorities strongly promote inter-Nordic collaboration, e.g. by implementing the Nordic passport union.
in 1954 which allows Nordic citizens to reside in any of the Nordic countries without a valid passport, by creating an inter-Nordic job exchange platform (Nordjobb, founded in 1985) and by emphasising the ideological role of using Nordic languages in inter-Nordic communication situations rather than a *lingua franca* such as English. Among other incentives, this was secured by the *Språkkonvention* (‘language convention’) that ensured that citizens of the Nordic countries are entitled to use their native language in written communication with authorities. In this case, authorities adopt the language of the client. Another communication pattern which also eliminates the usage of a *lingua franca* such as English is communication in the native language of the speaker. This is possible in language communities of closely related languages, and the closer the language varieties involved are to each other, the more effortless the communication works. An example would be a Dane speaking Danish to a Norwegian, who then replies in Norwegian. Haugen (1953) called this type of communication *semi-communication*. Danes, Norwegians and Swedes especially are likely to use their native language when communicating with each other and mutual intelligibility of Scandinavian language varieties has been a focus of linguistic research in Scandinavia over the past decades. Some of the investigations are summarised below.

Haugen (1953) investigated mutual intelligibility between the Nordic languages by asking members of *Föreningen Norden* (‘The Nordic Society’) how much of the neighbouring language they understood. *Föreningen Norden* is a non-governmental organisation which promotes cooperation between the Nordic countries. Haugen (1953) found that Danes reported to have slightly more difficulties understanding spoken Swedish than vice versa. He also elicited language attitudes from his participants and found that 42% of the Danish participants thought that Swedish sounded more beautiful than their own language, while none of the Swedish participants thought that Danish sounded more beautiful than their own language. His data thus does not provide support in favour of the hypothesis that language attitude and intelligibility are linked, as the group that held a more positive attitude towards the neighbouring language turned out to self-report that they understand less than the group that held a more negative attitude.

In contrast to Haugen (1953), who based his study on the participants’ self-assessment of their comprehension abilities, Maurud (1976) investigated mutual intelligibility between Danish and Swedish by testing the participants’ performance in a translation task and a multiple choice test. He reports that, while Danes understand on average 60% of spoken Swedish, Swedes only understand 48% of spoken Danish. That means, in contrast to Haugen (1953), Maurud (1976) finds a trend that Danes comprehend more spoken Swedish than vice versa. One of the major criticisms of Maurud’s (1976) investigation (Gregersen 2004), however, has been the fact that he compares comprehension of Swedish among Danes in Copenhagen to comprehension of Danish among Swedes in Stockholm. This means a substantial geographical asymmetry in the data as Copenhagen is located only 30 kilometres from the Swedish border, while Stockholm is located about 570 kilometres
from the Danish border. It can be assumed that people living in border regions have a higher amount of cross-border contact, such as travelling to the neighbouring country, talking to people visiting their own country, or even watching television in the neighbouring language. Maurud (1976), however, does not mention this geographic asymmetry as a factor that might cause or boost the comprehension asymmetry he reports, but rather hypothesises that attitudes towards the neighbouring languages are of major importance for Scandinavians’ ability to communicate with each other in their native languages, thereby indirectly subscribing to Wolff’s (1959) view.

In a large-scale investigation, Delsing and Lundin Åkesson (2005) elicited text comprehension as well as language attitudes among different groups of participants from the Nordic countries. These groups hailed from at least two different sites per country, except for Finland (three sites) and Åland, Greenland and the Faroe Islands (each one site). Danes were tested in Århus (340 km from Sweden via land route and 170 km via sea route) and Copenhagen, while Swedes were tested in Malmö (40 km from Danish mainland) and Stockholm. Thereby, the geographic asymmetry was somewhat neutralised, although Stockholm is still roughly two to three times as far from Denmark as Århus is from Sweden. Delsing and Lundin Åkesson (2005) confirmed Maurud’s (1976) finding that Danes understand more spoken Swedish than vice versa. They also report that Danes rate the Swedish language as more beautiful than vice versa. Delsing and Lundin Åkesson (2005) correlated linguistic performance and attitudes held towards the neighbouring language and the neighbouring country. They report that Danes’ comprehension abilities correlate significantly with their judgment of how beautiful the Swedish language sounds and that Swedes’ comprehension abilities correlate significantly with their willingness to move to Denmark, while neither Danes’ comprehension of Swedish and their willingness to move to Sweden correlated, nor Swedes’ comprehension of Danish and their judgment of the beauty of the Danish language. Unfortunately, no correlation coefficients were reported in their study, which makes it difficult to assess the degree of correlation. For the first time, however, empirical evidence was presented supporting the assumption that intelligibility and language attitudes are linked within the Scandinavian language area – although the nature of this link is still unclear. It is possible that listeners holding positive attitudes make a greater effort to understand the language in question than those holding negative attitudes, which is what Wolff (1959) suggests. It might also be the case, however, that those participants who understand the language better, simply perceive the language as being more beautiful because their comprehension makes them feel as part of the speech community and facilitates a development of positive feelings towards a said variety. The latter causality is in line with Giles and Niedzielski’s (1998) hypothesis.

In a recent study, Schüppert and Gooskens (in press) investigated attitudes towards the neighbouring language among Danish and Swedish 4-to-6 year old preschoolers as well as among adolescents aged between 17 and 20, all hailing from towns that were located at about 200 km from the Danish-Swedish border. They
reported that preschoolers held neutral and, importantly, symmetric attitudes towards the neighbouring language, while the Swedish language was rated as significantly more pleasant by Danish adolescents than the Danish language was rated by their Swedish peers. This suggests that the asymmetry in attitude towards the neighbouring language starts to develop in the age range of about 6 to 18 years. Schüppert and Gooskens (in press) also investigated auditory word recognition in a picture-pointing task in their participants and analysed reaction times to correctly recognised items. While Danish preschoolers recognised Swedish items as quickly as Swedish preschoolers recognised Danish items, the by now well-established asymmetry in comprehension scores was found among adolescents. Swedish adolescents had a significantly longer reaction time than Danish adolescents had. As it is generally assumed that the time it takes a participant to make a decision reflects the processing time and thereby the degree of complexity of the task (Gass and Mackey 2007: 22ff), these results indicate that Swedish adolescent participants had more difficulties in decoding the Danish stimuli than Danish participants had with the Swedish stimuli. Importantly, however, no significant correlation between a subject's mean reaction time of correctly translated items and his or her attitude towards the neighbouring language could be found. These findings are in conflict with Delsing and Lundin Åkesson (2005) who found a clear difference in attitudes between adolescents in the two countries. Linguistic factors were instead identified as the cause to the asymmetry in mutual intelligibility, i.e. an asymmetric speaking rate (Hilton, Schüppert and Gooskens submitted), less distinct vowel articulation (Vanhove et al.2010) or an asymmetric number of reduction processes (Bleses et al. 2008: 624). There are also indications that the conservative Danish orthography, which reflects proto-Nordic pronunciation in a more accurate way than the more modern Swedish orthography has, might work as an extra cue for Danish listeners when confronted with spoken Swedish (Gooskens & Doetjes, 2009; Schüppert et al. submitted).

The studies of language attitudes reported above (Haugen 1953; Maurud 1976; Delsing and Lundin Åkesson 2005; Schüppert and Gooskens in press) are all characterised by the shortcoming that there was not enough stimuli control in the test situations. The participants in the studies might have referred to different speech samples (from different speakers with different voices in different settings etc.) when making evaluative judgements about a language. For example, in Schüppert and Gooskens’ (in press) study, the Danish participants were confronted with a native speaker of Swedish and the Swedish participants were confronted with a native speaker of Danish in the word recognition experiment. After the experiment, they were asked whether they liked the language they had heard (a) more than their native language, (b) as much as their native language or (c) less than their native language. It is possible that the Swedish speaker accidentally happened to have more voice features that are generally judged as being more beautiful by native as well as non-native speakers of Swedish (such as a more variable intonation) than other speakers of Swedish have. This would result in low generalisability of the data.
The aim of this paper is to test whether intelligibility of a closely related language and the language attitudes held towards it correlate. This is done by linking mutual intelligibility between the two languages Danish and Swedish to Danes’ and Swedes’ attitudes towards their neighbouring language, respectively. To address the shortcoming of the studies previously conducted language attitudes of Danish and Swedish-speaking children and adolescents are elicited in a way that ensures that voice quality is kept constant across the two language samples. This is done by using the matched-guise technique (see Section 4.3.1), which, to our knowledge, had not been used hitherto to investigate language attitudes between Danish and Swedish-speaking listeners. After the attitude elicitation, the participants are tested on word recognition of the neighbouring language. By correlating individual attitudes with individual intelligibility scores, the intelligibility hypothesis, which states that a person’s attitude towards a language variety is linked to the intelligibility of that language variety, is investigated.

### 4.3. Method

#### 4.3.1. Matched-guise experiment

Within sociolinguistics research, it has been attempted to minimise biases due to differences in speech quality in voice evaluation tasks by employing the matched-guise technique (Lambert et al., 1960). This technique uses speech samples that are matched with regard to speech features. This is done primarily by using two speech samples from the same speaker. For instance, Lambert et al. (1960) used the technique to investigate stereotypical prejudices about English- and French-speaking Canadians held by people in bilingual Quebec. They (Lambert et al. 1960) investigated less consciously held language attitudes by instructing participants to rate English and French speakers with regard to personal speaker attributes, such as kindness, richness or beauty. All stimuli were produced by the same speaker, but this fact was not made clear to the participants, who believed that they judged different speakers - an Anglophone speaker and a francophone speaker. Lambert et al. (1960) found that the participants’ judgments of personality traits of the bilingual speaker were strongly influenced by the language that was spoken. Both English and French-speaking participants rated English more positively on status and solidarity traits, which is assumed to reflect the English language’s higher status in Quebec.

#### 4.3.1.1. Stimulus material

The auditory stimulus material for the matched-guise experiment was a short text, selected from the children’s book *Can’t You Sleep, Little Bear?* (Waddell and Firth 2005) and consisting of six sentences. Besides recordings of Danish and Swedish, which were made by one and the same speaker, recordings were also made by speakers of four other languages, namely Norwegian, Dutch, Frisian and Indonesian, that served as distracter stimuli. In total, the stimulus material thus comprised six different audio fragments representing six different languages. The six texts were presented to the Danish and Swedish participants in the following order: Norwegian,
Dutch, native language (Danish or Swedish, respectively), Frisian, Indonesian, neighbouring language (Swedish or Danish, respectively). In other words, the order in which the languages were presented to the children was adapted to the country in which the experiment was held, but the two guises which where the focus of the present study were always separated by two distracters.

4.3.1.2. Bilingual speaker
The Danish and the Swedish texts were produced by the same speaker: a young female Dane who had grown up in Southern Sweden but consistently spoke Danish with her Danish parents and siblings at home. A crucial factor in using the matched-guise technique is that reactions are attributable to the language itself. Therefore, much care was taken to ensure that the bilingual speaker sounded natively Danish and Swedish. This was done by organizing two so-called voice parades, which investigated whether the bilingual speaker sounded as native to listeners with both language backgrounds as other native speakers of the two languages did. It involved presenting native listeners (none of which participated in the matched-guise experiment) with a number of recordings of native speakers, including one by the bilingual, and instructing them to pick out one speaker that sounded non-native. We assumed that if the bilingual speaker is not chosen as the foreigner more often than on chance level, he or she sounds sufficiently native for our purpose. Two voice-parades were conducted, a Danish and a Swedish one. Five recordings were presented to 30 Danish and 15 Swedish listeners. For the Danish version, the four other recordings were produced by native female Danish speakers from the greater Copenhagen area, the same geographical area that the bilingual hailed from. The distracter recordings in the Swedish version were all recordings of female speakers from Southern Sweden. In both voice parades, the bilingual speaker was presented as the third speaker of five. The results of the tests are shown in Table 1, which demonstrates that the bilingual speaker was not judged as sounding less native than the distracter recordings. In the Swedish voice parade, the bilingual speaker was selected by none of the listeners as having a foreign accent; in the Danish voice parade she was chosen by 10% of the listeners, which is still clearly below chance level. Table 1 demonstrates that the recordings of the bilingual speaker were not rated significantly less native sounding than the other recordings by neither Danish or Swedish listeners. This suggests that both guises recorded for the experiment are perceived as native Danish and Swedish.

4.3.1.3. Procedure
The children were provided with separate rating questionnaires for every language consisting of six 5-point Semantic Differential Scales (Osgood 1957). Figure 1 shows the semantic differential scales employed. Semantic Differential Scales are similar to Likert scales in that several items can be used to evaluate the same target. An advantage of this technique is that, unlike with Likert scales, no explicit statements have to be formulated by the researcher, such as “The speaker sounds intelligent”. This way writing statements can be avoided; instead, respondents are asked to
indicate their answers by marking a scale between two bipolar adjectives as extreme values. Thus, the principle of Semantic Differential Scales is based upon the idea that most adjectives have logical opposites. Even if opposing adjectives are not obviously available, in Germanic languages they can easily be constructed by putting “not” or “un-” and their language specific equivalents in front of the original adjective (Dörnyei 2010: 30).

Table 1. Results of the voice parade for the Danish-Swedish bilingual speaker. Grey shaded cells indicate speakers that were picked at or above chance level. Note that each distracter represents two different speakers, i.e. a Danish and a Swedish one.

<table>
<thead>
<tr>
<th>Bilingual speaker</th>
<th>Distracter 1</th>
<th>Distracter 2</th>
<th>Distracter 3</th>
<th>Distracter 4</th>
<th>Chance</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danish</td>
<td>N 3 % 10 % 5</td>
<td>N 0 % 17</td>
<td>N 13 % 43</td>
<td>N 9 % 30</td>
<td>20 % 20</td>
<td>30</td>
</tr>
<tr>
<td>Swedish</td>
<td>N 0 % 0</td>
<td>N 3 % 20</td>
<td>N 12 % 80</td>
<td>N 0 % 0</td>
<td>20 % 20</td>
<td>15</td>
</tr>
</tbody>
</table>

Before the experiment started, every participant was familiarised with using semantic differential scales by listening to an oral introduction. This introduction was given by an experimenter living in the same region where the experiment was run, i.e. a Danish experimenter supervised the conduction of the experiment in Denmark and a Swedish experimenter did so in Sweden. The instructions were given as suggested by Dörnyei (2010: 31; 75pp.), which call for the inclusion of a friendly, respectable and involved local during all testing sessions. It was also ensured that detailed oral instructions were given and that the listeners had the opportunity to ask questions prior to the test. The children were instructed to judge six personality traits of the speaker on a five-point Semantic Differential Scale (see Figure 1). These six personality traits were chosen in such a way that three different dimensions were represented by two traits each. These three dimensions, namely dynamism (strange/normal, old-fashioned/modern), attractiveness (ugly/beautiful unkind/kind) and superiority (stupid/smart, poor/rich), are regarded as representative for eliciting attitudes (Zahn and Hopper 1985). Each speaker should be judged while the fragment was played to the participants.

Figure 1. Semantic Differential Scale that were provided for every text fragment.

How do you think the speaker sounds?

- strange  o  o  o  o  o  normal
- ugly  o  o  o  o  o  beautiful
- dumb  o  o  o  o  o  smart
- old-fashioned  o  o  o  o  o  intelligent
- unfriendly  o  o  o  o  o  friendly
- poor  o  o  o  o  o  rich
As can be seen in Figure 1, we chose to present our participants consistently with “positive” adjectives on the right-hand side of the sheet and their negative equivalents on the left. This has advantages and disadvantages. On the one hand, it is likely to produce a halo effect, i.e. a higher correlation between judgments on the variables strange/normal, old-fashioned/modern, ugly/beautiful, unkind/kind, stupid/smart and poor/rich than would occur if the positive adjectives were presented counterbalanced between left and right. On the other hand, by keeping positive adjectives consistently on the same side, it is easier for the participants to fill in the scales. As we test rather young subjects, we opted against a counterbalanced presentation to make the task as easy as possible for the participating children and accept the consequences of the halo effect. After having judged all six audio sequences, the children were instructed to provide some personal information such as date of birth, native language(s), how often they had been to the neighbouring country, and how often they had heard the neighbouring language via television. An English translation of the questionnaire can be found in the Appendix. All children remained anonymous. The experiment was conducted individually for the youngest (i.e. first grade) children, and with all pupils at once in all other grades. There was no indication that any participant in the matched-guise experiment became aware of the fact that they heard the same speaker twice, which is crucial for the experiment’s validity.

4.3.1.4. Participants
In total, 159 children participated in the matched-guise experiment, five of whom had to be excluded of further analysis for different reasons. One Swedish child had only lived in Sweden for a year, and therefore did not speak Swedish at the level of a native speaker of the same age; two Swedish children did not complete the questionnaire; one Danish child produced unreliable answers and indicated for example that she had heard Portuguese while the text was presented in her native language (i.e. Danish); and similarly, one Danish child claimed he did not recognise his own language. After excluding these five children, 154 children were left for further analysis of the matched-guise experiment.

