

The influence of extra-linguistic factors on mutual intelligibility: Some preliminary results from Danish and Swedish pre-schoolers

Anja Schüppert & Charlotte Gooskens

University of Groningen, The Netherlands

1. Introduction

Communication between speakers of the closely related Scandinavian languages often takes place in the language of the speaker rather than in English as a *lingua franca*. Also scientifically, it has been shown that Danish, Norwegian and Swedish are mutually intelligible to a large extent in adults (Bø 1978, Delsing & Lundin-Åkesson 2005, Maurud 1976). However, in previous investigations, mutual intelligibility between Danish and Swedish has been shown to be asymmetrical: Danes have fewer problems decoding Swedish than Swedes have decoding Danish.

A number of factors have been suggested to cause the asymmetry in spoken language (Bø 1978, Delsing & Lundin-Åkesson 2005, Gooskens 2006, Gooskens & Doetjes accepted, Maurud 1976). These factors can be divided into linguistic and extra-linguistic factors. Linguistic factors are language-inherent features of the language itself, such as supra-segmental features, average word length, and speech tempo. Danish and Swedish have common roots, but differ in many of these respects. Still, in previous research, mostly extra-linguistic factors have been considered to cause the asymmetry in mutual intelligibility between these two languages. We hypothesize that the asymmetry is mainly caused by linguistic factors. The aim of the experiment reported in this paper was therefore to exclude the influence of extra-linguistic factors on mutual intelligibility (see Section 2), and to compare our intelligibility scores to scores from previous studies, that have not done so. To place our experiment in perspective, we will start by giving an overview of the three mainly discussed extra-linguistic factors.

1.1 Attitude

In his investigation of two Nigerian tribes, Wolff (1966) showed that attitude towards a closely related variety may influence the effort the listeners makes to decode it, in the sense that a negative attitude would result in fewer efforts, whereas a positive attitude might encourage listeners to do their best in decoding it. Sweden has been and is still the Scandinavian country with the largest population and the strongest industry. Within Scandinavia, Sweden is often called *storebror* ('big brother'), indicating that Norway and Denmark might regard Sweden as more influential than their own country. The fact that Sweden

has a higher status in Scandinavia than Denmark might cause a bias in the willingness to understand the other variety, and thereby a bias in actual intelligibility.

Delsing & Lundin Åkesson (2005) gathered attitude and intelligibility scores from young Danes and young Swedes. All subjects were asked two questions to elicit their attitude towards the neighbouring country: (a) Do you think Danish/Swedish sounds beautiful? (b) Would you like to live in Denmark/Sweden? They showed that Danish-speaking subjects rated the Swedish language as more beautiful than Swedish-speaking subjects rated the Danish language. However, when the same subjects were asked if they would like to live in the neighbouring country, it turned out that Danes were less willing to move to Sweden than vice versa. Interestingly, Danish intelligibility scores were better predicted by the subjects' answer to question (a), whereas Swedish intelligibility scores were better predicted by the answer to question (b). The authors do not give an explanation for that. Gooskens (2006) re-investigated the relationship between attitude and intelligibility in the Danish and Swedish subjects studied by Delsing & Lundin Åkesson (2005). Whereas question (b) did not correlate significantly with intelligibility, Gooskens found a significant correlation between the answers to question (a) and the subjects' intelligibility scores ($r = .56, p = .02$). Generally, we conclude that asymmetry in attitude may cause an asymmetry in intelligibility.

1.2 Previous exposure

In earlier investigations, it has been assumed that previous exposure through watching tv, reading newspapers, visiting the neighbouring country or other forms of personal contact enhances the abilities to decode the variety in question. However, results from experimental research are contradictory.

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 had access to television programmes in the neighbouring variety, too. Bø 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.

Gooskens (2006) correlated intelligibility scores elicited by Delsing & Lundin Åkesson (2005) with four different contact scores (tv, newspapers, personal contact, visits), elicited by the same authors. In contrast to Bø (1978), she found no significant correlation between any of these contact forms and intelligibility scores. She explains this by the fact that the contact scores were very low in all subjects, i.e. the listeners in general had had little contact with the neighbouring

countries. This might indicate that a substantial amount of previous exposure is needed to enhance intelligibility of a closely related variety. However, since results are contradictory, it still has to be considered that the asymmetry in intelligibility might be caused or boosted by an asymmetry in the amount of contact with the neighbouring country.

