Charlotte Gooskens, Sebastian Kürschner & Renée van Bezooijen

Intelligibility of Swedish for Danes: loan words compared with inherited words

The Mainland Scandinavian languages, i.e. Danish, Norwegian and Swedish, are so closely related that the speakers often use their own language when communicating with each other (so-called semi-communication, Haugen 1966). However, communication is not perfect and sometimes fails. Previous investigations (Maurud 1976, Bø 1978, Börestam 1987, Delsing & Åkesson 2005) aimed at getting a general impression of how well Scandinavians understand each other. It appeared that mutual intelligibility is highest between Norwegians and Swedes, whereas Danish is relatively hard to understand, especially for Swedish-speaking subjects (Perridon 2000).[1]

The investigations just mentioned measured the overall intelligibility[2] of complete texts by means of open questions. Little attention was paid to the linguistic phenomena that can explain the differences in the level of understanding between the three languages. In complete texts, all linguistic levels (segmental, supra-segmental, lexical, morphological, syntactic) are combined and mixed, so that it is hard to assess the effect of separate linguistic phenomena. In the present article, we restrict the investigation of

---

1 Strikingly, intelligibility is not necessarily symmetric. For example, Danes understand Swedish better than Swedes understand Danish.

2 In this paper the term 'intelligibility' is not used in the phonetic meaning of decoding speech sounds but in the more general sense of attaching meaning to sound sequences.

intelligibility to the understanding of isolated words, trying to determine the role of a limited set of phonetic/phonological factors that may affect intelligibility in inter-Scandinavian communication. We focus on the intelligibility of Swedish words for Danes, and in particular on the intelligibility of inherited words compared to loan words.

As in most western countries, puristic movements in Scandinavia have taken action against the large number of loan words that have become part of the Scandinavian languages (cf. Section 1). However, from the point of view of semi-communication in Scandinavia, it could be argued that a large number of such words is an advantage for mutual intelligibility, at least if the languages have borrowed the same words. We can think of three reasons why this might be the case.

First, loan words may have specific segmental and/or prosodic properties that make them resistant to linguistic changes affecting inherited words. For example, the word accent of many loans from Greek, Latin or Romance differs from the Germanic languages. Whereas Germanic languages are characterized by an initial accent on the stem syllable, most French loans, e.g., are stressed on the final syllable, cf. Swedish *miˈjö*, Danish *miˈøu* 'environment'. While in Germanic languages vowels in unstressed syllables are often reduced (or as in Swedish, a limited number of full vowels is found in this position), this does not happen as easily in loan words with a different accent structure. Here, full vowels are often retained in unstressed syllables, even if the non-accentuated syllable is final, cf. Danish *ˈdato*, Swedish *ˈdatum* 'date' from Latin.

Second, inherited words have been part of the lexicon for a much longer time than loanwords so that certain historical sound changes which affected the inherited vocabulary were no longer active at the time the loans entered the language. As a consequence, loan words in the neighbouring language probably have more transparent phonetic correspondences with their counterparts in the mother tongue than inherited words. For example, the Swedish word *lag* 'law' pronounced as [lɑːg] and the Danish equivalent *lov* pronounced as [lɔw] may have become so different that they are no longer intelligible for the speakers of the neighbour language. In contrast, a loan word like *team*, pronounced as [tiːm] both in Swedish and Danish, is no doubt easily identifiable.

Third, loan words are often known not only from the native language but also from foreign languages that the speakers are familiar with. This may also have a facilitating effect on the recognition of loan words com-

pared to inherited words. For example, the recognition of the Swedish word *turism* 'tourism' may be facilitated because Danish listeners know the English equivalent.

To assess whether it is true that shared loan words are easier to understand than inherited words, we tested the intelligibility of 355 Swedish words (197 inherited words and 158 loan words) among Danish listeners. We also measured the phonetic distances between the Swedish stimulus words and their Danish counterparts to determine whether the distances are indeed larger for the inherited words than for the loan words. Finally, we correlated the intelligibility scores with the phonetic distances to see whether there is a relation between word intelligibility and phonetic distance. We tested the following hypotheses:

1. The percentage of correct identifications is higher for loan words than for inherited words.
2. The phonetic distances between Danish and Swedish loan words are smaller than between inherited words.
3. There is an inverse relationship between the percentage of correct identifications and phonetic distance.

1. LOAN WORDS IN DANISH AND SWEDISH

Swedish and Danish both belong to the North Germanic group, and due to their common origin the two languages share a large number of inherited words. In the course of history, Swedish and Danish have also borrowed extensively from other languages. Because of common language contact situations, the two languages share many loan words. In the early Middle Ages, both languages borrowed Latin and Greek words as part of Christianization. Throughout the Middle Ages, Latin remained influential because of its leading role in the church and in the sciences. In the late Middle Ages, Danish and Swedish both had intensive contacts with Low German. This was the language of the Hanseatic League, which constituted a strong economic power in the area of the Baltic Sea. From the seventeenth until the nineteenth century, French exerted a strong influence. From the time of the reformation and the Thirty Years' War (1618-1648) until the nineteenth century, High German was very influential. In the twentieth century, and especially after the Second World War, English

words started to be adopted, especially within the domains of industrialization, transport, technology, and sports. In fact, English is now almost the sole provider of loan words.