Table 2. Mean age and number of children that participated in the matched-guise experiment per grade and language.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Danish</th>
<th></th>
<th>Swedish</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>Mean age</td>
<td>$N$</td>
<td>Mean age</td>
</tr>
<tr>
<td>1st grade</td>
<td>10</td>
<td>8.1</td>
<td>11</td>
<td>7.8</td>
</tr>
<tr>
<td>3rd grade</td>
<td>22</td>
<td>10.1</td>
<td>15</td>
<td>9.8</td>
</tr>
<tr>
<td>5th grade</td>
<td>20</td>
<td>11.8</td>
<td>14</td>
<td>11.7</td>
</tr>
<tr>
<td>7th grade</td>
<td>22</td>
<td>13.8</td>
<td>15</td>
<td>14.1</td>
</tr>
<tr>
<td>9th grade</td>
<td>12</td>
<td>15.9</td>
<td>13</td>
<td>15.6</td>
</tr>
<tr>
<td>total</td>
<td>86</td>
<td>12.0</td>
<td>68</td>
<td>11.9</td>
</tr>
</tbody>
</table>

The Danish children ($N = 86$) were 12.0 years on average, while the Swedish children ($N = 68$) had a mean age of 11.9 years. We elicited data from children aged between 7
and 16, thereby filling the gap that Schüppert and Gooskens’ (in press) study left. The participants attended five grades: 1st grade, 3rd grade, 5th grade, 7th grade and 9th grade. Mean age and number of children per grade and language are indicated in Table 2.

All children were tested during school time. The schools were financially compensated for the time that they allocated to the experiment. The experiment was conducted in Odense Municipality (Denmark) and Kronoberg County (Sweden), both at approximately 200 km from the Danish-Swedish border.

4.3.2. Word recognition experiment

4.3.2.1. Stimulus material

The stimulus material for the word recognition experiment consisted of 50 auditory nouns and was derived in the following way. In a pre-experiment, 112 pictures, selected from the picture database developed at the Max-Planck-Institute for Psycholinguistics, were shown to five four-year-old Danish and five four-year-old Swedish children. None of these children participated in the word recognition experiment or the matched-guise experiment. The children were asked to label the depicted objects as spontaneously as possible. Then, a labelling consistency was calculated per picture for the most frequent label. For example, the object depicted in Figure 2 was labelled *kænguru* ‘kangaroo’ by four Danish children and *känguru* by two Swedish children. One Danish child labelled it *mus* ‘mouse’ and the remaining three Swedish children labelled it *struts* ‘ostrich’, *hare* ‘hare’ and *mammahare* ‘mommy hare’. That means that the Danish labelling consistency for the most frequent label *kænguru* was 80 percent, while the Swedish labelling consistency was 40%.

Figure 2. Example of one picture employed as target picture in the word recognition experiment.

To be selected for the experiment, a picture had to fulfil two criteria: It had to have a labelling consistency of at least 80 percent in both languages, i.e. it had to be given the same label by at least four of five children (*intra-language criterion*), and it had to be labelled with cognate words (*inter-language criterion*). Cognate words are words that share their etymology, such as Danish *hoved* and Swedish *huvud* (‘head’). By selecting target pictures (i.e. the pictures that corresponded to the auditory stimuli in the actual experiment) on the basis of these two criteria, it was ensured that they were recognised and produced by children even younger than the age group tested in the experiment. This procedure seemed most effective to ensure that all auditory stimuli presented in the experiment were highly frequent, as there were no frequency lists available that represent word frequencies in Danish and Swedish children.
Furthermore, this procedure ensured that all pictures were unambiguous to the children and could be identified easily if the auditory stimulus was intelligible to them. The example picture from Figure 2 met the inter-language criterion, but not the intra-language criterion and was therefore rejected. 53 pictures met these two criteria. Their labels were used as auditory stimuli, which were recorded by two female native speakers: a Danish and a Swedish speaker. Recordings took place in sound-attenuated rooms and the sound files were digitised at 44100 Hz and downsampling to 22050 Hz. The 59 pictures which had not met both criteria were supplemented with further 100 pictures from the same database and served as distracter pictures for the word recognition experiment. Three of the tokens were used in a demo version, leaving 50 stimuli for the experiment.

4.3.2.2. Participants
Of the 154 children that participated in the matched-guise experiment, 116 children (54 Danish and 64 Swedish) participated in the word recognition experiment. Mean age and number of participants per grade and L1 are shown in Table 3. Apart from the Danish 7-graders, who consisted of 9 subjects, all groups contained at least 10 participants.

4.3.2.3. Procedure
The intelligibility experiment was conducted after the matched-guise experiment. All children were tested individually and were presented with the 50 auditory stimuli in a picture-pointing task. The children were seated in front of a touch screen (LG L1510SF). Before the experiment started, the children were familiarised with the task through a short training session. The demo version of the experiment consisted of two trials with stimuli in the children’s native language, followed by one trial with a stimulus in the test language. Four pictures per stimulus were presented on the touch screen and remained on the screen until the participants touched the screen or for 10 000 ms. The children were instructed to point to the picture that best corresponded to the stimulus. Before the word-recognition experiment started, it was ensured that the children had understood the task, and, if necessary, further instructions were given. The auditory presentation of the stimuli was random, but every stimulus was presented together with the same set of four pictures across sessions and across languages.

Table 3. Mean age and number of children that participated in the intelligibility experiment per grade and language.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Danish</th>
<th>Swedish</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean age</td>
</tr>
<tr>
<td>1st grade</td>
<td>10</td>
<td>8.1</td>
</tr>
<tr>
<td>3rd grade</td>
<td>12</td>
<td>10.1</td>
</tr>
<tr>
<td>5th grade</td>
<td>10</td>
<td>11.9</td>
</tr>
<tr>
<td>7th grade</td>
<td>9</td>
<td>13.6</td>
</tr>
<tr>
<td>9th grade</td>
<td>12</td>
<td>15.9</td>
</tr>
<tr>
<td>total</td>
<td>53</td>
<td>11.9</td>
</tr>
</tbody>
</table>
4.4. Results

4.4.1. Matched-guise experiment

For all six fragments that were presented to them, the children indicated on a 5-point Semantic Differential Scale how normal, beautiful, smart, modern, kind and rich they think the speaker sounded. We coded the data by assigning the lowest score (1) for strange, ugly, stupid, old-fashioned, unkind and poor, the highest score (5) to normal, beautiful, smart, modern, kind and rich, and the remaining scores for answers given for any of the points between the extremes. Figure 3 shows the mean ratings of the six personality traits of the bilingual speaker when she spoke the neighbouring language (left panel) and the native language (right panel).

Figure 3. Mean judgments of the bilingual speaker when speaking the neighbouring language (left graph) and when speaking the listeners’ native language (right graph).

If the left and the right panels in Figure 3 are compared, it can be seen that the children judge the bilingual speaker more positively when she speaks the children’s native language (right panel) than when she speaks the neighbouring language. This is confirmed by six pairwise $t$-tests between ratings of the bilingual speaker when she speaks the native and the neighbouring language (all $p < .001$). Furthermore, six independent $t$-tests revealed that ratings of the bilingual speaker when she speaks the native language of the children were similar in the L1 groups (all $p > .1$). For three of the personality traits, however, namely normality, beauty and kindness, the ratings of the bilingual speaker when she spoke the neighbouring language differed significantly (all three $p < .01$) across the two L1 groups. As indicated by Figure 3 (left panel), Danish participants rate the bilingual speaker consistently as being kinder-sounding, more normal-sounding, and more beautiful-sounding when she speaks Swedish, than Swedish participants do when she speaks Danish.

Figure 4 shows mean ratings of all six personality traits of the speaker when she speaks the neighbouring language as a function of age for both L1 groups. It can be seen that there is a trend that children rate the speaker more negatively the older they are. Often, but not always, the highest mean is reached by the youngest group per L1,
and the lowest mean is found in the oldest group per L1. Interestingly, Swedes and Danes rate the speaker in a similar manner when they judge richness and modernity, but rather differently when they judge kindness, smartness, normality and beauty.

**Figure 4.** Mean personality trait ratings from Danish (dotted line) and Swedish children (solid line) of the bilingual speaker speaking the neighbouring language as a function of age.

Another trend we can detect is that the speaker is rated more positively by the Danes when she speaks Swedish than by the Swedes when she speaks Danish. This trend is particularly pronounced in the older groups, while the youngest group (7 to 8 year old children) often give similar judgments, especially for richness, kindness and smartness.

Finally, the Swedish-speaking group of 9 to 10 year old children seems to behave differently than would be expected from the overall trend in their L1. They rate the speaker almost as negatively as the oldest group for two of the personality
traits (smartness and normality), and even more negatively than the oldest group for one trait (kindness).

To test the detected overall trends that (1) younger children give more positive judgments than older children, (2) Danes rate the bilingual speaker speaking the neighbouring language more positively than the Swedish speakers do and (3) younger children tend to give more similar ratings in both groups of L1 than older groups do, and to evaluate our hypothesis that language attitude correlates positively with intelligibility, we reduced the data by conducting a principal component analysis (PCA) on the overall ratings on the six personality traits normality, beauty, smartness, modernity, kindness and richness, which served as input variables. It revealed that most of the six variables were significantly interrelated, but correlation coefficients never exceeded $r = .55$. The only variable that only correlates significantly with three of the remaining five variables (namely normality, beauty and richness) is modernity. This suggests that the five variables normality, beauty, smartness, kindness and richness measure the same phenomenon without entirely consisting of redundant information, which would be the case if the variables would correlate too highly.²

Two principal components have an eigenvalue of more than 1 and were extracted for further analysis. The first component has an eigenvalue of 2.64 and correlates highly (all $r \geq .70$) with the ratings for beauty, smartness, kindness and normality, medium with richness ($r = .57$) but low ($r = .34$) with modernity. This component therefore represents most personality traits well and seems to measure “overall attractiveness”. The second component has an eigenvalue of 1.05 and correlates highly with old-fashioned/modern but less highly with the remaining five personality traits (see Table 4). This component seems to measure mainly “modernity”.

Table 4. Component matrix with correlation coefficients between the two extracted principal components “attractiveness” and “modernity” and the ratings of the bilingual speaker when she spoke the neighbouring language with respect to the six personality traits.

<table>
<thead>
<tr>
<th></th>
<th>Attractiveness</th>
<th>Modernity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ugly/beautiful</td>
<td>0.80</td>
<td>0.13</td>
</tr>
<tr>
<td>Stupid/smart</td>
<td>0.74</td>
<td>-0.33</td>
</tr>
<tr>
<td>Unkind/kind</td>
<td>0.72</td>
<td>-0.39</td>
</tr>
<tr>
<td>Strange/normal</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>Poor/rich</td>
<td>0.57</td>
<td>0.28</td>
</tr>
<tr>
<td>Old-fashioned/modern</td>
<td>0.34</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Figure 5 displays the six eigenvalues for the maximum number of extractable components as a function of these six components (components 1 to 6). It can be seen that the “elbow” of the graph is located at the second extracted component. This

² This is confirmed by the fact that, for all six variables together, Bartlett’s test of sphericity resulted in $\chi^2(10) = 198.57 \ (p < .001)$, and that the Kaiser-Meyer-Olkin measure of sampling adequacy was .75.
suggests that this component could be excluded from further components (Field 2005), as should components 3 to 6, and that extraction of component 1 is a meaningful way to reduce the data (Hatcher 1994). However, as the second component has an eigenvalue of slightly more than 1, we opted for the extraction of the two first components. Together, both extracted components explain 61.6% percent of the variance.

Figure 5. Scree plot showing eigenvalues per component.

Now data from six variables are reduced to two components representing “attractiveness” and “modernity”. Both components consist of standardised values (z-scores), which means that the mean value for all 154 participants is 0 and the standard deviation for all participants is 1. These components form the basis of the remaining analyses and represent the ratings of the bilingual speaker when she speaks the neighbouring language with regard to six personality traits. Figure 6 shows an error bar plot of both components for the L1 groups of participants (Danish and Swedish). Circles represent the mean value while error bars represent the 95%-confidence interval. In other words, 2.5% of the values fall above the upper part of each error bar and 2.5% of the values are found below the lower part of each error bar. The Danish mean value given for the attractiveness of the bilingual speaker when she speaks Swedish is $\bar{z} = 0.2$, the Danish mean value for modernity is $\bar{z} = -0.1$. The Swedish mean value given for the attractiveness of the bilingual speaker when she speaks Danish is $\bar{z} = -0.3$, the Swedish mean value for modernity is $\bar{z} = 0.1$. In other words, Danes find the bilingual speaker attractive but old-fashioned when she speaks Swedish, while Swedes find her modern but unattractive when she speaks Danish.

An independent $t$-test on both components revealed that the difference between Danish and Swedish judgments of the speaker’s attractiveness are significantly different $t(149) = 3.3, p = .001$, two-tailed), while the bilingual speakers’ modernity was not judged significantly different across the two L1 groups. That means that Danish children judge the bilingual speaker as more attractive when she speaks Swedish than Swedish children judge the same speaker when she speaks Danish. This result confirms findings by Delsing and Lundin Åkesson (2005) and Schüppert and Gooskens (in press), who report that language attitudes were significantly more positive in Danish participants than in Swedish participants.
Figure 6. Error bars (95% CI) of the extracted principal component (representing the ratings of the bilingual speaker when she speaks the neighbouring language with regard to five personality traits) broken down by L1.

Figure 7 shows line diagrams of the extracted components “attractiveness” and “modernity” of the bilingual speaker speaking the neighbouring language as a function of age for both L1s. It can be seen that both Danish and Swedish children tend to judge the speaker as sounding less attractive when she speaks the neighbouring language as they get older. Indeed, the variables “age” and “attractiveness” correlate to a low degree but significantly negatively ($r = -.23, p = .004$), indicating that, generally, the attractiveness score decreases with increasing age. The extremely negative judgments of the Swedish-speaking group of 9 to 10 year old children are reflected very clearly in this graph. While they are the second youngest group, their ratings are almost as negative as the ratings from the oldest Swedish-speaking group. There are no indications as to why they behave the way they do, however. The ratings regarding the modernity of the bilingual speaker when she speaks the neighbouring language do not change with age, as this variable does not correlate significantly with attended grade. Before correlating our extracted component “attractiveness” with intelligibility scores, the intelligibility scores are investigated more closely.

Figure 7. Mean values of the extracted component “attractiveness” (left panel) and “modernity” (right panel) as a function of age for Danish (dotted line) and Swedish (solid line) participants.
4.4.2. Word recognition experiment

Of the 154 children whose data was analysed in the matched-guise experiment, 116 participated in the word recognition experiment. As the experiment was designed in such a way that all groups of participants could conduct it successfully, the older children (from grade 5 onwards) correctly identified more than 90% of the stimuli and thus performed near ceiling. The already well-documented asymmetry between Danish and Swedish-speaking listeners (Maurud 1976; Delsing and Lundin Åkesson 2005; Schüppert and Gooskens in press) was confirmed by our data. Danes decoded more items (\(\bar{x} = .90\)) than Swedes did (\(\bar{x} = .87\)) when confronted with the neighbouring language (\(t(114) = 1.71, p = .04\), one-tailed). Mean word recognition scores per L1 and per grade are illustrated in Figure 8, which also suggests that word recognition gets better with age. This is confirmed in a correlation analysis between the factor “age” and “word recognition” which results in \(r = .61 (p = .004)\).

![Figure 8](image.png)

4.4.3. Correlation between language attitudes and word recognition

A Pearson correlation conducted on the factors word recognition accuracy and the extracted principal components “attractiveness” and “modernity” revealed that they do not correlate significantly. Neither do word recognition scores and any of the personality traits. This is true for the group of 116 participants, as well as for subsets defined by L1 (two subsets), grade (five subsets) or both (ten subsets, all \(p > .01\)). However, two of three investigated factors, namely attractiveness of the speaker when speaking the neighbouring language and word recognition of this language, change significantly with age. Attractiveness of the bilingual speaker correlates negatively with age, which means that attitudes get more negative with age, while word recognition correlates positively with age, meaning that children have fewer difficulties to decode the neighbouring language the older they get.

In a last step, we test the hypothesis that language attitudes and word recognition are linked to each other. As both variables correlate significantly with age in opposite directions, a correlation analysis across all age group is likely to yield misleading results. To avoid this, we normalised for age by calculating z-scores of
three factors (attractiveness, modernity, word recognition) for all five age groups (7-8, 9-10, 11-12, 13-14, 15-16) separately. The standardised values of these variables show similar patterns as the non-standardised values do: Danish participants have fewer difficulties decoding spoken Swedish than Swedish participants have decoding spoken Danish ($t(114) = 1.9, p = .03$ one-tailed), and Danish participants find the bilingual speaker more attractive when she speaks Swedish than Swedish participants do when she speaks Danish ($t(149) = 3.3, p = .001$), while the speaker was judged as equally modern in both L1 groups when she spoke the neighbouring language.