1.3 Literacy

Written Danish and Swedish are rather similar and mutual intelligibility of written items is high and symmetrical: Danes decode around 56 percent of written Swedish, and Swedes decode around 54 percent of written Danish (mean figures from Bø, 1978; Delsing & Lundin-Åkesson, 2005; Maurud, 1976). Danish and Swedish also share a large part of their vocabularies, i.e. there is a great number of Swedish-Danish cognate word pairs. However, contemporary Danish and Swedish differ considerably in pronunciation of these cognates. Danish is characterised by a large amount of lenition and a great number of approximants. Consonants that appear in contemporary written Danish and Swedish are often pronounced clearly in Swedish, but lenited or left out in Danish. This results in an “opaque phoneme to grapheme relation in Danish” (Bleses & Thomsen 2004:79).

As can be seen from Table 1, Danish and Swedish have a roughly common orthography. Since the Danish pronunciation has developed further away from the former near-common pronunciation, there is a large distance between contemporary spoken and written Danish. This difference is smaller in Swedish, where the spoken language has developed less rapidly and the Swedish orthography reflects the pronunciation more accurately than in Danish.

Danish		Swedish	
orthography	pronunciation (IPA)	orthography	pronunciation (IPA)
<i>hoved</i>	[ho:(\square)əð]	<i>huvud</i>	[h \square :v \square d]
<i>stjerne</i>	[sdj \square \square \square :nə]	<i>stjärna</i>	[\square æ: \square a]

Table 1. Two words (‘head’ and ‘star’) written in Swedish and Danish contemporary orthography and IPA symbols to reflect pronunciation.

In the first example in Table 1, the Swedish word *huvud* and the Danish word *hoved* (‘head’) are both written CVCVC. The Swedish pronunciation is CVCVC, too, whereas the Danish pronunciation comes close to CV(V)VV, since the approximants [\square] and [ð] are (semi)vowels. The letter *v* is pronounced as a voiced labiodental fricative in Swedish, but as a labiodental approximant in Danish and can hardly be perceived at all in this word. The final letter *d* is pronounced as a voiced dental stop in Swedish, but as an alveolar approximant in

Danish. It can be assumed that a Dane has advantages when decoding the Swedish word because the *v* is pronounced as a voiced fricative at word onset in Danish, whereas a Swede has less support from the Swedish orthography when he hears the Danish pronunciation without the two consonants. This is especially important if we consider the findings by Van Ooijen (1994) and Gooskens, Beijering & Heeringa (submitted) that consonants are more important for the intelligibility of a word than vowels are. On the other hand, as illustrated in Table 1, the Danish word *stjerne* ‘star’, pronounced with the word initial consonant cluster [sdj], is presumably easier for a Swede to understand 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], whereas it can be assumed that for a Swede it is not equally unexpected that *stj* is pronounced [sdj], because this comes very close to an accumulated pronunciation of the three isolated phonemes [s][t][j].

Gooskens & Doetjes (accepted) calculated the distances between Swedish and Danish pronunciation and orthography using the Levenshtein algorithm (for details see Heeringa 2004). They showed that the orthographic distance between Danish and Swedish is 24 percent, and the phonetic distance between the two languages is 53 percent. They also calculated the distances taking into account the advantage that Danes and Swedes have from their native orthography and show that both Danish and Swedish literate listeners have advantage of their reading and writing skills when confronted with the neighbouring language. However, adult Danish listeners generally have a bigger advantage of their literacy when confronted with spoken Swedish than adult Swedish listeners have when confronted with spoken Danish. In other words, the *hoved–huvud* example given in Table 1 is more representative for the Swedish-Danish spelling-pronunciation situation than the *stjärna–stjerne* example. This leads to the fact that adult speakers of Danish have a larger advantage when listening to Swedish, than Swedes have when listening to Danish. The findings by Gooskens & Doetjes therefore might indicate that the asymmetry in mutual intelligibility is caused by an asymmetric degree of profiting from the native writing system.