To our knowledge, the proportion of loan words in Danish has not been quantified, but a comparison by Gooskens, Van Bezooijen & Kürschner (2010) of similarly constructed samples of formal spoken Swedish and Dutch showed that Swedish has many more loans than Dutch. Among the 1,500 most frequent words in a one-million-word database of each of the two languages 44.4% were loans in Swedish against only 27.9% in Dutch. The frequency data were gathered for lexemes rather than for word forms. This means that the frequencies of, for example, the Swedish word forms *hus* 'house' and *huset* 'the house' were combined. For Danish we expect proportions comparable with Swedish. For each word in the Swedish sample it was determined from which language it has been directly borrowed using etymological dictionaries. The largest group of loan words is formed by Low German (38.7%) followed by Latin (25.2%), French (14.6%), High German (14.3%) and Greek (4.6%). English loan words constitute only a small proportion of the loans (1.6%).

It is likely that loan words from different origins are characterized by different phonological features and that these differences may affect intelligibility. Unfortunately, there are no extensive studies which systematically compare loans from such different origins. Most of the newer studies concentrate on English loans only (cf. Heidemann Andersen & Rathje 2007, Heidemann Andersen & Jarvad 2008, Davidsen-Nielsen, Hanen & Jarvad 1999, Sørensen 1973) or on the graphematic level (cf. Gellerstam 2003). The latter we do not discuss here because we do not deal with the intelligibility of written forms but of spoken words. The available information suggests that there is an important distinction between loans from Low and High German and loans from all other loan-giving languages. Loans from German varieties are more similar to the Scandinavian varieties than other loans, and consequently they are less often identified as foreign. By contrast, Latin, French, and English loans are often clearly recognizable as loans and therefore more often subject to puristic tendencies (cf. Edlund & Hene 2005: 133-134, Hansen & Lund 1994: 124 & 133).

A comparison of consonant clusters introduced from foreign languages into Swedish reveals that many new clusters come from Greek and Latin or Romance languages (Edlund & Hene 2005: 105-109). Initial clusters im-

ported by means of Greek words are, e.g., /ps-/ (*psyke* 'psyche'), /pt-/ (*ptolemeisk* 'Ptolemaic'), /ks-/ (*xylofon* 'xylophone'), /tm-/ (*tmesis* 'tmesis'). Greek final clusters are, e.g., /-sm/ (*spasm* 'spasm'), /-tm/ (*rytm* 'rhythm'), /-rf/ (*morf* 'morph'). A large number of final clusters from Latin and Romance is found as well, such as /-rb/ (*verb* 'verb') or /-pt/ (*adept* 'adept'), and especially from French, e.g. /-ŋs/ (*chans* 'chance'). Some clusters are also introduced from German. Nearly all initial clusters from German begin in /ɧ/ (substituting German /ʃ/), cf. e.g. /ɧl-/ (*schlager* 'hit song'), /ɧn -/ (*schnitzel* 'cutlet'), but according to Edlund & Hene (2005:106) /ɧ/ is most often substituted by /s/ in such clusters, resulting in a genuine cluster. Final clusters from German include /-jf/ (*slejf* 'strap') and /-jt/ (*bojt* 'yell'). Edlund & Hene (2005) do not treat the borrowing of vowels, but it is important to note that Swedish has no genuine diphthongs. Therefore, it could be that words from languages which have a rich inventory of diphthongs, such as English, are differently adapted to the Swedish phonological system than words from languages with fewer diphthongs.

Hansen & Lund (1994: 86-103) summarize the most important phonetic adaptations of loan words in Danish. In the pronunciation of Greek, Latin and French loans words, a *stød* (a kind of creaky voice creating phonological contrasts) is often added according to the Danish rules (*cafe* 'café', *palæ* 'palace'). Accents are often placed on the first syllable following Germanic rules (*guitar* 'guitar' and *primær* 'primary'), but many words keep the original accent (*formular* 'form' and *autoritær* 'authoritarian' with accents on the last syllable). Nasal vowels are often rendered Danish by pronouncing the nasal consonant after the vowel as [ɳ] (*facon* 'shape'). Final *e* which is not pronounced in French is mostly pronounced as schwa in Danish (*massage* 'massage'). French *u* [y] is pronounced as Danish [u] (*robust* 'sturdy'). As far as the English loanwords are concerned, the pronunciation of some words is still rather close to the English pronunciation while others have adapted to the Danish pronunciation. The leading principle seems to be that English loan words are pronounced with the corresponding Danish sounds that are as similar to the English sounds as possible. Some words keep the original English accent while others shift the place of the accent according to the Danish rules.