If the age factor is controlled for in this way, a Pearson correlation between these three factors (word recognition and the two extracted principal components “attractiveness” and “modernity”) resulted in a low, but significant positive correlation between “attractiveness” and “word recognition” ($r = .19, df = 114, p = .04$) but no significant correlation between “modernity” and “word recognition”. That means that, in our population, listeners’ attitudes towards the neighbouring language explain 3.6% of their word recognition variance.

4.5. Discussion and conclusion

4.5.1 The Development of Language Attitudes

One important finding in this study has been that for both groups of participants (Danish and Swedish), attitudes towards the neighbouring language become more negative with age (confirming findings reported by Schüppert and Gooskens (in press)). What is more, it can be concluded that while the youngest participants hold symmetric attitudes (i.e. Danish and Swedish 7 to 8 year old children are equally neutral), the older participants hold asymmetric attitudes, Danish listeners having a more positive attitude towards Swedish than vice versa. This development of negative language attitudes seems to happen relatively independently of the development in comprehension ability, as no statistically significant correlation between the two factors in our data could be established if the age factor was not controlled for.

The age period in which language attitudes develop is difficult to pin-point on the basis of our data. It could be reasoned from our data that the asymmetry in language attitudes emerges in the age range between 12 and 15 years. This is in contrast to findings reported by Day (1982), who reviews studies investigating the age factor in language attitudes and reports that children develop the association of nonstandard with low socio-economic status (SES) and standard with high SES between the ages of 3 and 7, and that, in the same age, children who speak a non-standard variety generally change from a positive or neutral attitude towards their own variety to a positive attitude towards the standard variety. The studies reported by Day (1982), however, all investigated attitudes toward the children’s L1 and not a closely related language.

An exception from the general trend that (a) attitudes held towards the neighbouring language become more negative with age and (b) younger participants hold rather neutral and symmetric attitudes towards the neighbouring language is
the group of 9 to 10 year old children. Here we assume that it is the Swedish group of 9 to 10 year old participants that behaves differently from the rest of the subjects. We cannot exclude that factors such as a stressful day at school or a test in a previous class might have had an effect on these particular children’s mood and subsequently also their evaluative ratings.

4.5.2. The relationship between language attitudes and intelligibility of a closely related language

The main aim of this article was to investigate the relationship that exists between people’s attitudes towards a linguistic variety and their ability to comprehend that variety. Previous literature has suggested a link between attitudes held towards a specific language variety and the ability to decode that variety. Wolff (1959) and Giles and Niedzielski (1998) suggest opposing causal relationships. While Wolff (1959) suggested that a negative attitude might have a detrimental effect on intelligibility and vice versa, Giles and Niedzielski (1998) hypothesised that attitudes towards a specific language may be influenced by the degree of intelligibility of that language. In the current investigation we attempted to shed light on the relationship between attitudes and intelligibility by investigating the development of language attitudes alongside mutual intelligibility of a neighbouring language (Danes for Swedes and Swedish for Danes) in Scandinavia. The causal relationship between these two factors was not investigated, however.

Our data indicate discrepancies both in language attitudes and in comprehension scores between the two groups of children (Swedish and Danish). In a matched guise test, Danish participants find a bilingual speaker more attractive when she speaks Swedish than Swedish participants do when she speaks Danish. This suggests that Danish children hold a more positive attitude towards Swedish than Swedish children hold towards Danish and confirms earlier findings by Maurud (1976), Delsing and Lundin Åkesson (2005) and Schüppert and Gooskens (in press) who also concluded that Danes are more positive towards Swedish than Swedes are towards Danish.

Contrary to Schüppert and Gooskens (in press) and in line with Delsing and Lundin Åkesson (2005), we found a low but significant positive correlation between attitudes and intelligibility. Participants with a positive attitude towards the neighbouring language perform better in the word recognition experiment than those with a negative attitude and vice versa. Although our data cannot answer the question on the causal relationship between intelligibility and attitude, a weak link seems to exist between these two factors. This means that, at least indirectly elicited, language attitudes seem to have some bearing upon the degree to which adolescents comprehend a closely related linguistic variety.

It is important to note, however, that the effect of language attitude in our experimental setting was very limited, as it merely explains 3.6% of the variance. That means that more than 95% of the variance is explained by factors other than language attitude. These factors might be general verbal talent, the amount of contact
that the listeners have had with the neighbouring language, or mental access to other foreign languages (such as Norwegian). It is also a possibility that orthographic knowledge is activated during spoken word recognition (as indicated by Ziegler & Muneaux 2007, Pattamadilok et al 2009, Perre et al. 2009). This could mean that Danes might have an advantage from having learnt a spelling that reflects a proto-Nordic variety to a larger extent than Swedish spelling does. It has also been suggested that linguistic factors could play a role in explaining the asymmetry in intelligibility. Hilton et al. (accepted) found that Danish news readers produce more phonetic syllables per second than Swedish news readers do. As both groups of speakers produced equally many phonological syllables, this finding suggests that phonological syllables are reduced to a larger degree in spoken Danish than in spoken Swedish. Findings reported by Vanhove et al (2010) indicate that Standard Danish vowels occupy larger spaces in the articulatory area than Swedish vowels do, and that there are more overlaps found between Danish vowel spaces than between Swedish vowel spaces. However, Schüppert & Gooskens (accepted) provide evidence that these linguistic factors might play a role only in combination with extra-linguistic factors, as Danish and Swedish pre-schoolers in their experiment performed equally well in a spoken-word recognition task in the neighbouring language. An example for a combination of linguistic and extra-linguistic factors is the relatively high amount of reduction in Danish compared to Swedish, which is not reflected in conservative Danish orthographic rules. As Danish orthography generally reflects the proto-Nordic pronunciation more faithfully, orthographic rules might serve as additional cues for literate speakers of Danish when confronted with spoken Swedish, while this is not the case for literate speakers of Swedish, as their orthography is more adapted to current pronunciation of standard Swedish. For example, the morpheme *bad* in Danish *badekar* /be:dækãυ/ (‘bath tub’) is spelt word-finlally with a letter that also represents a a dental plosiv such as in *dal* /de:λ/ (‘valley’). When literate Danes are confronted with the spoken Swedish cognate *badkar*, pronounced with a dental plosive (/bækɔːr/), it might be easier for them to match *badkar* to *badkar* than it for Swedish-speaking listeners confronted with the unfamiliar approximant /ð/ in the Danish item *badekar*. Future research should look into the role of these other extra-linguistic and linguistic factors for intelligibility more in detail and specifically investigate whether L1 orthography serves as an additional cue during spoken word recognition of a closely related language.

All in all this study has given strong indications that language attitudes and comprehension ability are two factors that develop relatively independently of each other, but that certain age specific attitudes are linked to intelligibility. A thorough investigation of the causal relationship between intelligibility and language attitudes is highly desirable.

---

3 Note that the phonetic sound which is transcribed by the IPA character ð is an approximant articulatorily, and not a fricative as in English *the* /θe/.
Acknowledgements

We thank Erik Vedder for help with the data collection and Vincent van Heuven and Renée van Bezooijen for their helpful commentson an earlier draft of this paper.

References


Vanhove, J., T. Leinonen, V. J. van Heuven, C. Gooskens & A. Schüppert. 2010 Comparing the acoustic distinctiveness of Danish and Swedish vowels. Paper presented at Experimental Approaches to Perception and Production of Language Variation (ExAPP2010), Groningen.


5 Syllable Reduction and Articulation Rates in Danish, Norwegian and Swedish

The final version of this paper will be published as
Nanna Haug Hilton, Anja Schüppert & Charlotte Gooskens: “Syllable Reduction and Articulation Rates in Danish, Norwegian and Swedish” in Nordic Journal of Linguistics, Vol. 34, No. 2. 3 (Oct. 2011) by Cambridge University Press. All rights reserved.

Abstract

This investigation compares articulation rates of phonological and phonetic syllables in Norwegian, Swedish and Danish to investigate differences in degrees of syllable deletion (reduction) between these three languages. For the investigation two sets of data are used: one consisting of recorded speech from radio news and another consisting of read sentences. The results from the comparative investigation show that in both data sets Danish exhibits a much larger degree of syllable reduction in speech than Norwegian and Swedish do. The finding that syllable deletion processes take place in Danish that do not occur in Norwegian and Swedish is viewed as typological. The finding that Danish words are shorter than their Norwegian and Swedish counterparts is also viewed as a potentially contributing factor to problems that arise in inter-Scandinavian communication.

5.1. Introduction

Linguistic differences between the North Germanic languages Norwegian, Swedish and Danish have been described to a large extent in previous literature (e.g. in Bandle et al. 2005). There are, however, some linguistic features that have not yet been studied in depth or which comparative investigations are lacking. The degree to which syllables and phonemes are reduced in speech is one such feature. This article reports a comparative investigation of reduction and articulation speed in the three most widely spoken Scandinavian languages. Through a descriptive investigation of differences in speech tempo in Norwegian, Swedish and Danish our study aims to portray fundamental differences in the phonetic realisations of the three languages.

5.1.1 Linguistic Similarities between the Norwegian, Swedish and Danish

The three North Germanic languages spoken in mainland Scandinavia (here used to refer to Norway, Sweden and Denmark only) are typologically closely related: they share a great proportion of their lexis and have but few differences in their syntactic structures. They are mutually intelligible to a large extent (Maurud 1976; Delsing & Lundin Åkesson 2005). Historical events that we will not go into details about in this article have led to a situation where Norwegian and Danish are especially closely related on the lexical level. A review of descriptive literature of differences between the Scandinavian national varieties (Brønsted 1967; Lundeby 1969; Landmark 1970;
Fjeldstad and Hervold 1989; Fjeldstad and Cramer 1992; Cramer and Kirkegaard 1993; Mårtenson and Fjeldstad 1993; Braunmüller 1998; Zola Christensen 2007; Teleman 2008) indicates that while morpho-syntactic differences are rare, differences in phonology are more substantial between the three varieties. The accounts cited above indicate that the highest number of morpho-syntactic differences is identified between Norwegian and the other two varieties. The acknowledged differences between the language pairs generally comprise minor discrepancies in word order (particularly involving a negator or an infinitival marker), or the morphological marking of definiteness. Another example often mentioned is the placement of the particle in phrasal verb constructions. In Danish the particle must occur after the object while it comes before the object in Swedish and can occur in either place in Norwegian.

Phonological differences across the three varieties are generally considered to be larger, particularly between Danish and the other two languages, as Swedish and Norwegian share a number of phonological traits. For instance, both languages make use of two lexical tonal accents that can distinguish minimal pairs. Also, if we compare the vowel system of Standard Swedish (e.g. Leinonen 2010) with that of the Norwegian variety spoken around the capital Oslo (e.g. in Kristoffersen 2000) - a variety often referred to as Standard Norwegian (Røyneland 2005), we find very similar inventories. Both languages have 9 distinctive long vowels and 9 distinctive short vowels. The Danish phoneme inventory, however, is mostly described consisting of more contrastive vowels than Swedish and Norwegian, with 12 distinctive long vowels and 13 distinctive short vowels (e.g. Basbøll 2005:50). In addition to differences in the phoneme inventory, Danish has another feature that is different from anything found in Swedish and Norwegian, namely *stød*, which has been analysed as a realisation of creaky voice or laryngealisation by Grønnum (1998: 179) and Basbøll (2005: 83).

The languages in question, Norwegian, Swedish and Danish are among the languages that have been referred to as stress-timed languages. The syllable-timed or stress-timed distinction is not a dichotomy, but rather a continuum with two poles. The pole ‘syllable-timed’ can be described as speech where all syllables have the same duration. ‘Stress-timed’, on the other hand can be described as speech where the time duration between the stressed syllables is equal. Germanic languages are generally described on the ‘stress-timed’ end of the continuum (e.g. Ladefoged 1975). This implies that the languages employ some degree of reduction in fluent speech to achieve relatively similar durations between major stress groups in the utterance. Reduction is one of the features of Danish that has received some attention in previous studies (see below) while arguably all the Scandinavian North Germanic varieties exhibit some reduction in speech. East Norwegian, as well as most Swedish varieties, show an assimilation process, for instance, by which rhotics with succeeding alveolar consonants /d, l, n, s, t/ change into retroflex sounds /ð, l, ʂ, t/. Schwa-deletion and syllabification of /n/ can occur word-finally in all three
languages, for instance in the pronunciation of Danish *våben*, Swedish *vapen* and Norwegian *våpen* ‘weapon’.

Interestingly, reduction-related processes like assimilation, schwa reduction and lenition have received particular attention in work on spoken Danish, possibly due to the link between the processes and problems in speech development and speech perception. Bleses et al. (2008) suggest that the large number of reduction and assimilation processes in Danish causes or aggravates the delay in vocabulary development in Danish infants and children compared to that of their peers from ten European countries and from the US and Mexico. They point out that schwa-deletion and the vocalisation of consonants result in long vocalic stretches, making the Danish sound structure unclear with weak, or even no, cues for word and syllable boundaries. Grønnum (2007) gives numerous examples of phonetic reduction taken from the Danish Phonetically Annotated Spontaneous Speech (DanPASS) corpus. She notes how the phrase *behøver ikke* ‘does not need to’, /be.ʌ.ʊ.e.ɡ.a/ in its phonological form, is reduced in fluent speech to single syllable [b̠œ̝ɡ]. Similarly the phrase *kan jo ikke* ‘cannot’ is pronounced [ˈk̠a.j.o.ɛɡ] in its canonical, or full form, but reduced to [k̠e̝ɡ] in fluent speech. Notably four or even five phonological syllables in the two examples are reduced to one single phonetic syllable.

Descriptions of equally dramatic reduction processes have not been found for Norwegian and Swedish in the literature. However, if we consider the same examples from Norwegian and Swedish, a large number of reduction processes can also be seen. The phrase *behøver ikke* ‘does not need to’ can be reduced from its canonical form /bœ̝.hø.ʊ.ɾi.ka/ to [bœ̝.kə], from five to two syllables. In Swedish, the same applied where *behöver inte* ‘does not need to’ has five canonical syllables in phonological form /bœ̝.h̠.ʊ.ɾ.in.ta/ but can be reduced to bisyllabic [bœ̝n.ta] in fluent speech.

There are a number of articulatory processes that fall in the category of what is referred to as ‘reduction’. Articulatory weakening in the production of vowels is particularly referred to as reduction in phonetic literature. Laver (1994:157) lists shortening, pitch-lowering, centralisation or lowering of intensity as vowel-reduction processes. These form, along with syllable re-organisation processes, the compression, or reduction of syllables in fluent speech. The elision or lenition of consonants can also be viewed as reduction processes. This investigation is, however, not concerned with reduction processes on phoneme level, but considers instead the extent of deletion of syllables in spoken Norwegian, Swedish and Danish. By conducting a comparative investigation of the numbers of phonological and phonetic syllables produced per time unit in the three Scandinavian languages, we investigate whether the syllable reduction processes are indeed more extreme in Danish as opposed to in Norwegian and Swedish, as indicated by the previous literature described above.

To facilitate a measurement of speech reduction the current study therefore compares the number of canonical syllables in six data sets of natural speech from Norway, Sweden and Denmark with the number of phonetic syllables produced in the
same data. The phonetic syllables are measured automatically by the computer counting the number of voiced intensity peaks in the speech signal. The difference between the two measurements is an indication of the degree of reduction that occurs in the three languages.

By counting the number of syllables produced in natural speech we also measure the articulation rates at which these languages are spoken. This gives us further opportunity to compare the three languages on a phonetic level. Our study is in this way two-fold. Its main aim is to give an account of syllable reduction in Danish, Norwegian and Swedish, which in turn can be used for a comparative investigation of differences on a purely phonetic level, i.e. in speech tempos, between the languages.

5.1.2. Previous Research of Tempo and Reduction in Speech

Reduction and the tempo at which speech is produced go hand in hand. To increase the rate of our speech, our articulatory processes become less precise and some phonological content can be deleted or shortened in duration. Likewise, the more reduction that takes place in the production of an utterance, the shorter the time must be to produce it. In this paper, we compare the tempos at which Norwegian, Swedish and Danish are spoken.

Studies of the tempo at which speech is produced are generally concerned either with speech rate or articulation rate. Speech rate is defined as the number of items (words, syllables, phonemes etc.) produced during a specific time period. This means that pauses in speech are considered part of the signal and taken into the measurement. Articulation rate, on the other hand, is a measure of the amount of articulatory activity within a time frame, i.e. the number of speech items per time unit after silent intervals have been removed from the signal. This study is concerned with articulation rate and reports study of speech where silent intervals longer than 150 ms have been removed. Previous studies of articulation rates disagree on the exact measure of what constitutes a meaningful pause and what does not. Campione and Véronis (2002) claim for their study of pauses in German, Italian, English, French and Spanish that pauses shorter than 200 ms are difficult to discriminate from occlusives but that some brief pauses can be as short as 60ms in length. In the current study we follow Tsao and Weismer’s (1997) suggestion with a cut-off point at 150ms for pauses. This duration was chosen based on their claim that 150ms is longer than the typical stop closure interval, yet probably the lowest threshold of what constitutes a meaningful pause (Tsao and Weismer 1997:861). As the aim of our comparative investigation is to put side by side amounts of articulatory activity per time period in three different linguistic varieties, pauses are excluded from the speech signal to ensure that the data sets from the different languages are as comparable as possible for this purpose.

