Anja Schüppert

Linguistic predictors of non-native word recognition:
Evidence from Swedish pre-schoolers Listening to spoken Danish

In a number of perception studies, the Levenshtein algorithm has been employed successfully to calculate phonetic distance between language varieties. Recently, these phonetic distances have been used to predict word recognition in Scandinavian varieties among adults (Beijering et al. 2008), Gooskens & Heeringa 2006), Kürschner et al. 2008). If recognition scores are analysed per word, correlation coefficients with phonetic distances were reported to be significant, but lower than if they are averaged over several varieties. Since there is evidence that adult listeners benefit from their native writing skills when decoding spoken language (Perre & Ziegler 2008, Perre et al. 2009), we hypothesised that the Levenshtein algorithm is more suitable to predict word recognition in illiterate subjects.

This hypothesis was tested with 21 Swedish-speaking pre-schoolers that were auditorily presented with a set of Danish stimuli and instructed to choose the corresponding picture on a touch screen in a multiple choice picture-pointing task. Correlation coefficients between word recognition scores and phonetic distances confirmed our hypothesis that the Levenshtein algorithm predicts word recognition more accurately for naïve listeners ($r = -.62$, $p < .001$) than for literate adults with certain L2 knowledge of other languages. This finding confirms that word recognition of closely related languages is heavily influenced by extra-linguistic factors such as literacy. These factors need to be considered by educationalist developing language materials aimed at teaching closely related languages, both in Scandinavia and elsewhere.

1. Introduction

Danish, Norwegian and Swedish are closely related languages that are generally mutually intelligible to a certain extent. Therefore, Scandinavians often communicate with each other using their native languages rather than English as a lingua franca. This custom is strongly supported by the Nordic Council (*Nordiska ministerrådet*) and other Nordic authorities. However, communication between speakers of different varieties is not as unproblematic as communication between speakers of the same variety. How easy it is to decode a closely related variety depends, at least partly, on linguistic features of the varieties employed, such as the number of cognates between the two varieties, the deviance of pronunciation in cognate words, as well as on differences in prosody on word and sentence level (Gooskens & Heeringa 2006). Beijering et al. (2008) showed that it is possible to predict word recognition to a certain degree by integrating some of these factors. They evaluated the role of lexical and phonetic distances for word recognition and found that phonetic distance between two varieties is the best predictor of word recognition between these varieties. This finding was confirmed by Kürschner et al. (2008), who analysed the impact of 11 linguistic factors on the decoding abilities in Danish listeners confronted with Swedish stimuli.

To calculate phonetic distances, Gooskens and Heeringa (2006), Beijering et al. (2008), and Kürschner et al. (2008) used the Levenshtein algorithm (for more details see section 1.1). In these three studies, the participants were literate adults with varying degrees of L2 backgrounds. Danish listeners that can use their knowledge of English or German, for example, can be assumed to have less difficulty to recognise certain Swedish words than naïve listeners without that knowledge. In addition, as Perre & Ziegler (2008) have shown, literate listeners use their orthography when decoding spoken stimuli in their native language. It might be the case that speakers of a closely related variety also activate orthography when decoding spoken language. However, the Levenshtein algorithm models a listener without any knowledge of other languages and without writing skills - to name just a few extra-linguistic factors that are suggested to have an impact on decoding abilities (Schüppert & Gooskens fc.). In other words, this algorithm models a naïve listener.

The aim of the present experiment was to determine whether the Levenshtein algorithm predicts non-native word recognition more accurately in naïve listeners than in literate adults with different L2 backgrounds. To answer this question, we conducted an experiment with subjects that were chosen in such a way that they come close to being naïve listeners with respect to literacy and L2 knowledge, namely children from outside the border region. Furthermore, we aim at determining other factors that influence non-native word recognition in naïve listeners. Therefore, we investigate the role of five more factors that were shown to correlate significantly with Swedish word recognition among Danes, namely phonetic distances between cognate words, the Swedish tonemes, word length, differences in syllable number and the presence of unknown sounds (Kürschner et al. 2008). These factors will be discussed in detail below.