Regarding loans from German, it is worth mentioning that both Low German and High German words are characterized by reduced vowels in unstressed syllables by the times of intensive language contact with

Scandinavian. Also with respect to phonotactics, both varieties of German are more in parallel with Danish and Swedish than the other contact languages. Words from the German varieties were thus similar to those of the North Germanic languages, and integration was more easily possible. In fact, Andersson (1994: 312) even considers that Low German "loans have been totally assimilated to the native vocabulary" of Swedish.[3]

The examples of phonetic characteristics presented above show that loan words from different languages have been adapted in various ways to Danish and Swedish. In Section 2 we will show how we quantified the differences between Danish and Swedish loan words and inherited words in order to investigate the role these differences play in intelligibility. In addition to their phonological make up, loan words can differ from inherited words with respect to their length. To assess the potential influence of this lexical feature on word intelligibility, it was also included in the analysis presented in Section 3.

2. METHOD

The Swedish-Danish intelligibility experiment reported on in this paper is part of a large-scale Internet investigation designed to test the intelligibility of seven Germanic languages for different groups of subjects in the Germanic language area. We are interested in the degree of intelligibility at a first confrontation, i.e. when subjects have very little prior experience with the test language. We selected 384 test words from a database with parallel lists of 2575 frequent words in the seven Germanic languages. We annotated the lists with different kinds of linguistic information to investigate the role of various word characteristics for the intelligibility. Written and spoken forms of the test words were presented via the Internet to groups of subjects in Scandinavia, the Netherlands and Germany in a translation task. In this section, we first give a global description of the database (Section 2.1), the selection of the 384 words (Section 2.2) and the general setup of the Internet experiment (Section 2.3). Next, we provide details on the part of the Internet experiment that

[3] Of course, some early loans from other languages are just as well integrated, cf. Greek loans like Swedish/Danish *biskop* 'bishop' or Swedish *kyrka* / Danish *kirke* 'church'.

tested the intelligibility of Swedish spoken words among Danes, which constitutes the topic of the present paper (Section 2.4). Finally, we show how we measured phonetic distance (Section 2.5).

2.1 DATABASE

We selected our Swedish test words from a database with parallel lists of high-frequency words in seven languages (Dutch, Frisian, High German, Low German, Danish, Swedish and Norwegian). These lists were compiled for a large investigation on the mutual intelligibility of Germanic languages. The present investigation is a part of this investigation.

As people are confronted with both formal and informal speech in everyday life, we decided to include both kinds of speech in the database. All words were collected from large corpora of Dutch (and partly Swedish) and then translated to the other Germanic languages.

The informal speech was selected from the Corpus of Spoken Dutch[4]. It was produced in casual interactions between friends and relatives in a homely atmosphere. The formal speech consists of Dutch and Swedish monologues in the European parliament, sampled in the Europarl corpus.[5] We took the 1500 most frequent words from each of the two corpora. As there was some overlap, the combined list included 2575 words. These words were translated into the other Germanic languages in our investigation. Each word in each language was enriched with information about word class, pronunciation,[6] origin (native word or loan word) and historical relationship[7] (cognate = historically related with the corresponding words in the other languages or non-cognate = not historically related).

---

4 See http://lands.let.kun.nl/cgn/home.htm.

5 See http://people.csail.mit.edu/koehn/publications/europarl/.

6 Transcriptions representing the standard languages stem from Uppsala (Swedish) and Copenhagen (Danish).

7 The Swedish etymological information was found in Wessén (1960) & Hellquist (1980).

2.2 SELECTION OF TEST WORDS

For pragmatic reasons, the number of words to be tested in the Internet experiment had to be restricted. The test would have become too long if we had included all 2575 words. We decided to use simple nouns only (no proper names, no compounds), since in general they are central to the intelligibility of a language. The selection resulted in a database of 815 simple nouns. To make sure that all concepts referred to were familiar to our subjects, i.e. secondary school pupils between 15 and 18 years of age, we tested among a group of 24 Dutch secondary school pupils whether they were familiar with the nouns in question. The pupils were asked to indicate which concepts were unknown to them. Eighty-two words were rejected by one or more participants and were removed from the sample. From the remaining 733 words we made a random selection of 400 to be used as test words in the Internet experiment. To obtain the spoken versions, male, native speakers of the seven languages read the words in professional audio studios. Sixteen of the words were not recorded well in one of the languages, leaving us with the final set of 384 test words. The recordings of these words were used for the listening test.

2.3 INTERNET EXPERIMENT – GENERAL SETUP

The experiment was carried out via the Internet.[8] To keep the task manageable, each person was presented with only 96 spoken words (a quarter of the material) and 96 written words (another quarter of the material). We recruited our subjects from secondary schools. This made it relatively easy to find subjects with comparable backgrounds in all countries. The spoken words were presented via headphones and the written words were presented at the computer screen. The task was simple: the subjects were asked to type the translation of each test word into their own language in a box on the computer screen. The response time was the same for all words (10 seconds). When the response time had

[8] It is possible to participate in the experiment as a 'guest' via the website http://www.let.rug.nl/lrs, login name 'germanic', password 'guest'. Via 'settings', guests can choose the test language and the number of test words. We thank Johan van der Geest for developing the application.

elapsed, the next word was automatically presented, but the subjects could also proceed to the following word immediately by pressing the enter key.