Some previous studies have been conducted to investigate tempo of speech (either speech rates or articulation rates) cross-linguistically, many of which have concluded that tempos do not differ substantially across languages. Osser and Peng
(1964) compared the number of phonemes produced per minute by native speakers of American English and native speakers of Japanese but found no significant cross-linguistic difference in speech rate. Neither did Kowal et al. (1983), who re-evaluated findings from earlier studies based on spontaneous speech in English, German, French, Spanish, and Finnish. Den Os (1988) conducted a comparative study on Italian and Dutch speech rate. She analysed the tempo of reading aloud by native speakers of Italian and Dutch but did not find a significant difference when syllables per second were compared across languages. When phonemes per second were compared, however, articulation rate in Italian turned out to be significantly slower than in Dutch. This might have to do with the fact that Dutch has more complex consonant clusters than Italian does, and that the Dutch syllable thus generally exists of more phonemes.

What the studies above have in common is that they investigate tempo in fluent speech in languages that are not mutually intelligible, or even very closely related. However, by comparing syllable reduction and articulation rates in closely related varieties, the possibility that differences found between languages stem from different phonotactics or lexical structure in the languages can be excluded. A number of studies exist where variation in tempo between different regional varieties of the same language has been investigated. Verhoeven et al. (2004) compared articulation rates in Belgian and Netherlands Dutch and found that Netherlands Dutch varieties are articulated at a significantly higher rate than Belgian Dutch varieties are. Robb et al. (2004) also attest variation in English between varieties spoken in Christchurch (New Zealand) and that spoken in Connecticut (American English) and conclude that New Zealand English is spoken at a higher rate than American English. Robb et al. (2004) relate their findings back to reduction processes and suggests that because New Zealand English has a high degree of vowel raising, vowels tend to be shorter in New Zealand English than in American English, which again has an effect on articulation rates in the two varieties.

Although no comparison has ever been made of the tempo at which the three North Germanic Scandinavian languages are spoken, some previous work has been done to look into speech rates in the respective languages. One quantitative study of speech and articulation rates in Norwegian exists (Almberg 2000). Almberg (2000) reports tempo in the pronunciation by 60 informants from three different regions of Norway. Almberg’s (2000) data consist of read strings of numbers in addition to announcements of telephone numbers that the informants know by heart. He concludes that there is a significant correlation between articulation rates and utterance length in his data: longer utterances are produced at higher articulation rates. This finding is not new, and is most likely universally true (Fonagy & Magdics 1960). The mean articulation rates in Almberg’s (2000: 66) Norwegian corpus vary between 3.6 and 4.4 syllables per second depending on the utterance length. The study reports no regional differences found in speech tempo in Norway. Importantly for the current investigation, Almberg (2000) does not consider the amount of syllable reduction that occurs in fluent speech in his data.
Jande (2003) considers the effect of reduction and speed on perceived naturalness of synthesised Swedish speech. Jande (2003) concludes that reduced speech sounds more natural than canonical speech when the speed of the speech is fast (at medium or high speech rates). The reduction measured in Jande’s (2003) work consists of vowel or syllable deletion processes mainly, but the work gives no information about the degree of reduction that occurs in natural Swedish speech as opposed to in Danish or in Norwegian. One recent study has looked at articulation rates in Swedish (Hansson 2002). The study is primarily concerned with rate differences between words within the prosodic phrase and concludes that reduction in speech happens throughout the prosodic phrase and that Swedish shows some signs of phrase-final lengthening. Hansson (2002) makes no claims about variation (or consistency) in articulation rates across speakers or regions, however.

No previous investigation has reported empirical results on speech or articulation rates in Danish. As mentioned in the previous section, however, a number of studies have covered speech reduction in Danish. We hypothesise here that a large degree of reduction in speech also means that Danish is produced at a high articulation rate. As previous research has presented evidence that higher speech rate reduced our ability to successfully perceive speech (e.g. Vaughan & Letowski 1997, Gordon-Salant et al. 2007, Jones 2007), a higher number of reduction processes (and/or a subsequent higher rate of articulation) in Danish could be one contributory factor to why spoken Danish is generally viewed as a difficult language to understand by other Scandinavians (c.f. Delsing & Lundin Åkesson 2005).

5.2. Method

5.2.1. Material and speakers

The aim of the current investigation is to conduct a comparative investigation of syllable reduction across Norwegian, Swedish and Danish. For the investigation, equivalent types of data from the three national speech communities should be analysed. Our investigation focuses on standard-like speech in the three countries using data consisting of short sentences with cognate words read aloud as well as radio news broadcasts read by professional news readers. Both types of data were recorded in a highly controlled setting and so our findings will not necessarily reflect reduction and articulation rates found in less formal speech. An advantage of using these data is, however, that they are highly comparable across speaker communities.

The radio news broadcasts data set (1) was compiled of recordings made by the three previously state owned nationwide radio stations in Norway (NRK), Sweden (SR) and Denmark (DR). The Norwegian recordings used were originally aired on the station P1, the Swedish recordings on stations P1 and P3 and the Danish recordings on stations P1 and P4. A total of 26 minutes of fluent speech was used for the analysis. The lengths of the recordings vary between 21.43 to 44.85 seconds. The data were produced by 55 informants, distributed evenly across nations and genders save an additional male Danish speaker, as illustrated in Table 1.
The news broadcasts were all aired in spring 2010 and had been recorded with speakers who use a standard accent. This means that the Swedish and Danish broadcasters spoke their respective codified Standard varieties while the Norwegian informants all spoke a variety that used Bokmål features with an East Norwegian accent.

The second data set (2) used for this investigation consisted of 16 read sentences produced by 9 male speakers, 3 from each of the countries’ capital cities Oslo, Stockholm and Copenhagen. This second data set was added to the investigation to compare the larger corpus of radio news broadcasts consisting of different lexical contents (1) to data consisting of the same lexical (and morphological) contents (2). The sentences recorded by the 9 speakers for this control set all consisted of cognate words with the same number of canonical syllables in all three languages. The speakers who recorded the sentences all use similar varieties to the speakers who produced the data in the radio news broadcast corpus: Standard Swedish, Standard Danish and an urban East Norwegian variety based on Bokmål. The speakers were all students in their 20s.

The sentences recorded for the second data set were Semantically Unpredictable Sentences (henceforth SUS) generated by the method developed by Benoit et al. (1996). These sentences are generally used in sentence intelligibility experiments, but are also ideal for analysis of articulation rates. The SUS are syntactically correct but consist of phrases with concepts that are not likely to be semantically related to each other (cf. Gooskens et al. 2010 for a more detailed description of the material). It was hypothesised that reading these sentences would be equally difficult in all three languages and that speech rate could not be influenced by lexical combinations that might be more frequent in one language than in another. The SUS can be automatically generated using 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. An example of a SUS is given in (1).

(1) a. Danish Et folk deler et job som går.
   b. Norwegian Et folk deler et jobb som går.
   c. Swedish Ett folk delar ett jobb som går.
   d. English 'A people shares a job that walks.'

The sentences developed for the current experiment consist of Danish-Norwegian-Swedish cognate words only and vary in length having from 6 to 10 phonological

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denmark</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Norway</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Sweden</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>27</td>
</tr>
</tbody>
</table>
syllables (either 6 or 7 words). All sentences read by the informants can be found in the Appendix.

5.2.2. Measurements

Previous investigations have mainly measured articulation rate as the number of syllables produced per second (e.g. Kowal et al. 1983; Den Os 1988; Almberg 2000; Verhoeven et al. 2004). To enable comparison of our results with previous studies, we measure articulation rate in the same way here. Importantly, however, we make a distinction between phonetic syllables (actually produced) and phonological syllables (canonical syllables) to be able to measure the degree of reduction that takes place in the languages.

All sound recordings analysed in this investigation were transcribed in Praat (Boersma & Weenink 2008). Any pauses in the speech signal with durations of more than 150 ms were removed. The length of the recordings was established for each individual speaker, and canonical, or phonological, syllables were counted based on orthography. This count was checked against descriptions of the phonological syllables in the three languages (Grønnum 1998 for Danish; Kristofersen 2000 for Norwegian; Elert 1966 for Swedish). The number of syllables was subsequently divided by utterance duration to calculate the articulation rate of phonological syllables.

To calculate degree of reduction, and articulation rate of phonetic syllables, the number of syllables actually produced in the three languages was determined automatically using a Praat script developed by De Jong & Wempe (2009). This script counts the number of intensity peaks (with drops in intensity of 2 dB immediately before and after the peak) in the speech signal that have voicing (where F0 can be measured). To obtain individual phonetic articulation rates per speaker, the number of syllables produced by every speaker is simply divided by the duration of the analysed sample for this speaker. One advantage of an automated measure of syllables is that a comparable measurement can be made for all three languages without human interference. Human knowledge of phonology and underlying syllable structure could influence a researcher’s ability to objectively identify phonetic syllables in a speech signal. The disadvantage of the script lies in the same matter, however, since it can count somewhat differently from that which a human does. De Jong & Wempe (2009) found that automatic and human syllable detection correlate highly ($r > .71$). De Jong & Wempe’s (2009) results show that human and automatic syllable detection are not completely congruent processes. It remains unclear, however, whether the algorithm detects too few syllables, or whether humans detect too many under the influence of their phonological or orthographic knowledge. Importantly for our study, the discrepancy between human and automatic syllable detection would be the same across all languages measured as long as the same parameters are used for the measurement.
Examples of how the automated analysis deals with specific phonetic details in Norwegian, Swedish and Danish are shown in Figures 1-3 portraying the output of the automated analysis in PRAAT for three speakers producing the SUS ‘A free earth answers a friend’ taken from data set (2). We see that the automated phonetic syllable count finds 5 syllables for the Danish speaker, 7 syllables for the Norwegian speaker and 7 syllables for the Swedish speaker. Particularly relevant for the current study is perhaps the treatment of stød in Danish (see introduction) by the automated count. One worry is that the laryngeal activity occurring with stød could interfere with the automated measure of voiced intensity peaks in the signal. In Danish the words fri ‘free’ and jord ‘earth’ have stød. The occurrence of stød in the sentence in Figure 1 has no bearing on the syllable count for Danish, however. The analysis counts three phonetic syllables for the three nucleus vowels produced in the first half of the sentence en fri jord ‘a free earth’.

The automated count finds 7 phonetic syllables in both the Norwegian as well as the Swedish recording (cf. Figure 2 and Figure 3). From an auditory analysis, it is clear that the two speakers in the recordings use a particularly careful pronunciation of the second half of the illustration sentence. All canonical syllables are produced in svarer en venn / svarar en vän ‘answers a friend’ by both the Norwegian as well as the Swedish speaker. The Norwegian produces [svær.en.ven] and the Swedish [svær.en.ven]. Comparatively, the automated analysis finds only two phonetic syllables in this part of the utterance as spoken by the Danish speaker (Figure 1). In an auditory analysis, the authors counted only two phonetic syllables in this part of the Danish speaker’s utterance and thus transcribed his utterance: [svær.en.ven].

The examples illustrated by Figures 1-3 show findings in line with de Jong & Wempe’s (2009) claim that an automated syllable count correlates highly with a human count. In our opinion, both approaches to syllable counting have their strengths and weaknesses. A manual approach is error-prone due to human’s phonological knowledge that might interfere with the count, whilst an automated analysis could presumably be influenced by changes in voice quality where intensity becomes lower, or the Fo becomes weaker. With a large set of recordings an automated count of syllables is clearly preferable, and since this current study aims to investigate reduction in a fairly large data set we opt for an automated phonetic syllable count for our analysis.

Figure 1. Oscillogram, spectogram and phonetic syllable tier for the Danish sentence ‘En fri jord svarer en venn’.
Figure 2. Oscillogram, spectogram and phonetic syllable tier for the Norwegian sentence ‘En fri jord svarer en venn’.

Figure 3. Oscillogram, spectogram and phonetic syllable tier for the Swedish sentence ‘En fri jord svarar en vän’.

5.3. Results

5.3.1 Phonological Syllables

Table 2 shows the overall measurements in the radio news broadcast data set along with the mean number of phonological syllables produced per second in each recording. Table 3 below shows the same measurements in the read SUS data.

Table 2. Articulation rate of phonological syllables in the radio news broadcast data.

<table>
<thead>
<tr>
<th></th>
<th>Utterance length (s)</th>
<th>No. of canonical syll.</th>
<th>Articulation rate (syll/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danish</td>
<td>570.27</td>
<td>3543</td>
<td>6.21</td>
</tr>
<tr>
<td>Norwegian</td>
<td>502.64</td>
<td>2701</td>
<td>5.37</td>
</tr>
<tr>
<td>Swedish</td>
<td>500.44</td>
<td>2676</td>
<td>5.35</td>
</tr>
</tbody>
</table>

Table 3. Articulation rate of phonological syllables in the SUS data.

<table>
<thead>
<tr>
<th></th>
<th>Utterance length (s)</th>
<th>No. of canonical syll.</th>
<th>Articulation rate (syll/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danish</td>
<td>78.78</td>
<td>372</td>
<td>4.72</td>
</tr>
<tr>
<td>Norwegian</td>
<td>94.07</td>
<td>362</td>
<td>3.85</td>
</tr>
<tr>
<td>Swedish</td>
<td>105.76</td>
<td>357</td>
<td>3.38</td>
</tr>
</tbody>
</table>

The two tables show that the mean articulation rates of phonological syllables are rather different in the two types of recordings. Firstly, a faster syllable rate was produced in the radio news broadcasts than in the reading of sentences. This is in line
with earlier findings by Fonagy & Magdics (1960) and Almberg (2000), who reported that long utterances are produced at higher speed than short utterances. Furthermore, the SUS are semantically anomalous, which could also be a possible factor slowing speakers down. Secondly, and most importantly, in both data sets, the Danish speakers have a significantly higher articulation rate than Norwegians and Swedes have.

In a one-way ANOVA on the radio news broadcast data set with speakers’ individual means of phonological syllables produced per second, the difference is highly significant \( F(2, 52) = 22.56, p < .001 \). A Tukey post-hoc test reveals that only the difference between Danish and the two other languages is significant; the difference between Norwegian and Swedish is not. In Figure 4, a box plot illustrates the ranges in articulation rates for the three language groups. The line in the middle of the boxes is the median speech rate while the boxes represent middle two quartiles. As is visible from Figure 4, the Danish and Norwegian speakers have less variation in articulation rates than the Swedish speakers. The variation in Swedish speakers could be due to social background or age differences in the sample, but is not of significance for the typological discussion in the current article.

**Figure 4.** Box plot of articulation rates of phonological syllables in the news broadcast data by language.

A one-way ANOVA on the SUS data set with the nine speakers’ individual means of phonological syllables produced per second revealed that the differences between the languages are significant \( p < .05 \) in this data set as well \( F(2, 6) = 9.33 \). The post-hoc indicates, however, that only the difference between Swedish and Danish is significant in this data set. Norwegian lies in between and is neither significantly different from Swedish nor Danish.

### 5.3.2 Phonetic Syllables

Tables 4 and 5 below show the measurements of phonetic syllables as produced by the automated analysis with the Praat script.

**Table 4.** Articulation rate of phonetic syllables in the radio news broadcast data.

<table>
<thead>
<tr>
<th></th>
<th>Utterance length (s)</th>
<th>No. of phonetic syll.</th>
<th>Articulation rate (syll/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danish</td>
<td>570.27</td>
<td>2498</td>
<td>4.38</td>
</tr>
<tr>
<td>Norwegian</td>
<td>502.64</td>
<td>2215</td>
<td>4.41</td>
</tr>
<tr>
<td>Swedish</td>
<td>500.44</td>
<td>2242</td>
<td>4.48</td>
</tr>
</tbody>
</table>
Table 5. Articulation rate of phonetic syllables in the SUS data.

<table>
<thead>
<tr>
<th>Utterance length (s)</th>
<th>No. of phonetic syll.</th>
<th>Articulation rate (syll/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danish</td>
<td>78.78</td>
<td>309</td>
</tr>
<tr>
<td>Norwegian</td>
<td>94.07</td>
<td>373</td>
</tr>
<tr>
<td>Swedish</td>
<td>105.76</td>
<td>370</td>
</tr>
</tbody>
</table>

We tested whether articulation rates of phonetic syllables differed across the three groups of speakers in the two corpora (news readers and SUS): this turned out not to be the case. A one-way ANOVA on speakers’ mean articulation rates of phonetic syllables showed no significant differences between Norwegian, Swedish and Danish rates, neither for the news corpus, nor for the SUS corpus.