2. Research question

The aim of the experiment reported in this paper was to exclude the influence of extra-linguistic factors on intelligibility and thereby investigate the role that these factors play for mutual intelligibility. We hypothesized that linguistic factors account for a large part of the asymmetry and therefore expected the asymmetry in mutual intelligibility to be present to a similar extent even when extra-linguistic factors are controlled for. To test this hypothesis, we conducted a word comprehension experiment eliciting intelligibility scores from a group of

subjects that can be considered to be neutral with respect to the three extra-linguistic factors, because they (a) did not have a specific attitude towards the test language, (b) had not been previously exposed to the test language, and (c) were illiterate: pre-schoolers from outside the border regions. We assumed that young children do not have specific attitudes towards the neighbouring languages, but, as an extra precaution, we asked the children questions to elicit their attitudes. Furthermore, we expected pre-schoolers to have had less contact with the neighbouring language than adults, especially if children's movies are dubbed. However, this might not always be the case in the Scandinavian countries, so this assumption was asserted individually, too. Still, we assumed it to be easier to find children that have not had contact with the neighbouring language, than adult subjects. Finally, pre-schoolers that are not able to read and write cannot make use of the orthographic advantage.

By eliciting data from 5- to 6-year old children, we aimed at excluding the influence of extra-linguistic factors on the intelligibility asymmetry between Danish and Swedish. To our knowledge, neither in Scandinavia, nor in other language areas, children's intelligibility of closely related languages has been investigated.

3. Method

3.1 Subjects

We tested 16 Danish- and 20 Swedish-speaking 5- to 6-year-old preschoolers from outside the border regions. In order to exclude children that might have acquired some basic phono-graphemic skills for their native language, a questionnaire was filled in by the parents or care-takers of every participating child, asking if the child had learnt to read and write and, if so, to indicate how many words the child could write. Furthermore, after the experiment, the children were asked if they liked the language they had just heard (see section 3.3).

After questionnaire evaluation, three Swedish children had to be excluded because their parents indicated that they could write "many" words or "almost everything"¹, and one Swedish child had to be excluded due to extensive exposure to the Danish language, leaving 16 Danish and 16 Swedish children for the analysis. No children were excluded on the basis of their attitude. The Danish children ranged in age from 5.0 to 6.7 ($\mu = 5.9$, $sd = 0.6$, $SE = 0.1$), the remaining Swedish children were aged between 5.0 and 6.6 ($\mu = 5.9$, $sd = 0.4$, $SE = 0.1$).²

3.2 Stimulus material

The experiment was programmed and run in E-Prime 2.0. The auditory stimulus material consisted of 53 highly frequent Swedish-Danish cognate nouns. These

nouns had been proven to be highly congruent labels of pictures that were shown to five Danish and five Swedish four-year-old children in a pre-test. Only pictures that had a labeling consistency of at least 80 percent in this pre-test were included. By this, it was made sure that the 53 target pictures (i.e. the pictures that corresponded to the 53 auditory stimuli) were recognized and produced by children even younger than the age group tested in the experiment and labelled almost unambiguously by these children.

These labelings (henceforth ‘stimuli’) were read aloud by two female native speakers from Odense (Denmark) and Växjö (Sweden) respectively and presented twice with an interstimulus interval of 3000 ms. An analysis of the linguistic features of the stimulus material revealed that 43.4 % of the Danish words contained the typical Danish feature ‘stød’, which is a kind of creaky voice and 34.0 % of the Swedish words had accent 2, which is one out of two possible tonal contours in Swedish. Both features are generally regarded as marked. Mean word duration was 560 ms for Danish and 800 ms for the Swedish items. This is a significant ($t(49) = -8.36, p < .001$) but representative difference (see Schüppert & Gooskens, in preparation). The mean number of syllables and segments did not differ significantly between the two languages. For a detailed discussion of the influence of these linguistic features on word recognition see Kürschner, van Bezooijen & Gooskens (in press) and Schüppert & Gooskens (submitted).

Three of the stimuli were used in a demo version, leaving 50 stimuli for the experiment. Simultaneously to the 50 auditory stimuli, 200 pictures were presented visually to the subjects, so that every trial consisted of an auditory presentation of one stimulus and a visual presentation of four pictures randomly chosen. The pictures appeared at word onset and remained on the screen until the end of the trial. The pictures were taken from the picture database developed at the Max-Planck-Institute for Psycholinguistics in Nijmegen, the Netherlands. The stimulus material was presented randomised, but every stimulus was presented together with the same set of four pictures in every session and across languages.