To motivate the subjects, a reward was offered to the best-performing pupil in each group. Furthermore, all subjects stood a chance of winning a prize, regardless of their performance. The whole test lasted approximately 40 minutes.

The Internet program was designed in such a way that the accuracy of the responses was checked automatically and that the percentage of correct responses was reported to the participant immediately after the test. Subsequently, mother tongue speakers manually checked all responses categorized as wrong translations. When the reason for a wrong translation was a spelling mistake, the response was counted as correct. We defined spelling mistakes as instances where only one letter had been spelt wrongly without resulting in an existing word. The response *ærende* (correct *ærinde*) 'errand', for example, was considered a correct translation with a spelling mistake, because only one letter is spelt wrongly and *ærende* is no existing Danish word. By contrast, *aske* (correct *æske* 'box') was not considered a spelling mistake, although only one letter differs. Since *aske* is an existing Danish word meaning 'ashes', it was impossible to determine if a spelling mistake had occurred or if the subject had meant to translate the stimulus with *aske*.

For each subject and word, we obtained a score of 1 if a word was translated correctly, and a score of 0 if the response was a wrong translation or if no response was given. The mean score for all subjects represents the intelligibility per word.

Since the test was carried out via the Internet, people with different backgrounds might participate. To be able to make a selection of subjects meeting specific criteria, the participants were asked to fill in a questionnaire about their social and linguistic background.

2.4 INTERNET EXPERIMENT – TESTING SPOKEN SWEDISH WORDS AMONG DANES

Our study of the intelligibility of Swedish spoken words among Danes is based on the performance of 42 Danish subjects (30 females and 12 males). They were 16 to 19 years old, with a mean of 17 years. They all at-

tended a pre-university school. They had Danish as their mother tongue and spoke Danish with both parents. Since we are interested in intelligibility at a first confrontation, we needed subjects who had had little contact with the Swedish language. We therefore excluded subjects living on the island of Zealand, which is geographically close to Sweden. Most of the subjects (N = 35) were from peninsular Jutland, mostly from Århus and Vejle (250 kilometers from Sweden), and the rest (N = 7) were from Odense and surroundings on the island of Funen (180 kilometers from Sweden). As an extra precaution, we also had the subjects translate a number of Swedish non-cognates. Such words should be unintelligible to subjects with no prior experience with the language. Indeed, hardly any of the non-cognates were translated correctly. An exception is formed by the word *flicka* 'girl' (Danish *pige*), which was translated correctly by 75% of the subjects. This word is probably known to most Danes as a stereotypical Swedish word, among others because it was used in the popular Danish pop song *sköna flicka* ('beautiful girl') by Kim Larsen. Furthermore, four other non-cognates were translated correctly once by three different subjects. On the basis of these results we decided not to exclude any of the 42 subjects.

Four subgroups of subjects each listened to 96 different spoken words, i.e. one quarter of the total set of words. Unfortunately, the computer application did not divide the 42 subjects equally across the test words. Some test words were presented to as many as eighteen subjects, while others were presented to no more than five subjects. The mean number of translations per word was 10. In total 384 words were tested. We include only the results of the 355 cognates in our analysis, since we are interested in the relation of intelligibility with phonetic distance and it is not meaningful to measure phonetic distance between historically unrelated words.

2.5 PHONETIC DISTANCE

To measure phonetic distance the so-called Levenshtein algorithm was used. Levenshtein distance is calculated automatically by computer on the basis of the phonetic transcriptions of corresponding word pairs (see Heeringa 2004 for an extensive discussion). The distance between corresponding words is based upon the minimum number of phonetic symbols that need to be inserted, deleted or substituted in order to transform the

word in Swedish into the corresponding word in Danish. In order to obtain distances which are based on linguistically motivated alignments, the algorithm was adapted so that a vowel may only match with a vowel and a consonant with a consonant. The approximants [j] or [w] could be matched with a vowel or a consonant. In this way, undesirable matches (e.g. a [p] with an [a]) are prevented. Word length was normalized by dividing the total sum of costs by the number of symbol alignments. Our measurements are based on segmental transcriptions only, i.e. we did not consider any suprasegmental information such as Swedish tonal accents or Danish *stød*, a kind of creaky voice that occurs in long vowels and in voiced (sonorant) consonants.[9] All distances were given the same weight. This means for example that a long vowel is considered to be equally deviant from its short counterpart as from any other vowel in the system, whether long or short. We will return to this in Section 3. As an example we present the calculation of the distance between the pronunciations of the Swedish word *butik* 'boutique' pronounced as [b~~ʉ~~tiːk] and the Danish equivalent pronounced as [butig]:

Alignments	1	2	3	4	5
Swedish	b	~~ʉ~~	t	iː	k
Danish	b	u	t	i̯	g
Costs		1		1	1

The sum of costs (1+1+1=3) is divided by the number of alignments (5). The result is a distance of 60%. We calculated the phonetic distances between the 355 Swedish words in the Internet experiment and their Danish cognate equivalents. The mean distance between the two languages is the mean distance over all 355 cognate word pairs.