5.3.3 Reduction

With a count of phonological and phonetic syllables produced per time unit in all three languages, the extent of reduction in every speaker group was calculated. The articulation rates of phonological and phonetic syllables are compared in Figure 5. The figure makes clear that the difference between the number of phonological and phonetic syllables is larger in the news corpus than in the SUS corpus. Furthermore, Figure 5 indicates that the difference between the number of phonological and phonetic syllables is the largest in the Danish language group. The difference between the number of phonological and phonetic syllables counted in the radio news broadcasts is significant for all three languages in paired samples t-tests of speakers’ mean articulation rates of phonological versus phonetic syllables within every language group (Danish: $t(18)=-16.184$, $p<.001$; Swedish: $t(17)=-5.302$, $p<.001$; Norwegian: $t(17)=-9.174$, $p<.001$). In paired-samples t-tests of speakers’ mean articulation rates for the SUS data, only Danish shows a significant difference between phonological and phonetic syllables ($t(2)=-9.442$, $p<.03$).

Figure 5. Reduction ratio (no. of canonical syllables minus no. of phonetic syllables for both data sets.

To calculate whether the difference between phonological and phonetic syllables, i.e. the degree of reduction is larger for the Danish speakers than for the Norwegian and Swedish speakers in the radio news corpus, the reduction ratio was calculated by subtracting the number of phonetic syllables from the number of phonological syllables counted per speaker. A one-way ANOVA was subsequently conducted on these individual differential scores (i.e. the individual reduction ratio) with language as an independent factor in the analysis. The results reveal that the differences
between the number of phonological and phonetic syllables is indeed significantly larger for the Danish speakers (mean 1.8) than for the Norwegian (mean 1) and Swedish speakers (mean 0.9): \( F(2, 52) = 15.793, p<.001 \). A post-hoc Tukey test shows that the difference between Swedish and Norwegian is not significant, and that Danish has significantly more reduction than both Swedish and Norwegian.

5.4. Discussion

5.4.1 Reduction in Danish, Norwegian and Swedish

The results presented above show that all three Scandinavian varieties have a degree of reduction in speech, but that this is the largest in the Danish data. In the comparison of the articulation rates of phonological and phonetic syllables made in the current investigation, the ratio of the number of phonetic syllables per phonological syllables is by far the largest for Danish speakers. This difference can be seen as a measurement of the elision of syllables that occurs in fluent speech. Previous literature has indicated that the reduction of syllables is a particularly noticeable feature in spoken Danish. The current investigation can substantiate this claim with data that shows that the amount of syllable reduction that occurs in Danish is indeed larger than that which occurs in Swedish and Norwegian.

In the radio news broadcasts the number of phonological syllables differs significantly from the number of syllables actually produced, i.e. phonetic syllables. This is the case for all three speaker groups and indicates that syllable reduction happens to some degree in fluent speech in all the varieties in question. This is not surprising, but makes the findings from the SUS data more interesting: the automated count of phonetic syllables shows that there are no real differences between the number of phonetic and phonological syllables in the Swedish and Norwegian SUS data. Sentence reading is a rather formal task (as opposed to speaking freely or speaking with notes only) where a large deal of attention is paid to speech. It thus seems that in Norwegian and Swedish most, if not all, phonological syllables are realised in such speech situations. This is not the case for Danish, however. The results from the SUS data indicate that in Danish there is a large degree of syllable reduction even in careful speech styles.

Our results thus indicate a substantial typological difference between the three Scandinavian languages: Danish words are produced as shorter than Norwegian and Swedish words. There is no \textit{a priori} reason to believe the number of underlying syllables should be different for Danish than for Norwegian and Swedish. Indeed, Grønnum (1998) defines the Danish phonological syllable as built around the nucleus vowel (or in some cases a sonorant consonant), which is also the case for the Norwegian (Kristoffersen 2000) and Swedish syllable (Elert 1966). Grønnum (1998:211) indicates that Danish speakers are able to identify the correct number of canonical syllables in words, and there seem to be no grounds for claiming that this is different for Swedish or Norwegian speakers. However, Danish speakers reduce the canonical syllables to a much larger degree than Swedish and Norwegian speakers do.
It could be that this large-scale reduction has repercussions for the canonical representations of syllables over time. Reduction processes such as apocope, syncope or haplology result in syllables disappearing from the canonical syllables and words. McMahon (1994) uses the haplology example Eng-la-lond from Old English undergoing loss of the medial syllable and resulting in today’s canonical England. It could be that we are witnessing Danish undergoing a severe sound change where reduction processes, even in the formal speech styles investigated here, happen at a large scale. More research is needed to investigate which types of syllables become elided in Danish and exactly which stylistic constrains there might be on such reduction.

5.4.2 Articulation Speed in Danish, Norwegian and Swedish

By investigating the number of syllables produced per time unit in a corpus of free speech in Danish, Norwegian and Swedish, we have also measured articulation rates in the three languages. Some of the results presented above could indicate that Danish speakers actually talk faster than Norwegians and Swedes do, but this would be a too simplistic rendition of our findings. What we have found is rather that in fluent Standard Danish speech a lot more phonological information is transferred per time unit than in fluent standard-like Norwegian and in fluent Standard Swedish. The Danish news broadcasters produce almost an entire phonological syllable more per second than their Scandinavian peers on average (0.84 more syllables per second than Norwegians, and 0.86 syllables than Swedes). This indicates that more semantic information is also transferred per time unit. The fact that this difference is also found in the highly unpredictable SUS-corpus where all words that were analysed had the same number of canonical syllables indicates that word length differences in the news corpus have had no bearings on the results. On the other hand, if one simply measures the number of syllables actually produced in fluent speech without taking into consideration the content of the speech signal, there are no differences between Danes, Swedes and Norwegians in their articulation rates. The automated analysis of voiced intensity peaks in the recordings of the three languages shows no significant difference in the number of phonetic syllables produced per time unit by speakers of different nationalities. Phonetically speaking, therefore, Danish is not spoken faster than Norwegian and Swedish in our data.

In conclusion, one might put forward another measure of speed in speech: communication rate. If one were to use a measure of the amount of phonological content which was transferred per time unit to determine speed, our data indicate that Danes communicate the fastest in Scandinavia and that this happens through large-scale syllable elision.

5.4.3 Comparison of our findings with previous studies

To give an indication of whether the differences reported above are indeed reliable reflections of typological traits of Danish and Norwegian and Swedish, it is useful to
compare our findings with those made in previous studies. Almberg (2000) found articulation rates between 3.6 and 4.4 phonetic syllables per second depending on utterance length in his Norwegian data of realised strings of numbers. This rate is comparable to that found for Norwegian in the current study. The mean number of phonetic syllables measured for our study is 3.97 and 4.41 in the radio news broadcast and SUS data respectively. Even if the number of phonetic syllables is counted automatically in our corpus and manually in Almberg’s (2000), the numbers of phonetic syllables produced per second in the two studies match perfectly well.

Unfortunately, no previous quantitative studies exist of articulation rates for neither phonetic nor phonological syllables in Danish and Swedish. The consistency of our Norwegian data in relation to that of Almberg (2000) would lead us to expect, however, that the overall measurements made for the Danish and Swedish are equally constant.

### 5.4.4 Implications of this research

The sections above have shown that on the one hand, there is a substantial difference between the articulation rates of phonological syllables and degree of phonetic reduction that occurs in Danish and that which occurs in Norwegian and Swedish. On the other hand, the phonetic rate differences between the three languages are negligible. Speakers of the three languages produce similar numbers of phonetic syllables per time unit in both data sets. We argue, however, that the differences found in rates of phonological syllable production and the subsequent phonetic reduction are more disadvantageous for the mutual intelligibility of Danish, Norwegian and Swedish than differences in rates of phonetic syllables would have been. The fact that about one in every four phonological syllables is deleted on the phonetic surface is likely to have an effect on speech comprehension. Non-native listeners who learn the phonological syllables of Danish through its orthography, or listeners from Norway and Sweden who rely on the phonological syllables in their own varieties, could struggle when confronted with the largely elided or reduced realisation of these phonological syllables.

Even if the three Scandinavian languages are very closely related and share a large portion of their lexical inventories, large-scale reduction in Danish speech production could be one (of many) explanatory factors for communication problems that occur between Norwegians, Swedes and Danes when speaking their native languages. Delsing & Lundin Åkesson (2005) report that Norwegians and Swedes understand less than half of the spoken Danish presented to them in test situations. Similar results have been reported by Maurud (1976) and Bø (1978). If Norwegian speakers rely on their own knowledge of phonological syllables for instance for the phrase behøver ikke ‘does not have to’ they might expect to hear something like its full form /bo.hø.ve.ri.ko/ or even a reduced form like [bœk.ko] with two produced syllables. In spoken Danish, however, canonical /be.hø.ve.e.go/ can be reduced to the single syllable [beg]. This shortening might cause comprehension problems for a Norwegian listener. More empirical research is needed to establish the degree to
which syllable reduction (and articulation rates) influence comprehension of closely related varieties.

5.5. Conclusion

The North Germanic languages spoken in Scandinavia exhibit differences in their degree of reduction in fluent speech: Danish speech contains more reduction than Swedish and Norwegian speech does. Phonetically, however, the three languages exhibit very similar patterns, the number of phonetic syllables produced per time unit does not differ substantially between speakers of the three languages. Although the latter finding suggests some degree of similarity between the three languages, we suggest that the large degree of reduction in Danish speech as opposed to Norwegian and Swedish speech could be an indicator of phonological change in progress occurring in Danish but not in the other Scandinavian national languages Norwegian and Swedish. Further research is needed to establish the nature and extent of syllabic reduction in Danish speech and its relationship to changes in phonology.

Differences in reduction or articulation rates of phonological syllables are also likely to negatively affect speech comprehension in Scandinavia. It is possible that syllable reduction is one of the traits of Danish speech that causes intelligibility problems for non-native as well as other Scandinavian listeners. The relative effect of this trait on intelligibility must be investigated in greater detail in future studies. This could be done by manipulating the duration of utterances recorded at different speech styles (for instance slow and accurate articulation and quick and highly reduced articulation), and thereby tearing apart the variables articulation speed and reduction. In this way, the relative contribution of both factors could be evaluated.

Acknowledgements

We would like to thank Laura de Klerk and Jan Vanhove for their help with the data collection and transcription. We would also like to thank Vincent van Heuven and Renée van Bezooijen for their helpful comments on an earlier version of this paper. Finally, we would like to thank the two anonymous reviewers for their helpful comments to this manuscript.

References


Den Os, E. A. *Rhythm and tempo of Dutch and Italian: A contrastive study*. PhD. Utrecht University


Gooskens, C., V. van Heuven, R. Van Bezooijen and Jos Pacilly. Is spoken Danish less intelligible than Swedish? *Speech Communication*, 52, 1022-1037


90


Appendix

Semantically Unpredictable Sentences, Norwegian version
En fri jord svarer en venn.
En god hånd ønsker et gulv.
En rød natt tar et navn.
En sjef eier en smak.
Et egg krever en tekst som håper.
Et folk deler en jobb som går.
Et grep kjører en jakt som lyser.
Et hjørne savner et bord som sitter.
 Hvordan når et brudd et fint ben?
Kjenn en skyld eller et krav.
Når legger en prins en lav skog?
Når retter en lengde en dyp feil?
Skriv et land og en munn.
Støtt et valg og en strøm.
Hvor elsker en stein et sent hull?
Vis en sol og en bok.

Semantically Unpredictable Sentences, Swedish version
En fri jord svarar en vän.
En god hand önskar ett golv.
En röd natt tar ett namn.
En stark chef äger en smak.
Ett ägg kräver en text som hoppas.
Ett folk delar ett jobb som går.
Ett grepp köper en jakt som lyser.
Ett hörn saknar ett bord som sitter.
Hur når ett brott ett fint ben?
Känn en skuld eller ett krav.
När lägger en prins en låg skog?
När rättar en längd ett djupt fel?
Skriv ett land och en mun.
Stöd ett val och en ström.
Var älskar en sten ett sent hål?
Visa en sol och en bok.

Semantically Unpredictable Sentences, Danish version
En fri jord svarer en ven.
En god hånd ønsker et gulv.
En rød nat tager et navn.
En stærk chef ejer en smag.
Et æg kræver en tekst som håber.
Et folk deler et job som går.
Et greb køber en jagt som lyser.
Et hjørne savner et bord som sidder.
 Hvordan når et brud et fint ben?
Kend en skyld eller et krav.
Hvornår lægger en prins en lav skov?
Hvornår retter en længde en dyb fej
Skriv et land og en mund.
Støt et valg og en strøm.
Hvor elsker en sten et sent hull?
Vis en sol og en bog.
Semantically Unpredictable Sentences, English Translation

A free earth answers a friend.
A good hand wishes a floor.
A red night takes a name.
A strong chef owns a taste.
An egg demands a text that hopes.
A people shares a job that walks.
A grip buys a hunt that illuminates.
A corner misses a table that sits.
How does a breach reach a nice leg?
Know a blame or a demand.
When does a prince lay a low forest?
When does a length correct a deep mistake?
Write a land and a mouth.
Support a choice and a stream.
Where loves a stone a late hole?
Show a sun and a book.
The role of L1 orthography for spoken word recognition of a closely related L2

This paper will be submitted as

Abstract
Danish and Swedish are closely related languages that are generally mutually intelligible. It has consistently been reported, however, that Danish-speaking listeners decode more Swedish items than vice versa. One reason for this asymmetry might be that spoken Swedish is closer to written Danish than spoken Danish is to written Swedish. We hypothesise here that literate speakers of Danish use their orthographic knowledge of Danish to decode spoken Swedish. L1 Danish speakers were confronted with spoken Swedish in a translation task. The critical manipulation was performed on cognates (words sharing form and meaning across languages), whose pronunciation differed in one phonetic segment only (e.g., the word mild is pronounced /mild/ in Swedish but /mil/ in Danish). Half of the Swedish cognates were pronounced in a way that would be consistent with the spelling of the Danish word (i.e., orthographically consistent cognates), while the other half were pronounced in a way that would not be consistent with the spelling of the Danish word (i.e., orthographically inconsistent cognates). Event-related-brain potentials (ERPs) in the translation task were obtained for these consistent and inconsistent cognates to study the on-line brain responses during decoding operations (i.e., the first 1000ms). Our data showed that ERPs to inconsistent words were significantly more negative than ERPs to consistent words between 750 and 900 ms after stimulus onset. Together with higher word recognition scores for consistent items, our data provides strong evidence that on-line activation of L1 orthography enhances word recognition of spoken Swedish in literate speakers of Danish.

6.1. Introduction
6.1.1. Mutual intelligibility between Danish and Swedish
Danish and Swedish are closely related languages that have been shown to be mutually intelligible to a large extent. There is evidence, however, that mutual intelligibility is asymmetric in that Danish-speaking listeners understand more spoken language items when they are confronted with Swedish than Swedish-speaking listeners do when they are confronted with spoken Danish (Maurud 1976, Bø 1978, Delsing & Lundin Åkesson 2005). To explore this asymmetry in mutual intelligibility, recently, differences in articulation rates have been investigated in a comparative study of spoken Danish and Swedish (Hilton et al. in press). Hilton et al. (in press) found that native speakers of Danish and Swedish produce the same number of phonetically realised syllables per second, while speakers of Danish produce more canonical syllables than speakers of Swedish do. This indicates not only that a specific message is transferred more quickly in Danish than in Swedish, which might have detrimental effects on intelligibility of spoken Danish for Swedish-
speaking listeners, but also that a larger number of reduction processes, such as deletion or assimilation of phonetic segments towards a preceding or following segment, takes place in Danish compared to in Swedish. Phonological reduction is a characteristic of colloquial language use in most languages (in English e.g. [ʃə.ɹjə.ɹi] or even [ʃə.ɹi] for /fju.ɹə.ɹi/, *February*) but it seems that such processes occur with different frequencies across languages. More specifically, the findings reported by Hilton et al. (in press) suggest that they occur more frequently in Danish than in Swedish.

Reduction processes such as schwa-assimilation and the vocalisation of consonants are well-documented phenomena in Danish (Basbøll, 2005; Grønnum, 1998; Grønnum, 2007). What is more, there is evidence that [ɑj] and [aw] are currently becoming subject to monophthongisation and that the unvoiced unaspirated plosives /b/, /d/ and /ɡ/ are increasingly reduced, at least in Copenhagen Danish (Pharao 2010), where e.g. *helt* /heː.ɹi/ (Engl. ‘completely’) is reduced to /heː.ɹ/. Apart from these recent developments, there are many reduction processes in standard Danish that took place several centuries ago, and are so well-established that pronunciation dictionaries only indicate this reduced pronunciation. In the word *mild* /miː.ɹ/ (Molbæk Hansen 1990), for instance, the word final phonetic segment was dropped several centuries ago, while it is preserved in its Swedish cognate word *mild* which is pronounced /mild/ in colloquial Swedish (Hedelin 1997). The words’ orthographic structure is CVCC in both languages, but while the number of segments in spoken Danish has been reduced to three, namely CVC, the number of phonetic segments remains unreduced in Swedish.