3.3 Procedure

The testing session consisted of a stimulus-response experiment followed by a short interview with every child.

Before the experiment started, the children were familiarised with the task by being presented a demo version of the experiment. The child sat in front of a touch screen (LG L1510SF) wearing ear phones. During the demo version, two audio files in the children’s native language were presented, followed by one audio file in the test language. Simultaneously, four pictures per stimulus were presented on the touch screen. The children were shown how to select the corresponding picture. After the demo version, the experiment started. The

experimental design and the children's task were the same as in the demo version.

After the experimental part, the children were asked the following question in order to determine their attitude towards the test language: „The language you heard in the experiment was Danish/Swedish. Do you think it sounded (1) less nice than Swedish/Danish (i.e. native language), (2) as nice as Swedish/Danish, or (3) nicer than Swedish/Danish?“ The question could be answered by choosing one out of three possible answers that were assigned the scores indicated above (1, 2, or 3). A fourth possibility was to answer ‘no opinion’.

4. Results

Generally, the intelligibility scores of the children were high, so we judged the task to be appropriate to this age group. Our primary research question was to assess whether mutual intelligibility in children is asymmetrical as in adults, or not. The Danish children decoded 64.0 percent ($sd = 0.15$, $SE = .04$) and the Swedish children decoded 70.8 percent ($sd = 0.12$, $SE = .03$) of the presented stimuli. This difference is not significant $t(30) = 1.41$, $p = .17$. That means, that, in contrast to adult Danes, Danish children encounter the same amount of problems when decoding spoken Swedish as their peers from Sweden do when confronted with spoken Danish. The results are illustrated in Figure 1. On the left hand side, the mean intelligibility scores from Maurud 1976, Bø 1978, and Delsing & Lundin Åkesson 2005 for Danish and Swedish adults are displayed. On the right hand side, intelligibility scores from this experiment are shown. Swedish listeners are represented by black bars, Danish listeners by white bars. Note that absolute intelligibility cannot be compared between adults and children, because the experimental designs differed widely.

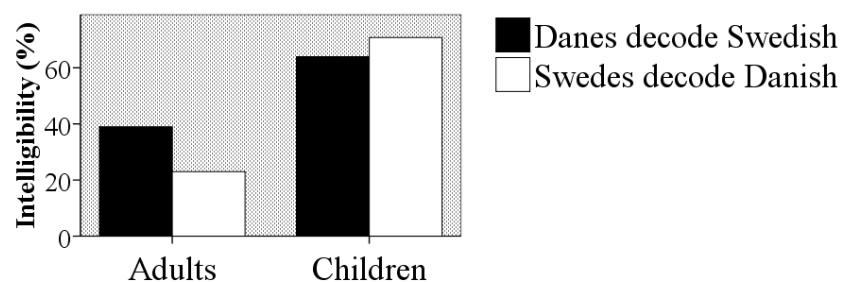


Figure 1. Adult and child intelligibility scores of Swedish listeners confronted with Danish stimuli, and of Danish listeners confronted with Swedish stimuli. Adult scores represent the mean over Maurud 1976, Bø 1978, and Delsing & Lundin Åkesson 2005.

To test whether we succeeded in excluding the three extra-linguistic factors literacy, previous exposure, and attitude, we first ran an independent *t*-test analysing the parents' information on the questionnaire as well as the children's answer to the attitude question. For the analysis of the attitude results, answer option 'no opinion' (n = 11) was neglected. The results are displayed in Table 2.

	L1	μ	<i>sd</i>	<i>SE</i>	<i>t(df)</i>	<i>p</i>
Intelligibility result (%)	Danish	0.71	0.15	0.04	1.4(30)	.17
	Swedish	0.64	0.12	0.03		
Parents' report of previous exposure (holidays, friends etc.)	Danish	0.50	0.52	0.14	1.10(30)	.29
	Swedish	0.27	0.47	0.08		
Parents' report of previous exposure (TV)	Danish	0.08	0.29	0.08	-0.67(30)	.51
	Swedish	0.18	0.41	0.12		
Parents' report of writing skills (number of words)	Danish	14.00	21.87	6.3	1.60(30)	.14
	Swedish	3.64	4.59	1.4		
Childrens' answer to attitude question	Danish	1.86	0.66	0.18	-0.48(19)	.63
	Swedish	2.00	0.85	0.22		

** Significant at the .01 level (two-tailed).

Table 2. Means, standard deviations, standard errors of mean, as well as results from an independent *t*-test analysis for information provided by parents and children.