3. RESULTS

To test the hypotheses formulated at the end of the introduction, we calculated (1) the percentages of correct translations of the Swedish words, (2) the phonetic distances between the Swedish words and their Danish

[9] Cf. Fischer-Jørgensen (1989) and Basbøll (2003) on stød in Danish. On the origin of the tone accents cf. Perridon (2006a), on the origin of the West-Jutlandic stød cf. Perridon (2006b).

equivalents, and (3) the correlation between these two measures. The percentage of correct answers and the phonetic distance values are presented in Table 1, broken down for loan words and inherited words. In Figure 1 the loan words are broken down for the loan-giving languages, i.e. the languages from which the words have been borrowed directly. We left out eight words with mixed or unknown origin.

	Loan words (N=197)	Inherited words (N=158)	Total (N=355)
% correct	76.7	68.0	72.8
phonetic distance	54.1	59.2	56.4

TABLE 1. Percentages of correct answers and mean phonetic distances for loan words, inherited words and total.

PERCENTAGES OF CORRECT ANSWERS

The mean percentage of correct answers is higher for loan words (76.7) than for inherited words (68.0). The difference is significant at the .01 level (t = 2,764, df = 353, $p < .01$). These results confirm the first hypothesis formulated in the introduction, that the mean percentage of correct identifications is higher for loan words than for inherited words. However, in Figure 1 we see that not all groups of loan words are better understood than the inherited words: fewer Low German loans than inherited words are understood correctly.

PHONETIC DISTANCES

The mean distance between the 355 Swedish and Danish words is 56.4% (see Table 1). As expected, the mean distance between the loan words is smaller (54.1%) than between the inherited words (59.2%). The difference is significant (t = 2,314, df = 353, $p < .05$). In Figure 1 the phonetic distances are presented for inherited words and loans words of different origin. We see that each single group of loan words has a smaller phonetic distance than the group of inherited words. This confirms the second hypothesis in the introduction, that the phonetic distances between Danish and Swedish loan words are smaller than between inherited words.

RELATIONSHIP BETWEEN PERCENTAGE OF CORRECT ANSWERS AND PHONETIC DISTANCES

Unfortunately, some groups of loans in Figure 1 contain relatively few words, which makes it somewhat hazardous to draw conclusions. This pertains especially to English (N = 13), Greek (N = 14) and High German (N = 8). However, the results presented in Figure 1 do suggest a systematic, inverse relationship between intelligibility and phonetic distance. The inherited words and the Low German words form the groups with the lowest percentage of correct answers and the largest phonetic distances. In all other groups, percentages of correct answers are higher as the phonetic distances are smaller. The relationship between intelligibility and phonetic distance is confirmed when correlating the percentages of correctly translated words with the phonetic distances. The correlation is significant (r = -.322, df = 354, p < .001). This confirms the third hypothesis in the introduction, that there is an inverse relationship between the percentages of correct identifications and phonetic distances. It is tempting at this point to interpret the correlational relationship as a causal relationship. Loan words are better understood *because* they have a smaller phonetic distance.

However, there is an alternative explanation that needs to be considered. Previous research has shown that word length plays a role in word recognition in that longer words are better identified than shorter words. This, in turn, is explained in terms of the relationship between word length and the number of 'neighbors', i.e. competing word forms that are very similar to the stimulus word (for an extensive description of the neighborhood activation model, see Luce & Pisoni 1998). Longer words have fewer neighbors than shorter words (Vitovitch & Rodriguez 2005) and there is therefore less chance that a listener chooses a wrong response. Neighborhood density is often defined as the number of words which deviate from the stimulus word in only one sound, disregarding the correct response. For example, the Swedish word *säng* 'bed' with the correct Danish translation *seng* has four Danish neighbors: *syng* 'sing', *senge* 'beds', *hæng* 'hang', and *stæng* 'close', while the Swedish word *adress* 'address' has no neighbors. Since Kürschner, Gooskens & Van Bezooijen (2008) showed that both word length and etymology (inherited words versus loan words) correlate significantly with the identification of Swedish words among Danes, we wanted to make sure that the greater intelligibility of loan words compared

with inherited words in the present study was not merely due to a concomitant difference in word length. We therefore computed the length of all 355 stimulus words and checked whether the various groups of loan words were longer than the inherited words. The results are presented in Table 2.

Inherited words (N = 157)	5.03
English (N = 13)	4.69
High German (N = 8)	5.00
Low German (N = 54)	5.33
French (N = 37)	5.54
Latin (N = 64)	6.75
Greek (N = 14)	6.86

TABLE 2. Mean word length of inherited words and loan words

Of the six groups of loan words, only two had a significantly longer word length than the inherited words (mean length of 5.03 symbols), namely Latin (mean length of 6.75 symbols, $t = –3.077$, $df = 219$, $p < .01$) and Greek (mean length of 6.86 symbols, $t = –2.719$, $df = 169$, $p < .01$). No significant differences with the inherited words were found for the other groups of loans. Overall the correlation between word length and word intelligibility was low but nevertheless significant ($r = .155$, $df = 354$, $p < .01$). This means that word length might covary with the subdivision of the lexicon into inherited words and loan words.