Obviously, the phonetic distance between the two spoken items is symmetric, i.e. the phonetic distance remains the same whether we transform Danish /miː.ɹ/ into Swedish /mild/, or vice versa. That means, if solely the phonetic representations of the two items are used for word recognition, a native speaker of Danish has to overcome the same distance when confronted with the non-native item as a native listener of Swedish has. The same goes for the orthographic distances. It can be assumed that, if word recognition relies solely on orthographic input of the two items, Danish-speaking and Swedish-speaking participants encounter the same number of problems when confronted with the neighbouring language. In the *mild* – *mild* case, no problems should occur as the words are spelt in the same way. This is illustrated in Table 1, which schematically sketches the transformation of the spoken words /miː.ɹ/ into /mild/ and vice versa, as well as the transformation of the written words *mild* into *mild* and vice versa. To calculate phonetic and orthographic distances, the Levenshtein algorithm was employed, which identifies the ‘cheapest’ way to

---

4 The usage of IPA-symbols in this article slightly deviates from Danish and Swedish norms as the norms differ slightly across the two languages (cf. section 4.2.2.).

5 The *stød* denoted by the phonetic character // is a supra-segmental feature of the Danish language and is not found in Swedish. It is usually described as a realisation of creaky voice or laryngalisation (Grønnum 1998: 179) and Basbøll (2005: 83). There are monosyllabic and polysyllabic minimal pairs which differ only with regard to absence or presence of *stød*. However, in some Danish regiolects no *stød* is employed. Gooskens & Kürschner (2010) showed that the presence or absence of the *stød* does not significantly impact intelligibility – this is the case for Danish subjects listening to Swedish expecting the *stød* in certain words, as well as for Swedish subjects listening to Danish confronted with an unfamiliar phenomenon.
transform one string into another and counts the number of ‘costs’, i.e. operations (substitutions, insertions and deletions) needed for this transformation. For a detailed discussion of the application of the Levenshtein algorithm for measuring phonetic distances, see Nerbonne & Heeringa (2010).

Table 1. Phonetic and orthographic distance between Danish mild /mil/ and Swedish mild /mild/ as calculated by the simple version of the Levenshtein algorithm.

<table>
<thead>
<tr>
<th></th>
<th>Phonetic distance</th>
<th>Orthographic distance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Danish</strong></td>
<td>m i l’</td>
<td>m i l d</td>
</tr>
<tr>
<td><strong>Swedish</strong></td>
<td>m i l d</td>
<td>m i l d</td>
</tr>
<tr>
<td><strong>Number of transformation operations</strong></td>
<td>0 0 0 1</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td><strong>Levenshtein distance</strong></td>
<td>4/5=25%</td>
<td>0/4=0%</td>
</tr>
</tbody>
</table>

However, if a listener has access to both phonetic and orthographic information, the distance between the two items becomes asymmetric, as listeners can use their native orthography as an additional cue during spoken word recognition (see Table 2). We argue that this is the case for literate listeners.

Table 2. Grapho-phonetic distances between written Danish and spoken Swedish (left) and between written Swedish and spoken Danish (right) for the cognate word pair mild - mild.

<table>
<thead>
<tr>
<th></th>
<th>Danish listeners</th>
<th>Swedish listeners</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L1 orthography</strong></td>
<td>m i l d</td>
<td>m i l d</td>
</tr>
<tr>
<td><strong>L2 pronunciation</strong></td>
<td>m i l d</td>
<td>m i l’</td>
</tr>
<tr>
<td><strong>Number of transformation operations</strong></td>
<td>0 0 0 0 0</td>
<td>0 0 0 1</td>
</tr>
<tr>
<td><strong>Levenshtein distance</strong></td>
<td>1 / 4 = 0%</td>
<td>1 / 4 = 25%</td>
</tr>
</tbody>
</table>

The first three phonetic segments of the spoken forms /mil/ and /mild/ are identical in Danish and Swedish, but while the word ends with /d/ in Swedish, the final plosive has been deleted in Danish. Nevertheless, the phonetic segment /d/ is found frequently in spoken Danish and generally written with the letter d, e.g. in dansk /daensg/, ‘Danish’. Therefore, it can be assumed that Danes hearing the Swedish word /mild/ are able to match the word-final /d/ to the grapheme of their native orthography. In other words, the Danish spelling mild is consistent with Swedish pronunciation /mild/ for literate Danish listeners, while the Swedish spelling mild is considered to be inconsistent with the Danish pronunciation /mil/ to literate Swedish listeners. Table 2 illustrates that the transformation from spoken Swedish /mild/ to written Danish mild is assumed to be less costly than the transformation from spoken Danish /mil/ into written Swedish mild, if native orthography is assumed to be activated during spoken language recognition of the neighbouring language. It can therefore be assumed that literate speakers of Danish match the Swedish word more quickly or more accurately to its native cognate than Swedes do. In this example, Swedish and Danish pronunciation are inconsistent with regard to one segment (the phonetic segment [d]), but, importantly, spoken Swedish /mild/ is consistent with written Danish mild, whereas spoken Danish /mil/ is inconsistent with written Swedish.
Gooskens & Doetjes (2009) calculated grapho-phonetic distances between spoken and written Danish and Swedish on the basis of 86 cognate words and report that spoken Swedish generally is closer to written Danish than spoken Danish is to written Swedish. Following this, it could be assumed that literate Danes generally have a larger advantage from their native orthography when hearing Swedish than literate Swedes have when hearing Danish. This asymmetric advantage from an additional cue could explain the asymmetry in mutual intelligibility between Danish and Swedish-speaking participants reported in a number of previous studies (Maurud 1976, Bø 1978, Delsing & Lundin Åkesson 2005). However, the assumption that L1 orthography is activated during spoken word recognition of a closely related language has not been tested experimentally. If it proofs to be true, it would also predict that spoken word recognition is more symmetric in illiterate than in literate listeners. This prediction is confirmed by data reported by Schüppert and Gooskens (2010) who investigated spoken word recognition of the neighbouring language in Danish and Swedish-speaking pre-schoolers and adolescents. They found that while pre-schoolers performed equally well in a picture-pointing task when confronted with 50 cognate words of the neighbouring language, Danish adolescents and young adults clearly outperformed their Swedish peers in the same task. This suggests that literate speakers of Danish use additional cues in spoken word recognition of Swedish. Their L1 orthographic knowledge could be such an extra cue.

The hypothesis that literate speakers of Danish use their L1 orthography to decode spoken Swedish has not been tested experimentally, but fits in well with findings from experiments that investigated the influence of native orthography for native spoken word recognition. Those studies are summarised in section 6.1.2.

6.1.2. On-line activation of orthography during spoken word recognition

A number of studies have shown that native orthography is activated during native spoken language processing. In a rhyme detection task, for instance, Seidenberg and Tanenhaus (1979) showed that reaction times to orthographically consistent pairs of words, such as pie-tie, were shorter than for orthographically inconsistent pairs of words, such as pie-rye. Jakimik et al. (1985), Slowiaczek et al. (2003), and Chéreau et al. (2007) found that auditory lexical decision responses to targets such as tie were faster when the prime and the target shared both orthography and phonology (e.g., pie-tie) than when they shared only phonology (sigh-tie). In recent studies, Perre and Ziegler (2008), Pattamadilok et al. (2008), and Perre et al. (2009) showed that, in lexical and semantic decision tasks, orthography is activated early, i.e. before lexical access, during spoken word recognition. Perre & Ziegler (2008) also showed that differences in event-related brain potentials (ERPs) between stimuli with multiple possible spellings (‘inconsistent words’) and stimuli with only one possible spelling (‘consistent words’) occurred time-locked to the inconsistency, that is, differences in ERP amplitudes between orthographically consistent and inconsistent items occurred earlier for items that had the orthographic inconsistency in the onset than for words that had the orthographic inconsistency in the rhyme.

These findings all indicate that native orthography is involved in native spoken word recognition. It is likely that this also is the case in non-native word recognition, but to our knowledge, this hypothesis has hitherto not been tested experimentally.
The aim of this paper is therefore to test the hypothesis that L1 orthography is activated during spoken word recognition of a closely related L2. We test this hypothesis by presenting two conditions of spoken Swedish items to literate speakers of Danish in a translation task. The stimuli were manipulated with regard to their L1-grapheme/L2-phoneme consistency in that half of the Swedish cognates were pronounced just like the Danish word was spelt (i.e., orthographically consistent cognates), while the other half was pronounced slightly differently (i.e., orthographically inconsistent cognates).

To investigate the on-line brain processes that occur during the first 1000 ms after stimulus onset, we used ERPs. This is particularly pertinent as the translation task involves a number of processes (such as typing the response) that might potentially mask early and transient consistency effects that occur during perception. On the basis of the results by Perre and Ziegler (2008), we expected to find differences in ERP amplitude between consistent and inconsistent cognates before the N400. However, given that lexical access is delayed in L2 (e.g., Midgley et al. 2010), we expected to find consistency effects in a later time window than in the study by Perre and Ziegler (2008) who investigated consistency effects in L1.

6.2. Experimental Procedure

6.2.1. Participants

Twenty-six Danish-speaking students participated in the experiment. They were recruited from a database of 112 students at the University of Copenhagen, who had completed an online questionnaire. In this questionnaire, they were asked to indicate their L1(s), as well as which other languages they spoke, which languages they had learnt at school or elsewhere, and which languages they did not speak, but were able to understand. These questions were multiple-choice questions and were asked solely to register how much knowledge of Swedish every student had. However, the participants filled out the questions for a set of nine additional European languages whose function was to divert attention from the fact that Swedish was at issue. The participants were also asked to indicate whether they were right or left-handed, had hearing problems, or dyslexia. Seventy students indicated that they were right-handed, neither had hearing problems nor dyslexia, never had learnt the neighbouring language and did not speak it. From these, 26 students participated in the experiment. Among them, 21 (72 percent) of the students indicated that they could not understand the neighbouring language. The participants mainly hailed from the Danish capital of Copenhagen (83 percent). The subjects were paid for their participation, and travel expenses were reimbursed. The participants were 23.5 years old on average and 13 (45%) of the participants were male.

6.2.2. Material

6.2.2.1. Speaker

Standard Danish has a uvular /u/, which phonologically corresponds to standard Swedish alveolar /r/. It is not clear whether Danish listeners interpret the standard
Swedish alveolar /r/ as correspondent to their native grapheme r. To keep this factor constant across languages, the Swedish material is produced by a Southern Swedish speaker, as Southern Swedish varieties have a uvular /ʁ/ similar to the Danish one.

6.2.2.2. Stimuli

Stimuli were 112 closely related cognate words (see Appendix). The stimuli were recorded in a sound-proof room at the Humanities Lab at Lund University. All items were recorded with a sampling frequency of 22050 Hz. The material was selected to form two conditions: words where listeners were expected to have an advantage from their native orthography because their native spelling was consistent with the phonemic realisation in the neighbouring language (O+ condition), and words where the listeners were expected to have no advantage from orthography because their native spelling was inconsistent with the phonemic realisation in the neighbouring language (O− condition). In both conditions, native and non-native pronunciation form minimal pairs that are written in exactly the same way, but differ in their phonemic realisation in exactly one phonetic segment. It was the realisation of this critical phonetic segment that was either consistent or inconsistent with native orthography, thus forming the two conditions.

Table 3. Examples of O+ and O− stimuli. The critical phonetic segments which differ across the languages are underlined. Pronunciation of these critical segments differs across the two languages in all items, but while Swedish pronunciation of the critical segment is consistent to Danish orthography in the O+ items, this is not the case for the O− items.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Danish orthography</th>
<th>Danish pronunciation</th>
<th>Swedish pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>O+</td>
<td>tolv</td>
<td>tål</td>
<td>tolv</td>
</tr>
<tr>
<td></td>
<td>hat</td>
<td>hæ̝d</td>
<td>hatː</td>
</tr>
<tr>
<td></td>
<td>liv</td>
<td>liːw</td>
<td>liːv</td>
</tr>
<tr>
<td></td>
<td>mild</td>
<td>mil</td>
<td>mild</td>
</tr>
<tr>
<td>O−</td>
<td>skærm</td>
<td>sɡæ̝m</td>
<td>ʃæ̝m</td>
</tr>
<tr>
<td></td>
<td>ros</td>
<td>rɔs</td>
<td>ʁɔs</td>
</tr>
<tr>
<td></td>
<td>kemisk</td>
<td>kemsɨ̝ːk</td>
<td>ɕemisk</td>
</tr>
<tr>
<td></td>
<td>gift</td>
<td>ɡifːd</td>
<td>jift</td>
</tr>
</tbody>
</table>

Every minimal pair of stimuli had to fulfil two selection criteria: (1) The two spoken forms differed in exactly one segment, and (2) the spoken non-native form and the
written native form were either consistent (O+ condition) or inconsistent (O- condition) with regard to this single segment. That means that Swedish pronunciation of O+ words was consistent with the Danish spelling of the corresponding cognate.

There were no significant differences across the conditions with regard to frequency, number of phonemes, and duration across languages and conditions. This is indicated in Table 7. As there are no up-to-date frequency lists for Swedish, frequency was assessed by averaging the number of hits in the Internet search engine Google and by averaging the number of hits in the Språkbanken corpus. None of the features differed significantly across the two conditions. In total, the participants heard 56 O+ items, and 56 O– items. For each stimulus, the point of the inconsistency occurrence was determined. This was defined as the onset of the critical phonetic segment and was annotated in Praat (Boersma & Weenink 2008). For plosives, the onset was set between the silent interval and the noise burst.

Table 4. Mean values of stimulus features across conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Frequency (Språkbanken)</th>
<th>Frequency (Google)</th>
<th>No. of phonemes</th>
<th>Duration (ms)</th>
<th>Adjusted duration</th>
<th>Inconsistency point (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O+</td>
<td>1.6</td>
<td>27.1</td>
<td>4</td>
<td>732</td>
<td>639</td>
<td>-</td>
</tr>
<tr>
<td>O–</td>
<td>0.9</td>
<td>17.2</td>
<td>4.1</td>
<td>754</td>
<td>652</td>
<td>105</td>
</tr>
<tr>
<td>5. p</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.2.3. Procedure

The participants were seated in front of a computer screen. The words were presented auditorily in randomised order through loudspeakers placed in front of the participants. The participants were instructed to translate the word they heard into their native language. In addition to the 112 experimental items, 9 items were presented in a training block prior to the experiment. After this training block, participants were asked whether the task was clear and, if necessary, further instructions were given.

Every trial consisted of the following steps. A fixation cross was presented at the centre of the screen for 2000 ms to reduce the amount of head and conscious eye movements. The fixation cross remained on the screen during stimulus presentation and for 1500 ms after stimulus offset. After that, the fixation cross disappeared and a typing mask appeared in the centre of the screen. There was no time limit for typing in the translation but the trial ended when the participants hit the enter key. The participants were instructed to fixate the cross as long as it was presented, then to type in their translation and to blink with their eyes before hitting the enter key, to reduce involuntary eye movements.

6.2.4. EEG Recordings

Continuous electroencephalogram (EEG) was recorded using Geodesic EEG Net Station software version 4.4.2 (EGI, Eugene, OR) with a 128-channel Geodesic Sensor Net (Tucker, 1993). The impedance of all electrodes was kept below 50 kΩ prior to the recordings. All recordings were on-line referenced to Cz, then average-
rereferenced off-line. Bioelectrical signals were amplified using a Net Amps 300 amplifier (EGI, Eugene, OR) and were continuously sampled (24 bit sampling) at a rate of 250 Hz throughout the experiment.

6.2.5. EEG data reduction and analyses

The electrode montage included 4 active midline sensors and 4 sensors over each hemisphere. Analyses were conducted on correct trials solely. EEG data were digitally filtered using a 0.3–30 Hz bandpass filter. Averaged ERPs were formed off-line from correct trials free of ocular and muscular artefact (less than 3% of correct trials were excluded from the analyses).

In total, EEG responses to 112 stimuli were recorded from 26 participants. From the resulting 2912 trials, those containing more than ten bad channels or any ocular and muscular artefacts were excluded from the analysis. A semi-automated rejection algorithm was used for this purpose, together with a visual inspection procedure. Also, trials with incorrect translation were excluded. In total, this resulted in the rejection of 46.0% of the trials for ERP analysis, leaving 1572 trials for the analysis.

Continuous EEG data were divided off-line into epochs beginning 100 ms prior to target presentation and ending 1100 ms post stimulus onset. Data were baseline corrected to a 100-ms pre-stimulus interval. ERPs were calculated by averaging the EEG time-locked to 100 ms before stimulus onset and lasting 2100 ms, i.e. from -100 to 2000 ms. The 100-ms prestimulus period was used as a baseline. Separate ERPs were formed for the two experimental conditions. Artefacts were screened using automatic detection methods (Net Station, Electrical Geodesics, Inc.). Segments containing eye blinks and movement artefacts were excluded from analyses. Bad channel data were replaced using spherical spline interpolation of neighbouring channel values (Perrin, Pernier, Bertrand & Echallier 1989). In order to select appropriate time windows for the ERP analyses, a preliminary stepwise analysis was performed comparing the mean amplitude obtained in the consistent condition with that of the inconsistent condition using pairwise $t$-tests. Mean amplitude was measured for 12 electrode sites (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, O2, see Figure 6) in short and successive epochs, namely 12 epochs with a duration of 50 ms each, in the interval from 500 to 1100 ms after stimulus onset. This procedure allowed us to identify the precise moments at which the consistency effect appeared at different electrodes.