1.1 Phonetic Distances

Generally, cross-linguistic word recognition is facilitated by smaller phonetic distances, i.e. a non-native word whose pronunciation resembles the corresponding word found in a listener's native vocabulary tends to be easier to decode than a word which is pronounced very differently. Usually, two words that resemble each other share etymology and are considered cognate words. However, even among cognates, the range of phonetic similarity differs widely. Danish and Swedish share a large part of their vocabulary, but some of the cognate words are pronounced so differently that they are hard to recognise for listeners without any previous knowledge of the neighbouring language. One out of many examples for such word pairs is Danish *abe* [ɛbə] and its Swedish cognate *apa* [ɑːpa] ('monkey'). While the orthographic forms are somewhat easier to recognise for speakers of the neighbouring language, there are no shared phonemes in the standard pronunciations of Danish and Swedish in these words. If the acoustic features (such as place and manner of articulation) are considered, this means that the phonetic distance between the two words is 100%. The quantify the phonetic distance between pairs of words (or pairs of languages), the Levenshtein algorithm has been employed successfully by previous researchers.

The Levenshtein algorithm is a string edit distance measure that calculates the minimal costs required to change one string of symbols into another. There are three types of operations: insertions, deletions and substitutions of phonetic segments. It can be used to quantify the distance between the pronunciations of cognates in closely related varieties. Kessler (1995) employed the algorithm for measuring linguistic distances between several Gaelic dialects in Ireland. Since then it has been used on Sardinian (Bolognesi & Heeringa, 2002), Dutch (Heeringa, 2004: 213- 278), and German (Nerbonne & Siedle, 2005) data. It has also turned out to be useful for modelling intelligibility between speakers of different Scandinavian varieties. For example, it has been employed to predict non-native word recognition in Scandinavian dialects by adult Danes (Beijering et al. 2008) or to predict the intelligibility of Swedish words by Danish adults (Kürschner et al. 2008).

When choosing between insertions, deletions, and substitutions, the Levenshtein algorithm selects the operation that transforms one string into another at minimal cost, i.e. in such a way that the sum of the operations chosen by the algorithm is minimal. For example, the Danish word *nøgle* ('key') can be transformed into its Swedish counterpart *nyckel* in several ways, but some of these possibilities require more operations than other. This is illustrated in Table 1, which also gives an example of the kind of input that the Levenshtein algorithm processes and how it transforms one segment chain into another. The Swedish word *nyckel* and the Danish cognate *nøgle*, transcribed in X-SAMPA (extended SAMPA[1]), form the input for the algorithm and are aligned in such a way that, preferably, vowels correspond to vowel, and consonants to consonants. The algorithm then detects the number of necessary operations to change one word into its counterpart in the other variety. In the normalised version, the algorithm divides this number by the total number of segments after the two items have been aligned. By this, a distance percentage is obtained, namely the Levenshtein distance or phonetic distance. Cognates with 0 percent

[1] SAMPA (Speech Assessment Methods Phonetic Alphabet) is a machine-readable phonetic alphabet using ASCII symbols.

distance are identical in pronunciation and cognates with 100 percent are very different from each other. The cognates *nyckel* and *nøgle* have a phonetic distance of 67 percent, as illustrated in Table 1. For a detailed review of the implementation of the Levenshtein algorithm in dialectometry see Nerbonne & Heeringa (2010).

	(a) Segment-by-segment transformation					*(b) Lowest cost transformation*					
Swedish transcription IPA	n	y	k:	ɛ	l	n	y	k	ɛ	l	
Danish transcription IPA	n	ɔ	j	l	ə	n	ɔ	j		l	ə
Swedish transcription X-SAMPA	n	y	k:	E	l	n	y	k	E	l	
Danish transcription X-SAMPA	n	O	j	l	@	n	O	j		l	@
Necessary operations	0	1	1	1	1	0	1	1	1	0	1
Normalisation					4 / 5				4 / 6		
Obtained Levenshtein distance					= 80%				= 67%		

Table 1: Two possible ways of transforming one string into another. Version (b) forms the basis of the Levenshtein distance as employed here, as vowels are aligned to vowels and consonants to consonants.