To assess if the effect of the subdivision of the lexicon (inherited vs. loan words) was primary or only secondary to word length, we conducted an analysis of variance (ANOVA) including the subdivision of the lexicon as a fixed factor and word length as a covariate. In the first step of this analysis, the dependency of the intelligibility scores is corrected for the effect of word length. Subsequently, the effect of the subdivision of the lexicon is analyzed on the corrected intelligibility scores, i.e. on the residuals. If the effect is still significant, we can conclude that the subdivision of the lexicon plays a role even if word length is controlled for.

The analysis confirmed the significant relation of the covariate, word length, to the intelligibility scores, $F(1, 352) = 6.19$, $p < .05$. However, after

controlling for the effect of word length, there was also a significant effect of the subdivision of the lexicon, $F(2, 352) = 5.14$, $p < .05$. This means that there is indeed an effect of the subdivision of the lexicon independent of word length, and we may thus conclude that etymology plays a significant role in the intelligibility of Swedish words among Danes.

Considering this in the light of the findings that phonetic distance is lower for loan words than for inherited words, phonetic distance thus seems to be a determining factor.

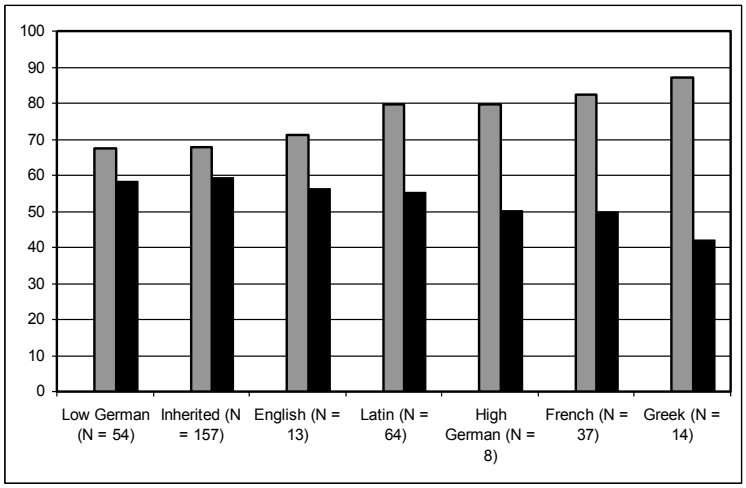

FIGURE 1. Percentages of correct answers arranged from lowest to highest (grey bars) and mean phonetic distances for inherited words and loan words of different origin (black bars), leaving out eight words with mixed or unknown origin.

4. DISCUSSION

The results of our analyses show that in general, Danish and Swedish loan words are both better understood and more similar than Danish and Swedish inherited words. This confirms our first two hypotheses in the introduction, namely that the percentage of correct identifications is higher

for loan words than for inherited words and that the phonetic distances between Danish and Swedish loan words are smaller than between inherited words. In the introduction we suggested three possible explanations: 1. Deviations in the phonological structure which prevent loanwords from taking part in sound changes leading to Danish-Swedish divergences; 2. The integration of loans into the language at a late point of time, i.e. after the relevant sound changes have happened; 3. (Only applicable to intelligibility) Knowledge of a loan from a foreign language. Since we do not have data on the third factor, we can only discuss the first two explanations here.

With respect to deviating phonological structures, in Section 1 we reported from relevant research that loans from German varieties are described as more similar to the Swedish (and Danish) inherited words than other loans. Loans from German varieties are even only rarely identified as loan words. For the group of Low German loans, which is even slightly less intelligible to Danes than the inherited words, this impression is confirmed. Both with respect to intelligibility, phonetic distance, and word length, this group resembles the inherited words of Swedish to a high extent. The data reveals that many of these words are first integrated differently into the respective language system, and then struck by phonological changes causing differences between the Swedish and Danish vocabulary. For example, Low German *Schapp* [ʃap] 'cupboard' has developed into Danish *skab* [sgɛːʔb] and Swedish *skåp* [skoːp]. The High German loans deviate from this impression. Maybe this is because Low German loans have entered Danish and Swedish earlier (i.e., before 1600) than High German ones (after 1600), so that they were subject to more sound changes. Still, since the number of High German loans (N = 8) is much lower than that of the Low German ones (N = 54), the deviation might also be a product of chance.

The other groups of loans are more similar in Danish and Swedish, and accordingly better recognized. This might be a hint that they also deviate more strongly from the Swedish and Danish inherited words, which might have protected them from certain sound changes. Most importantly, certain consonant clusters and different stress patterns might have prevented them from taking part in the general sound changes, cf. Section 1. A Greek loan like *problem* illustrates this, being pronounced as Danish [pʁo.ˈbleːʔm] and Swedish [prʊ.ˈbleːm]. In both languages, unlike usual

the second syllable is stressed, and full vowels are found in the unstressed syllable as well as in the stressed syllable. In an inherited word, stress would be expected on the first syllable, and only a restricted set of vowels would be allowed in the unstressed syllable.