Figure 1. Schematic illustration of the topographic distribution of the twelve analysed sensors.
6.3. Results

6.3.1. Behavioural data

Mean correct reaction time and accuracy for both conditions are displayed in Table 3. Reaction times (RTs) were measured from stimulus onset to response offset, i.e. until the participants hit the enter key after having typed the translation. This resulted in relatively long RTs. RTs larger than three standard deviations (SDs) beyond the global mean of a participant were discarded (7.5% of the data).

An independent t-test revealed that participants decoded items whose pronunciation was consistent with their cognates’ L1 orthography more accurately (63%) than inconsistent items (50%). This difference was highly significant ($t(25) = 4.83, p < .001$). The vast majority of the participants (23 out of 26) showed this effect. There were no significant differences in reaction times, however.

Table 5. Mean correct reaction time and accuracy for both conditions.

<table>
<thead>
<tr>
<th></th>
<th>Accuracy (%)</th>
<th>Correct RT (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistent</td>
<td>63</td>
<td>4451</td>
</tr>
<tr>
<td>Inconsistent</td>
<td>50</td>
<td>4437</td>
</tr>
</tbody>
</table>

6.3.2. EEG data

One of the first visible components was a central-posterior negativity (N1) peaking approximately 170 ms post stimulus onset on Cz. The N1 was followed by a central-posterior positivity that peaked at about 300 ms (P2) at Cz. Following the P2, a negative-going wave was visible on centro-posterior electrode sites that peaked at about 520 ms.

Based upon previous research (e.g. Perre & Ziegler 2008, Perre et al. 2009, Pattamadilok et al. 2008), differences in mean amplitude across the two conditions were expected in the time window 300-350 ms post stimulus onset. However, as consistency effects in this time window are considered to reflect prelexical processing, we expect the effect to occur time-locked to the inconsistency. As the inconsistency in our material occurred somewhat earlier than in the cited studies, we analysed voltages recorded in a larger time window stretching from 200 to 350 ms. A repeated-measures ANOVA with the within-subject factors condition (two levels: O+, O-), laterality (three levels: left hemisphere, midline, right hemisphere) and anteriority (four levels: anterior, central, posterior, occipital) with Greenhouse-Geisser correction yielded no significant main effect for condition and no significant interaction effects of condition with laterality or anteriority, but an interaction effect of all three factors that approached significance ($F(6,150) = 2.33$, Greenhouse-Geisser $p = .08$). Inconsistency tended to elicit a negativity on right hemisphere electrodes (except for O2) and a positivity on occipital electrodes. Mean voltages per condition and electrode site are given in Table 4. These data are visualised in Figure 1, which shows a scalp plot of difference waves (O− voltages minus O+ voltages) at 260 ms after stimulus onset.
Table 6. Mean voltages per condition and electrode site in the time window 200-350 ms post stimulus onset.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Laterality</th>
<th>Anteriority</th>
<th>Voltage (µV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O+</td>
<td>Left</td>
<td>Anterior</td>
<td>1.319</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central</td>
<td>1.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posterior</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Occipital</td>
<td>-1.290</td>
</tr>
<tr>
<td></td>
<td>Midline</td>
<td>Anterior</td>
<td>1.123</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central</td>
<td>1.388</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posterior</td>
<td>-0.428</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Occipital</td>
<td>-1.258</td>
</tr>
<tr>
<td></td>
<td>Right</td>
<td>Anterior</td>
<td>1.276</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central</td>
<td>1.231</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posterior</td>
<td>0.436</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Occipital</td>
<td>-1.349</td>
</tr>
<tr>
<td>O-</td>
<td>Left</td>
<td>Anterior</td>
<td>1.736</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central</td>
<td>1.231</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posterior</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Occipital</td>
<td>-1.000</td>
</tr>
<tr>
<td></td>
<td>Midline</td>
<td>Anterior</td>
<td>1.041</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central</td>
<td>1.633</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posterior</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Occipital</td>
<td>-1.193</td>
</tr>
<tr>
<td></td>
<td>Right</td>
<td>Anterior</td>
<td>1.265</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Central</td>
<td>0.654</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Posterior</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Occipital</td>
<td>-1.031</td>
</tr>
</tbody>
</table>

Figure 2. Voltage map based on difference waves (inconsistent minus consistent) at 260 ms post stimulus onset.

A second time window in which main and interaction effects of condition have been reported (e.g. Perre & Ziegler 2008, Perre et al. 2009, and Pattamadilok et al. 2008) stretches from 500 to 750 ms post stimulus onset. A repeated measures analysis of variance (ANOVA) yielded neither any significant main effect nor interactions in this time window. However, in the following time window from 750 to 900 ms post stimulus onset, the consistent vs. inconsistent condition produced a significant main effect ($F(1,25) = 15.47, p = .001$) and a significant interaction with anteriority ($F(3,75) = 3.64, p = .03$). Mean voltages averaged across all electrode sites were 0.05 µV in the O+ condition and −0.62 µV in the O− condition, indicating a broadly distributed negativity for inconsistent items. Mean voltages per condition and per electrode site are displayed in Table 5, which shows that inconsistency elicited a positivity on anterior electrodes and a negativity on central, posterior, and occipital electrodes in this time window. This is further illustrated by a scalp plot in Figure 2.
Table 7. Mean voltages per condition and electrode site in the time window 750-900 ms post stimulus onset.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Anteriority</th>
<th>Voltage (µV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O+</td>
<td>Anterior</td>
<td>-2.066</td>
</tr>
<tr>
<td></td>
<td>Central</td>
<td>-1.171</td>
</tr>
<tr>
<td></td>
<td>Posterior</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td>Occipital</td>
<td>2.451</td>
</tr>
<tr>
<td>O-</td>
<td>Anterior</td>
<td>-1.948</td>
</tr>
<tr>
<td></td>
<td>Central</td>
<td>-1.974</td>
</tr>
<tr>
<td></td>
<td>Posterior</td>
<td>-0.334</td>
</tr>
<tr>
<td></td>
<td>Occipital</td>
<td>1.778</td>
</tr>
</tbody>
</table>

Figure 3. Voltage map based on difference waves (inconsistent minus consistent) at 820 ms post stimulus onset.

Significance values for pairwise t-tests that were conducted for nine centro-posterior and occipital electrode sites in 16 subsequent time windows with durations of 50 ms each confirmed that a consistency effect was visible in two different time windows. A short positivity restricted to occipital electrodes occurred between 300 and 350 ms post-stimulus onset and a broadly distributed negativity on right and midline central and posterior electrodes was found between 750 and 1000 ms. These results are given in Table 6. Grand average ERPs of the centro-posterior electrodes sites are presented in Figure 3. Figure 4 shows ERPs associated with inconsistent and consistent words at Cz and Figure 5 shows voltage maps based on difference waves (inconsistent minus consistent) at different times during word decoding.

Table 8. Time course of the consistency effect at 16 different electrode sites as confirmed by pairwise t-tests. Grey-shaded cells show significance values where inconsistency produced a negativity, black-shaded cells show significance values where inconsistency produced a positivity.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>O1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>O2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>P4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>O2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. - p > 0.05, * p < 0.05, ** p < 0.01, *** p < 0.001.
Figure 4. ERPs to inconsistent and consistent words at frontal (F3, Fz, F4), central (C3, Cz, C4), posterior (P3, Pz, P4) and occipital (O1, Oz, O2) electrode sites.

Figure 5. ERPs for inconsistent and consistent words time-locked to word onset at Cz.

Figure 6. Voltage maps based on difference waves between inconsistent and consistent words.

6.4. Discussion

The earliest differences between inconsistent and consistent words in our data were found in the 200- to 350-ms window. This finding is in line with Perre and Ziegler (2008), Perre et al. (2009), and Pattamadilok et al. (2008), who typically reported an effect of orthographic inconsistency on centro-posterior and occipital electrodes in a 300-350ms time window. In contrast to these studies, however, inconsistency did not
elicit a negativity but a positivity. This effect only reached significance on two occipital electrodes in a pairwise t-test with condition as independent and voltage as dependent factor.

Between 750 and 900 ms, however, the consistency effect reached significance across the whole time window, was broadly distributed topographically and highly significant, particularly in a smaller window stretching from 800 to 900 ms post-stimulus onset. Here, orthographically inconsistent items evoked significantly lower voltages, i.e. significantly more negative-going potentials on centro-posterior and occipital sites than consistent items did. The voltage differences across the two conditions were large enough as to also produce a significant main effect, i.e. if averaged across the scalp, voltages elicited by inconsistent items were more negative than for consistent items.

The fact that the second effect of consistency appeared about 300 ms later in our data than reported by Perre and Ziegler (2008) and Perre et al. (2009) could be due to the fact that the task was more complex in our experiment. Both cited studies used a lexical decision task in which L1 French-speaking participants decided whether or not a word was a word in their native language, while participants in our experiment were confronted with a foreign language. Furthermore, a translation task can be assumed to require more resources and therefore cause longer latencies than decoding native language stimuli than a lexical decision task. This interpretation is supported by findings by Midgley et al. (2010) who report delayed lexical L2 access in a Go/No-go semantic categorisation task. The longer latency of the consistency effect compared to Perre & Ziegler (2008) could also indicate a late post lexical or decisional locus of the orthographic effect, which would be consistent with Pattamadilok et al.’s (2011) interpretation of an orthographic effect in the time window 375-750 ms. The fact that inconsistent words produced a negativity that was maximal in 800-900 ms time window might be an indication that the inconsistency between L1 orthography and L2 pronunciation is lexically resolved at this stage, or, rather, the consistency between L1 orthography and L2 pronunciation is integrated into lexical retrieval.

Our data confirm that literacy, i.e. access to native orthography, changes the way the brain processes spoken words in a closely related language. More specifically, in the case of literate Danes confronted with spoken Swedish, this access enhances spoken word recognition of the neighbouring language. Our findings indicate that native orthography is not only involved in native language spoken word recognition, but also in non-native word recognition, if the two languages are closely related. Our findings thus expand findings by Perre and Ziegler (2008), Pattamadilok et al. (2009), and Perre et al. (2009), who showed that L1 orthography is activated during L1 spoken word recognition, to the area of L2 processing.

Recalling that Gooskens and Doetjes (2009) reported that spoken Swedish is closer to written Danish than spoken Danish is to written Swedish, our results support the hypothesis that the asymmetry in mutual intelligibility between spoken Danish and Swedish, where Danes have fewer difficulties in decoding spoken Swedish than vice versa (see section 6.1) can at least partly be explained by differences in the depth of the speakers’ native orthographic systems.
As Elbro (2006:33) pointed out, Danish orthography was already old when a national norm was established around the year 1200. This norm reflected several obsolete pronunciations when it was formed, and has not changed dramatically since then. That means that Danish has a more conservative orthography than Swedish has, while at the same time spoken Danish has developed further away from its East Nordic root than spoken Swedish has (Elbro 2006). Findings reported by Pharao (2010) suggest that this process is still productive and that colloquial Danish incorporates more reduction rules today than just a few years ago. The combination of these two factors (conservative orthography and ongoing reduction during the last centuries until today) makes Danish orthography much deeper (i.e. less transparent) than Swedish orthography (Elbro 2006). In line with this, Elley (1992) showed that Danish children have more difficulties acquiring Danish orthography than their peers from other Nordic countries have, and findings by Allerup et al. (2001) and UNICEF (2002) suggest that reading comprehension skills are also poorer in Danish adolescents and adults compared to readers in the other Nordic countries. However, it seems that, once speakers of Danish finally have mastered the complex orthographic system of their native language, it serves as an additional cue for spoken language recognition in Swedish. Our findings emphasise the strong link between literacy and spoken language processing not only for native, but also for non-native target languages.

Acknowledgements

This work was supported by the Netherlands Organisation for Scientific Research (NWO) [grant number 276-75-005]. We thank Tom Sköld for producing our Swedish stimuli, Jan Vanhove for help with the preparation of the material and Richard Andersson for technical support. Furthermore, we are very grateful that Electrical Geodesic, Inc. generously provided us with a Net Station guest license to analyse our data outside the laboratory facilities. Finally, Dries Gankema from the ICT service at the University of Groningen unbureaucratically made the hardware needed for the analysis available to us.

References


Gooskens C. & G. Doetjes. 2009. Skriftsprogets rolle i den dansk-svenske talespråksförståelse [The role of orthography in spoken language comprehension between Danish and Swedish], *Språk och stil* 19, pp. 105-123.


### Appendix

**List of stimuli employed in the experiment.**

<table>
<thead>
<tr>
<th>Danish orthography</th>
<th>O+ phonemic realisation in Southern Swedish</th>
<th>Danish orthography</th>
<th>O- phonemic realisation in Southern Swedish</th>
</tr>
</thead>
<tbody>
<tr>
<td>fisk</td>
<td>fisk</td>
<td>ske</td>
<td>fje:</td>
</tr>
<tr>
<td>drift</td>
<td>drifth</td>
<td>skærm</td>
<td>fjærm</td>
</tr>
<tr>
<td>glimt</td>
<td>glimt</td>
<td>skælmi</td>
<td>fjelm</td>
</tr>
<tr>
<td>brist</td>
<td>bøist</td>
<td>skægles</td>
<td>fjegles</td>
</tr>
<tr>
<td>disk</td>
<td>disk</td>
<td>skæg</td>
<td>fjæg:</td>
</tr>
<tr>
<td>kvist</td>
<td>kvist</td>
<td>skyldig</td>
<td>fjylidi</td>
</tr>
<tr>
<td>frisk</td>
<td>frisk</td>
<td>skæmt</td>
<td>fjæmt</td>
</tr>
<tr>
<td>finsk</td>
<td>finsk</td>
<td>sko</td>
<td>squr</td>
</tr>
<tr>
<td>tid</td>
<td>tid</td>
<td>stol</td>
<td>squrl</td>
</tr>
<tr>
<td>død</td>
<td>død</td>
<td>sol</td>
<td>sul</td>
</tr>
<tr>
<td>syd</td>
<td>syd</td>
<td>ros</td>
<td>usu</td>
</tr>
<tr>
<td>led</td>
<td>leød</td>
<td>okse</td>
<td>ugso</td>
</tr>
<tr>
<td>gled</td>
<td>gled</td>
<td>ost</td>
<td>ust</td>
</tr>
<tr>
<td>smed</td>
<td>smed</td>
<td>porto</td>
<td>puto</td>
</tr>
<tr>
<td>ned</td>
<td>ned</td>
<td>pol</td>
<td>poul</td>
</tr>
<tr>
<td>gned</td>
<td>gned</td>
<td>fiol</td>
<td>fiul</td>
</tr>
<tr>
<td>mild</td>
<td>mild</td>
<td>viol</td>
<td>viul</td>
</tr>
<tr>
<td>pund</td>
<td>pond</td>
<td>klo</td>
<td>kluc</td>
</tr>
<tr>
<td>mætte</td>
<td>mættel</td>
<td>sno</td>
<td>smuc</td>
</tr>
<tr>
<td>flotte</td>
<td>fløtta</td>
<td>ko</td>
<td>ku:</td>
</tr>
<tr>
<td>jette</td>
<td>jetta</td>
<td>tro</td>
<td>tru:</td>
</tr>
<tr>
<td>hæfte</td>
<td>hefta</td>
<td>ro</td>
<td>ru:</td>
</tr>
<tr>
<td>loftte</td>
<td>loftta</td>
<td>bro</td>
<td>bu:</td>
</tr>
<tr>
<td>kniv</td>
<td>kniv</td>
<td>bo</td>
<td>bu:</td>
</tr>
<tr>
<td>elev</td>
<td>elev</td>
<td>glo</td>
<td>gluc</td>
</tr>
<tr>
<td>slev</td>
<td>slev</td>
<td>tempo</td>
<td>tempo</td>
</tr>
<tr>
<td>drev</td>
<td>dræv</td>
<td>giraf</td>
<td>fjaøf:</td>
</tr>
<tr>
<td>liv</td>
<td>liv</td>
<td>kemisk</td>
<td>cemisk</td>
</tr>
<tr>
<td>dov</td>
<td>dov</td>
<td>kemi</td>
<td>cemi</td>
</tr>
<tr>
<td>lev</td>
<td>lev</td>
<td>kiosk</td>
<td>ciosk</td>
</tr>
<tr>
<td>rev</td>
<td>øenv</td>
<td>kælpe</td>
<td>cælø</td>
</tr>
<tr>
<td>brev</td>
<td>bøev</td>
<td>kæmpø</td>
<td>cæmpø</td>
</tr>
<tr>
<td>aktiv</td>
<td>agtv</td>
<td>kærv</td>
<td>cærv</td>
</tr>
<tr>
<td>skrev</td>
<td>sgøev</td>
<td>kilo</td>
<td>cilo</td>
</tr>
<tr>
<td>greve</td>
<td>greve</td>
<td>soprøn</td>
<td>sopøn</td>
</tr>
<tr>
<td>ræv</td>
<td>ræv</td>
<td>parat</td>
<td>parøt</td>
</tr>
<tr>
<td>kvist</td>
<td>kvist</td>
<td>karat</td>
<td>kæøt</td>
</tr>
<tr>
<td>kat</td>
<td>kat</td>
<td>kran</td>
<td>kran</td>
</tr>
<tr>
<td>nat</td>
<td>nat</td>
<td>pirat</td>
<td>piaøt</td>
</tr>
<tr>
<td>das</td>
<td>das</td>
<td>nitrat</td>
<td>niøt:</td>
</tr>
<tr>
<td>pas</td>
<td>pas</td>
<td>gratis</td>
<td>gæøt:</td>
</tr>
<tr>
<td>hal</td>
<td>hal</td>
<td>drama</td>
<td>drøma</td>
</tr>
<tr>
<td>plat</td>
<td>plat</td>
<td>moral</td>
<td>moøgel</td>
</tr>
<tr>
<td>skat</td>
<td>skat</td>
<td>klar</td>
<td>klaøt</td>
</tr>
<tr>
<td>debat</td>
<td>debat</td>
<td>graf</td>
<td>graøt</td>
</tr>
<tr>
<td>glans</td>
<td>gløst</td>
<td>tjekkisk</td>
<td>cækkisk</td>
</tr>
<tr>
<td>hat</td>
<td>hat</td>
<td>gømøn</td>
<td>jæøm</td>
</tr>
<tr>
<td>tolv</td>
<td>tolv</td>
<td>gødsøl</td>
<td>jøsøl</td>
</tr>
<tr>
<td>gulv</td>
<td>gølv</td>
<td>gift</td>
<td>jif:</td>
</tr>
<tr>
<td>stolpe</td>
<td>stølpe</td>
<td>giftig</td>
<td>jiftí</td>
</tr>
<tr>
<td>droppe</td>
<td>droppe</td>
<td>gehøð</td>
<td>jehøð</td>
</tr>
<tr>
<td>tanke</td>
<td>tanke</td>
<td>gæs</td>
<td>jes:</td>
</tr>
<tr>
<td>nakke</td>
<td>nakke</td>
<td>gæst</td>
<td>jest</td>
</tr>
<tr>
<td>virke</td>
<td>virke</td>
<td>igen</td>
<td>ijen</td>
</tr>
<tr>
<td>gips</td>
<td>gips</td>
<td>jips</td>
<td>jips</td>
</tr>
</tbody>
</table>
Conclusion