Generally, suprasegmental features are not incorporated into the Levenshtein distance. That means that the Danish stød, a realisation of creaky voice or laryngealisation (Grønnum 1998: 179; Basbøll 2005: 83) will be disregarded. A Danish word with stød can therefore have a phonetic distance of 0% to its Swedish counterpart although there are no words in Swedish that have stød. Their pronunciation is thus not completely similar, but in standard pronunciation, their strings of phonetic segments are.

1.2 TONEMES

Swedish, unlike Danish, is a pitch accent language. That means that every word has its specific pitch contour which might distinguish it from a segmentally similar, yet semantically completely different word. There are two forms of these word-specific pitch contours, called toneme 1 and toneme 2. However, the realisation of the tonemes varies depending on regional variety. While toneme 1 in disyllabic words tends to have a peak in F_0 near the onset of the first syllable in Southern standard Swedish, F_0

peaks in the middle of the first syllable in central standard Swedish. In accent 2, the F_0 peaks at the end of the first syllable in Southern standard Swedish, but peaks twice in central standard Swedish, namely in the first and in the second syllable (cf. Meyer 1959).

While monosyllabic words are always assigned toneme 1, polysyllabic and compound words can have both tonemes. However, there is regional variation not only in the form of the two tonemes, but also in the distribution of them. Central standard Swedish tends to assign accent 2 to compounds to a larger degree than standard Southern Swedish (Perridon 2003).

Real minimal pairs are rare (Elert 1972: 163ff, Schüppert 2003: 12). Therefore, if a different pitch contour than the regionally employed one is assigned to a word, it will seldom cause misunderstandings, but it might puzzle the listener and thereby delay word recognition and influence how native the speaker sounds (Svensson (1978: 279). Figure 1 illustrates the pitch contours of two Swedish words with different tonemes by a native speaker from Småland. The word *bebis* ('baby') has toneme 1 and the word *äpple* ('apple') has toneme 2. It can be seen that the pitch in *bebis* 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. Monosyllabic words are always assigned toneme 1 and it is commonly agreed that this is the unmarked toneme (cf. Lahiri et al. 2005 for a critical review on this matter).

In contrast to Swedish, Danish does not have any tonemes. Words with an emphasis on the first syllable are pronounced basically with the same pitch contour. This is illustrated in Figure 2, which displays the pitch contours of the Danish words *baby* ('baby') and *æble* ('apple') by a native speaker from Funen. The pitch rises very subtly in the second syllable in both words before it falls again slightly.

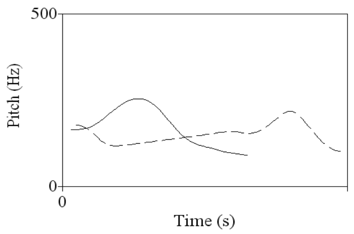

Figure 1: Pitch contours of the Swedish words *bebis* ('baby', solid line) and *äpple* ('apple', dashed line).

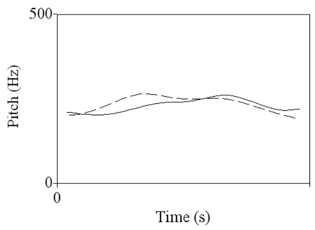

Figure 2: Pitch contours of the Danish words *baby* ('baby', solid line) and *æble* ('apple', dashed line).

From figures 1 and 2, 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 toneme 1 nor the contour of toneme 2. Gooskens & Kürschner (2010) provided evidence that adult speakers of Swedish encounter more problems when they are confronted with Danish stimuli that are cognates with Swedish words that have toneme 2, than when confronted with the correspondences of toneme 1 words. This suggests that the Danish pitch contour is be interpreted by the Swedish listeners as toneme 1. We hypothesised that this will be the case for our subjects as well.