Considering the time of integration, Low German loans were introduced much earlier than High German, French and English loans, giving them more time to diverge. For French (and, based on a small sample, High German) this is reflected in smaller phonetic distances and higher intelligibility compared with Low German. However, English words still have large phonetic distances. According to Vikør (1995, p. 181), English loan words are assimilated to the Swedish phonological system more easily than to the Danish system. Jarvad (2007: 207-218) refines this conclusion. She shows that on average Danish and Swedish adapt the pronunciation of a selection of vowels and consonants in English loan words to the native pronunciation to approximately the same degrees. However, some phonetic variables are adapted more to Danish than to Swedish, while some other variables are adapted more to Swedish than to Danish. For example, Danes pronounce the *th* in *death metal* as [θ] while Swedes mostly pronounce this sound as a [t]. On the other hand English *bacon* is pronounced with a full vowel in the second syllable in Danish according to the Danish rules but as a schwa in Swedish following the English pronunciation. This explains why the phonetic distances between the English loan words are rather large despite the fact that they have been borrowed so recently.

The oldest loans in the history of Swedish and Danish are of Latin and Greek origin. If the phonetic distance and intelligibility would mainly depend on the point of time when a word entered a language, we would expect these groups to be the least intelligible of all, and to show the largest distances. Still, while Latin holds a medium position, Greek loans are actually best understood and provide the smallest distances between Danish and Swedish. A reason for this might be that both Latin and Greek served as loan-giving languages over a very long time-span, starting much earlier than contact with Low German, but actually continuing until today. Nevertheless, we would expect the very old loans to be at least as divergent as the Low German ones. In our data, none of the Greek words are as well-integrated into Swedish as the oldest Greek loans (like *kyrka* 'church'), so based on our data we cannot draw conclusions about the relation between age and intelligibility in Greek loans. We should also con-

sider another factor, namely word length. As mentioned above, the Greek and Latin loans are significantly longer than the inherited words. Since the phonetic distances are normalized by dividing the distance with the length of the alignment (see Section 2.5), phonetic differences weigh less heavily in long words than in short words. This may cause smaller phonetic distances for the Latin and Greek words than for the other groups. Examples of long Greek loans are *bibliotek* 'library' and *strategi* 'strategy', for Latin loans *koordination* 'coordination' and *universitet* 'university'. These words have so much substance that differences in the pronunciation of one or two sounds hardly affect intelligibility (94, 100, 82 and 100% correct translations).

To sum up, the phonetic divergence and the intelligibility of certain groups of loan words can be explained by the degree of phonological integration, the point of time when the words entered the language, and the mean word length. None of these factors can explain the results in isolation.

5. CONCLUSIONS AND FURTHER RESEARCH

The hypotheses formulated in the introduction have been confirmed. Our results show that it is easier for Danish listeners to identify and understand Swedish cognate loan words than inherited words and that the phonetic distances between Danish and Swedish loan words are smaller than between inherited words. This general trend pertains to loan words from all origins. We found support for our claim that recent loans in Swedish have diverged less and are therefore phonetically more similar to the corresponding words in Danish than inherited words and older loan words. We also found support for the claim that the level of integration of loan words plays an important role for the phonetic distances between corresponding loan words in the two languages. Furthermore, the length of loans seems to be relevant.

Phonetically similar words can be expected to be easier to recognize than deviant words and, this is confirmed by the significant correlation between the percentage of correct identifications and phonetic distance per word pair. So we conclude that, from a communicative point of view, the large number of loan words in Swedish and Danish is in general an advantage.

REFERENCES

Andersson, Erik. 1994. Swedish. In: König, Ekkehard & Johan van der Auwera (eds.), *The Germanic languages*, 271-312. London/New York: Routledge.

Basbøll, Hans. 2003. Prosody, productivity and word structure: the stød pattern of Modern Danish. *Nordic Journal of Linguistics* 26:1, 5-44.

Bø, Inge. 1978. *Ungdom og naboland* [Youth and Neighboring Country]. (= *rapport*, 4) Stavanger: Rogalandsforskning.

Börestam, Ulla. 1987. *Dansk-svensk språkgemenskap på undantag* [Danish-Swedish Language Community as a Special Case]. Uppsala: Uppsala Universitet.

Davidsen-Nielsen, Niels, Erik Hansen & Pia Jarvad (eds.). 1999. *Engelsk eller ikke engelsk? That is the question.* Copenhagen: Gyldendal.

Delsing, Lars-Olof & Katarina Lundin Åkesson. 2005. *Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska* [Does the Language Keep the Nordic Countries Together? A Research Report on How Well Young People Understand Danish, Swedish and Norwegian]. Copenhagen: Nordiska ministerrådet.