The papers forming this thesis investigate the asymmetry in mutual intelligibility between spoken Danish and Swedish. It has been reported consistently in previous research that speakers of Danish decode more spoken Swedish than vice versa. Previous investigations, however, have focused on extra-linguistic factors such as the amount of language contact across the border and different attitudes held towards the neighbouring language, but could not always confirm convincing links between intelligibility and these factors. Furthermore, the causal relationship between these extra-linguistic factors and intelligibility is still unclear.

Initially, we hypothesised that linguistic factors also play an important role for this asymmetry, as most of the suggested factors in previous studies either did not correlate significantly or could only explain a small amount of variation (cf. the correlation coefficients’ significance values in Delsing & Lundin Åkesson 2005). To test this hypothesis, we conducted an intelligibility experiment with Danish and Swedish illiterate pre-schoolers from outside the border regions (namely from Odense kommune in Denmark and from Växjö kommun in Sweden), in which extra-linguistic factors such as the amount of contact and language attitudes were kept constant across the two groups of participants (Danish and Swedish). At the same time, linguistic factors such as the distributions of the Swedish tonemes and the Danish stød in the material, differences in word length and duration were kept similar to Gooskens & Kürschner’s (2008) study in which a significant asymmetry in intelligibility scores was reported. We expected that the various relevant linguistic differences such as word length, the number of phonetic neighbours, speaking tempo or the supra-segmental features stød and the tonal accents would cause asymmetric intelligibility results in our participants. However, this turned out not to be the case. Danish and Swedish pre-schoolers performed equally well on a picture-pointing word-recognition task with 50 auditorily presented cognate words in their neighbouring language. Contrary to our initial hypothesis, we had to conclude that the asymmetry in mutual intelligibility is caused by extra-linguistic factors after all. This experiment and its conclusions are reported in chapter 2.

Subsequently, the role of language attitudes was investigated more closely. In the paper reported in chapter 3, the data from the picture-pointing word-recognition task reported in chapter 2 was extended with data from more pre-schoolers and, additionally, with data from adult participants. After the word-recognition experiment, the participants were asked how they thought the neighbouring language sounded compared to their native language. Reaction times to correctly identified stimuli were analysed and correlated with self-reported language attitudes held towards the neighbouring language (Danish for Swedish participants and vice versa). It turned out that attitudes held by children generally were more positive than the attitudes held by adults, a tendency which was particularly pronounced in the Swedish participants. Swedish adults held a significantly more negative attitude
towards Danish when asked overtly how they liked the language they heard in the experiment. However, contrary to our reformulated hypothesis, there was no significant correlation between intelligibility score and attitude towards the neighbouring language when the age factor was controlled for. From this experiment, we could not confirm that language attitudes play a role for intelligibility. Based on this data, language attitudes do not explain the asymmetry in mutual intelligibility between spoken Danish and Swedish.

A restriction in this experiment was the fact that we did not control for individual speech properties in our two speakers such as the liveliness of the intonation, the general pitch employed, or speaking tempo etc. It is likely that the elicited attitudes are partly influenced by speaker-specific speech features and therefore are not comparable. To investigate the link between language attitudes and intelligibility more accurately, and to avoid the shortcomings from the experiment reported in chapter 3, we conducted a matched-guise experiment, which is reported in chapter 4. Seven- to 16-year-old participants hailing from similar geographical areas as participants who participated in the experiments reported in chapters 2 and 3 participated in the same word recognition experiment. However, this time, we elicited their subconsciously held attitudes towards the neighbouring language before their word recognition was tested. This was done by presenting them with the same text recorded in six different languages (Danish, Swedish, Norwegian, Dutch, Frisian and Indonesian) and instructing them to indicate on a five-point scale how normal, modern, beautiful, kind, smart and rich they thought the speaker was. Importantly, the Danish and the Swedish recordings were produced by the same speaker, a Danish-Swedish bilingual. It was ensured that the participants did not discover this. In this experiment, we took care not to ask the participants' attitude towards the language in question in an open manner, but to have the participants judge personality traits of all five speakers instead. As they judged the bilingual speaker twice, we could draw conclusions as to their attitudes regarding the neighbouring language. We found a significant but low correlation between intelligibility and language attitudes, which explained merely 3.6% of their word recognition variance. Interestingly, the Danish participants clearly outperformed the Swedish participants in the intelligibility experiment once again, thereby confirming results from Maurud (1976), Bø (1978), Delsing & Lundin Åkesson (2005), and our results reported in chapter 3. Again, we had to modify our hypothesis about which factors play an important role for the asymmetry in mutual intelligibility between spoken Danish and Swedish and had to conclude that language attitudes are of no major importance for intelligibility and, hence, cannot explain the asymmetry in mutual intelligibility.

In the work reported in chapter 5, we turned to the investigation of two linguistic factors, namely articulation rate and the degree of reduction in spoken language. Having observed that all Danish stimuli in our previous experiments tended to have a shorter duration than the Swedish stimuli, although the number of orthographic syllables generally seemed similar across the languages, we hypothesised that Danes generally communicate more quickly than Swedes do. If this
turned out to be the case, it is likely to impair the intelligibility of spoken Danish compared to spoken Swedish. We tested this hypothesis by measuring the number of canonical syllables and the number of phonetically realised syllables in two corpora, namely in radio news broadcasts and in recordings of semantically unpredictable sentences made in a highly controlled setting. In this experiment, we also included Norwegian, although this is not the focus of this thesis. We found that speakers in our Danish corpora generally produced more orthographic syllables per time unit than the speakers in the Swedish corpora did. At the same time, they did not produce significantly more phonetic syllables, which were defined as intensity peaks in voiced sounds. This means that Danes generally manage to transfer a specific message faster than Swedes do, without increasing their speaking rate. We assumed that this can only be done if some information is deleted or reduced, e.g. by assimilating sounds towards the surrounding phonetic segments, by leniting consonants to approximants or by deleting certain segments completely. By subtracting the number of phonetic syllables produced per time unit from the number of canonical syllables produced per time unit, we calculated a reduction ratio for all three languages. It turned out that the reduction ratio is significantly higher in Danish speakers than in Norwegian and Swedish speakers. That means that spoken Danish in our corpora was significantly more reduced than Swedish and Norwegian, which is likely to have a detrimental effect on the intelligibility of spoken Danish. These findings thus could be part of the explanation why mutual intelligibility between Danish and Swedish adults is typically reported to be asymmetric. However, it is not clear why Danish and Swedish preschoolers’ intelligibility of the neighbouring language was found to be symmetrical. This question was addressed in the following paper.

In the work reported in chapter 6, we investigated the link between literacy and intelligibility of the neighbouring language. Following up on the findings reported in chapter 5, we investigated the role of well-established reduction processes in spoken Danish, which are not reflected in the orthographic system. Danish orthography has been described as particularly conservative, both compared to its pronunciation (orthography is generally considered more conservative than the spoken form of the same language, Elbro 2006) as well as compared to its neighbouring languages (Danish orthography is more conservative than Swedish orthography is, Elbro 2006). The orthographic system represents pronunciation which was used several centuries ago (such as the spelling mild for pronunciation [mil], where the word final [d] had been dropped entirely in the spoken form of the word, while it is still preserved in its written form). At the same time, spoken Danish has developed further away from its East Nordic root than spoken Swedish has (Elbro 2006), which generally makes Danish orthography deeper (i.e. less transparent) than Swedish orthography. Figure 1 shows a schematic illustration of the distances between the two spoken and written systems.
Figure 1. Schematic illustration of the distances between spoken and written Danish and Swedish.

Figure 1 illustrates that the distance between the two orthographic systems is smaller than between the spoken forms of the languages. This is in line with the intelligibility results reported by Maurud (1976), Bo (1978) and Delsing & Lundin Åkesson (2005), who showed that the mutual intelligibility between written Danish and Swedish is higher than the mutual intelligibility of spoken Danish and Swedish. It can also be seen from Figure 1 that Danish orthography is deeper (i.e. less transparent) than Swedish orthography is (Elbro 2006), as the distance between spoken and written Danish is larger than the distance between spoken and written Swedish. Finally, Gooskens & Doetjes’ (2009) cross-linguistic calculations of grapheme-sound correspondences showed that the distance between spoken Danish and written Swedish is larger than the distance between spoken Swedish and written Danish. This is illustrated by the two one-headed arrows crossing each other. The assumption is that literate speakers of Danish can use their orthographic knowledge from their conservative orthography to decode spoken Swedish. Although it has been reported that native orthography indeed is activated during native language spoken word recognition (Perre and Ziegler 2008, Pattamadilok et al. 2009, Perre et al. 2009) it has not been shown whether this is also the case when listeners are confronted with a foreign language. In chapter 6, we therefore investigate whether literate speakers of Danish can use their native orthography as an additional cue when confronted with Swedish. Results revealed that this is the case, as their electrophysiological responses to words where (their native) Danish orthography was consistent with Swedish pronunciation (such as mild [mild]) were significantly different than signals to words where Danish orthography and Swedish pronunciation were inconsistent (such as gift [jift]). More specifically, inconsistent items elicited a broadly distributed negativity on centro-posterior and occipital electrodes in the time window 750-900 ms post-stimulus onset, which was interpreted as neural correlates of the resolution of the inconsistency between L1 orthography and L2 pronunciation, or the integration of the consistency between L1 orthography and L2 pronunciation into lexical retrieval. Furthermore, the participants’ word recognition scores were higher for consistent than for inconsistent words. This provides strong evidence that Danish listeners have an advantage from their native orthography when confronted with spoken Swedish. As this is not the case for illiterate pre-schoolers, our data suggested
that on-line activation of L1 orthography is part of the explanation of the well-documented asymmetry in mutual intelligibility between spoken Danish and Swedish.

The papers in this thesis indicate that both extra-linguistic as well as linguistic factors are linked to the asymmetry in mutual intelligibility between spoken Danish and Swedish. If other factors are controlled for, *language attitudes* have a small, yet significant impact on the degree of intelligibility of Danish to Swedish-speaking listeners and of Swedish to Danish-speaking listeners. More specifically, the more positive listeners think about the speaker, the greater efforts they seem to make in understanding the language he or she speaks.

Literacy usually is considered an extra-linguistic factor as it is based in the subjects and not in the material. However, it is not literacy alone which has an impact on mutual intelligibility between Danes and Swedes. Instead, it is the degree of orthographic conservatism, along with the amount of reduction in the spoken form which seems to have an effect on the degree of intelligibility of Danish and Swedish. These are clearly linguistic factors and they are likely to account for a part of the variance in mutual intelligibility. More specifically, a conservative native orthography and a conservative pronunciation of a closely related language together work in favour of the listeners, while a progressive orthography which is updated regularly give listeners fewer cues when they are confronted with the spoken form of a progressive closely related language. In our case, it turned out that literate Danes can use their orthographic knowledge when decoding spoken Swedish.

The results from the last experiment reported in chapter 6 fit nicely with the results from the first experiment reported in chapter 2, where we found that mutual intelligibility of spoken Danish and Swedish is symmetrical in illiterate pre-schoolers. This supports the assumption that the access to orthographic knowledge plays a major role for the intelligibility of Danish and Swedish in Scandinavian listeners by serving as an extra cue during spoken language recognition. The distance between the two spoken language forms is naturally symmetric, but this extra cue facilitates word recognition. As the degree of facilitation is not equal in the two groups of listeners (Danish and Swedish), it makes intelligibility scores asymmetric in literate listeners. Importantly, as Gooskens & Doetjes (2009) showed, Danish listeners benefit from it to a larger extent than Swedish listeners do. If we turn this finding around, the intelligibility of spoken Danish is impeded by the large amount of reduction and assimilation processes in contemporary Danish, which are unrecognisable for Swedish listeners as they are neither reflected in Swedish pronunciation nor orthography.

It can be concluded that two factors which are interlinked with each other account for a large part of the asymmetry in mutual intelligibility between spoken Danish and Swedish. Firstly, Danish pronunciation has changed substantially during the past centuries. Among other things reduction is reflected in the large number of processes such as schwa-assimilation (Basbøll 2005) and consonant lenition and deletion (Pharao 2010). This development has lead to a articulatory drifting apart of
spoken Danish from its East Nordic root on the one hand, and from its neighbouring languages on the other hand. The high number of reduction processes goes hand in hand with a higher articulation rate in Danish compared to Swedish and leads to a large phonetic distance of cognate words across the two languages. This makes these cognate words less intelligible to pre-schoolers, and, as they have to overcome the same (symmetric) distance, they encounter the same problems during spoken word recognition of the neighbouring language. Secondly, many of these historical reduction processes are not reflected in Danish spelling, which makes the distance between Danish orthography and Swedish pronunciation smaller than the distance between Swedish orthography and Danish pronunciation. Literate speakers of Danish therefore can use their orthographical knowledge when confronted with spoken Swedish as an additional cue, which results in asymmetric intelligibility compared to literate speakers of Swedish.

By investigating which factors cause the asymmetry in mutual intelligibility between spoken Danish and Swedish, this thesis sheds light on the factors that play a role for language perception in general, and intelligibility of a closely related language in particular. Future work should reinvestigate the symmetry in intelligibility scores that we reported for pre-schoolers in an experiment which controls for reduction processes in the stimulus material in a more structured manner. In this way, the relative contribution of linguistic and extra-linguistic factors could be investigated quantitatively.

Furthermore, the link between reduction and intelligibility should be investigated more closely in order to tease apart the factors speaking tempo and reduction. This could be done by reducing the duration of utterances recorded with clear and slow speech (thereby forming a condition with high speaking tempo but little reduction) and, likewise, by extending the duration of utterances recorded with reduced and quick speech (thereby forming a condition with low speaking tempo but relatively much reduction). By analysing how intelligible utterances from these four conditions are to native speakers as well as speakers of the neighbouring language, the relative contribution of reduction and speaking tempo could be established.

Finally, a thorough investigation of mutual intelligibility between other closely related languages could give valuable information on the role of different linguistic and extra-linguistic factors. Asymmetric intelligibility scores have been reported anecdotally for the Nigerian languages Kalabari and Nembe (Wolff 1959), and experimentally for Spanish and Portuguese (Jensen 1989), Dutch and German (Gooskens, Van Bezooijen & Van Heuven submitted), Dutch and Afrikaans as well as Frisian and Afrikaans (Gooskens 2007). Studying factors that cause the asymmetric intelligibility in other closely related languages is likely to shed more light on the processes underlying the asymmetry in mutual intelligibility between spoken Danish and Swedish, and thereby on factors influencing intelligibility and language perception in general.
References