From Table 2, it can be seen that the two groups of subjects did not differ significantly with regard to literacy and previous exposure, as reported by the parents. The Danish children were slightly more literate than the Swedish children, and the Swedish children were slightly more positive towards the neighbouring language than their Danish peers, but these differences were not significant. We therefore succeeded in making sure that the three extra-linguistic factors suggested did not have an influence on the asymmetry in mutual intelligibility.

5. Discussion

Adult mutual intelligibility between Danish and Swedish has been shown to be asymmetric in such a way, that Danes have fewer difficulties decoding spoken Swedish than Swedes have when decoding spoken Danish. In our study, we successfully excluded the influence of three extra-linguistic factors that have been suggested to cause this asymmetry: literacy, previous exposure, and attitude. We hypothesised that extra-linguistic factors do not account for a large part of the asymmetry, but we could not confirm this hypothesis with our data,

because our results show symmetrical intelligibility results. This finding suggests that extra-linguistic factors play an important role in mutual intelligibility between adult Danes and Swedes.

Our subjects turned out to have almost equally neutral attitudes towards, and an equal amount of previous exposure to their neighbouring language. Furthermore, they have acquired roughly the same amount of writing skills. It might be the case, that, if one or several of these variables change with age, they might give rise to an asymmetry in mutual intelligibility. To test this hypothesis, a similar word comprehension test should be conducted with literate subjects, whose attitudes towards the neighbouring language is asserted.

¹ This bias in writing skills was mainly due to the fact that the experiment session in Sweden was split up into two parts at different times of the year. Therefore, the children tested in the second session were older and somewhat more literate.

² All age specifications are decimals, i.e. 5.75 years represents 5 years 9 months.

References

- Bleses, D. & P. Thomsen. 2004. The acquisition of spoken forms and written words: An empirical study of opacity in the speech/reading/writing interface in Danish. *Written Language & Literacy* 7(1), 79-99.
- Bø, I. 1978. *Ungdom og naboland. En undersøkelse av skolens og fjernsynets betydning for nabospråksforståelsen*. Rogalandforskning 4, Stavanger.
- Delsing, L. & K. Lundin Åkesson. 2005. *Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska*. TemaNord 573. Köpenhamn : Nordiska ministerrådet.
- Gooskens, C. 2006. Linguistic and extra-linguistic predictors of Inter-Scandinavian intelligibility. In J. van de Weijer & B. Los (eds.), *Linguistics in the Netherlands 23*, Amsterdam: John Benjamins, 101-113.
- Gooskens, C., K. Beijering & W. Heeringa. In press. Phonetic and lexical predictors of intelligibility. *International Journal of Humanities and Arts Computing*.
- Gooskens, C. & G. Doetjes. In press. Skriftsprogets rolle i den dansk-svenske talesprogsforståelse. *Språk och stil*.
- Heeringa, W. 2004. *Measuring dialect pronunciation differences using Levenshtein distances*. Groningen: Groningen dissertations in linguistics (Grodil).
- Kürschner, S., R. van Bezooijen & C. Gooskens. In press. Linguistic determinants of the intelligibility of Swedish words among Danes. *International Journal of Humanities and Arts Computing*.

- Maurud, Ø. 1976. *Nabospråksforståelse i Skandinavia. En undersøkelse om gjensidig forståelse av tale- og skriftspråk i Danmark, Norge og Sverige.* Nordisk utredningsserie 13. Stockholm : Nordiska rådet.
- Schüppert, A. & C. Gooskens. Submitted. Linguistic predictors of word recognition. Evidence from Swedish pre-schoolers confronted with spoken Danish. *Speech Communication*.
- Schüppert, A. & C. Gooskens. In preparation. Do Danes speak more quickly than Swedes?
- Van Ooijen, B. 1994. *The processing of vowels and consonants. PhD thesis, University of Leiden.* Den Haag: Holland Academic Graphics.
- Wolff, H. 1959. Intelligibility and inter-ethnic attitudes. *Anthropological Linguistics* 1: 34-41.