Edlund, Lars-Erik & Birgitta Hene. 2005. *Lånord i svenskan. Om språkförändringar i tid och rum* [Loan words in Swedish. About language changes in time and space]. Stockholm: Norstedts.

Fischer-Jørgensen, Eli. 1989. *A phonetic study of the stød in Standard Danish.* Turku: University of Turku.

Gellerstam, Martin. 2003. Anpassningen av lånord i Svenska Akademiens ordlista. In: Sandøy, Helge (ed.), *Med 'bil' i Norden i 100 år. Ordlagning og tilpassing av utalandske ord*, 64-72 [By car in Scandinavia for 100 years. The making of words and addaption of foreign words]. Oslo: Novus.

Gooskens, Charlotte, Renée van Bezooijen & Sebastian Kürschner. 2010. The reflection of historical language contact in present-day Dutch and Swedish. In: Norde, Muriel, Bob de Jonge & Cornelius Hasselblatt (eds.), *Language contact. New perspectives*, 103-118. Amsterdam: John Benjamins.

Hansen, Erik & Jørn Lund. 1994. *Kulturens gesandter. Fremmedordene i dansk* [The ambassador of culture. Loan words in Danish]. Copenhagen: Munksgaard.

Haugen, Einar. 1966. Semicommunication: The language gap in Scandina-

via. *Sociological Inquiry* 36, 280-297.

Heeringa, Wilbert. 2004. *Measuring dialect pronunciation differences using Levenshtein distances*. Groningen: Grodil.

Heidemann Andersen, Margrethe & Pia Jarvad. 2008. Tilpasning af engelske importord i dansk skriftsprog [The adaptation of English loan words in Danish written language]. In: Omdal, Helge & Helge Sandøy (eds.), *Nasjonal eller internasjonal skrivemåte? Om importord i seks nordiske språksamfunn*, 105-124. Oslo: Novus.

Heidemann Andersen, Margrethe & Marianne Rathje. 2007. Siger danskerne *et fancy* eller *et fancyt bælte*? Tilpasning af importord i dansk talesprog [Do the Danes say *et fancy* or *a fancyt bælte*? The adaptation of loan words in Danish spoken language]. In: Jarvad, Pia & Helge Sandøy (eds.), *Stuntman og andre importord i Norden. Om udtale og bøjning*, 103-127. Oslo: Novus Forlag.

Hellquist, Elof. 1980. *Svensk etymologisk ordbok* [Swedish etymological dictionary]. Lund: LiberLäromedel.

Jarvad, Pia. 2007. Tilpasning eller ikketilpasning. Resultater [Adaption or no adaption? Results]. In: Jarvad, Pia & Helge Sandøy (eds.), *Stuntman og andre importord i Norden: om udtale og bøjning*, 201-220. Oslo: Novus.

Kürschner, Sebastian, Charlotte Gooskens & Renée van Bezooijen. 2008. Linguistic determinants of the intelligibility of Swedish words among Danes. *International Journal of Humanities and Arts Computing,* 2:1-2, 83-100.

Luce, Paul A. & Pisoni, David B. 1998. Recognizing spoken words: The Neighborhood Activation Model. *Ear and hearing* 19, 1-36.

Maurud, Øivind. 1976. *Nabospråksforståelse i Skandinavia: en undersøkelse om gjensidig forståelse av tale- og skriftspråk i Danmark, Norge og Sverige* [Mutual Intelligibility of Neighbouring Languages in Scandinavia. A Study of the Mutual Understanding of Written and Spoken Language in Denmark, Norway, and Sweden]. Stockholm: Nordiska rådet.

Perridon, Harry. 2000. Het Zweeds en het Nederlands als buurtalen [Swedish and Dutch as neighbouring languages. In: Hees, Annelies van & Harry Perridon (eds.), *Noordse Letters. Opstellen over Skandinavische taal- en letterkunde aangeboden aan An Duits*, 123-132. Amsterdam: Universiteit van Amsterdam.

Perridon, Harry. 2006a. On the origin of the Scandinavian word accents. In: Vaan, Michiel de (ed.), *Germanic tone accents. Proceedings of the first inter-*

national workshop on Franconian tone accents, Leiden, 13-14 June 2003, 91-105. Stuttgart: Steiner.

Perridon, Harry. 2006b. On the origin of the vestjysk stød. *Amsterdamer Beiträge zur älteren Germanistik* 62, 41-50.

Sørensen, Knud. 1973. *Engelske lån i dansk* [English loans in Danish]. Copenhagen: Gyldendal.

Vikør, Lars S. 1995. *The Nordic languages. Their status and interrelations*. Oslo: Novus.

Vitovitch, Michael S. & Eva Rodriguez. 2005. Neighborhood density effects in spoken word recognition in Spanish. *Journal of Multilingual Communication Disorders* 3, 64-73.

Wessén, Elias. 1960. *Våra ord: deras uttal och ursprung: kortfatted etymologisk ordbok* [Our words: their pronunciation and origin: a short etymological dictionary]. Stockholm: Norstedt.