1.3 WORD LENGTH

Kürschner et al. (2008) showed that Danish word recognition scores for long Swedish words were higher than those for short words. This result is 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 at least partly due to the fact that short words are less unambiguous because they have more 'neighbours', i.e. words that are segmentally similar, but semantically different. Kürschner et al. (2008) give the example of the Danish word *seng* ('bed') which has four neighbours: *syng* ('sing'), *senge* ('beds'), *hæng* ('hang'), and *stæng* ('close'). The word *motorcycle* ('motorbike'), on the other hand, is the word with the most syllables in our experiment. There is no word that is as similar to the word as *syng* is to *seng*. It is therefore less likely that the semantics of the word *motorcycle* are confused with those of another words, than that the meaning of *syng* is interpreted as meaning *seng*.

We expect the general tendency that the number of syllables correlates negatively with word recognition in this experiment. However, since we designed an experiment where the influence of the neighbourhood effect is minimised or excluded (see section 2.3), our data can reveal whether the effect of word length is mainly due to the neighbourhood effect, or not.

1.4 Difference in syllable number

Kürschner et al. (2008) provided evidence that the difference in syllable number has a hampering effect on word recognition in Swedish participants confronted with Danish stimuli. The correlation coefficient reported by Kürschner et al. (2008) is -.17, suggesting that stimuli that do not have the same number of syllables in the participants' native variety as in the test variety are more difficult to recognise than stimuli that have the same number of syllables. We hypothesise that the same is true for our participants.

1.5 Unknown sounds

Another factor scrutinised by Kürschner et al. (2008) is the influence of language-specific sounds that only exist in one of the two varieties. When decoding spoken language, listeners categorise the sounds they hear. Models of cross-linguistic speech recognition such as the Perceptual Assimilation Model (PAM) developed by Best (1994) or the Speech Learning Model formulated by Flege (1995) posit that non-native sounds are categorised into the native sound categories and that discrimination of a non-native contrast depends on the perceived phonetic similarity between the sounds in the contrast and the L1 categories.

Since there are a number of sounds that exist in Danish, but not in Swedish, we assume that listeners will experience difficulties to categorise some sounds correctly. For example, we hypothesise that the presence of the Danish approximant [ð][2] will have a deteriorating effect on word recognition, because the listeners cannot match it to one of their native phonemes.

1.6 LITERACY

There is evidence that literacy influences the way we process spoken language (Perre & Ziegler 2008, Perre et al. 2009). Doetjes & Gooskens (2009) showed that there are instances where Swedish listeners have an advantage from their native orthography when decoding spoken Danish. An example of this is the Danish word *stjerne* ('star'), pronounced [sdjɛɒ̯:nə]. The Swedish counterpart *stjärna* looks very similar to the Danish pronunciation in writing, but a Swedish phonological rule changes the consonant cluster /stj/ into the typical Swedish fricative [ɧ]. It is therefore pronounced [ɧæ:ɳa]. It could be assumed that literate Swedish listeners can use their orthographic knowledge to decode the Danish word [sdjɛɒ̯:nə]. In the present experiment, we are interested in modelling a naïve listener. Therefore we aim at excluding the influence of literacy on word recognition (see section 2.1).

2. METHOD

2.1 PARTICIPANTS

We tested 26 Swedish-speaking 4- to 6-year-old preschoolers living in the city of Växjö, 200 km from the Danish border, who had not been exposed to the neighbouring language before. To make sure that the subjects cannot use their native writing system to decode the stimuli they are confronted with, we conducted our experiment with illiterate children. In order to exclude children that might have acquired some basic phonographemic skills for their native language, a questionnaire was filled in by

[2] It should be pointed out that the usage of IPA-symbols is usually not congruent for English and Danish. Note that the IPA symbol 'ð' used in this article does not refer to the dental fricative as in English the [ðə], but to the Danish approximant.

the caretakers of every participating child before the experiment, asking if the child had learnt to read and write, and, if so, to estimate how many words the child could write. In the same way, information concerning previous exposure to the Danish language, via a holiday in Denmark, Danish friends or relatives, or TV, was elicited from the parents to be able to exclude children that had been exposed to the Danish language before. Furthermore, the parents were asked to indicate whether their child speaks another language than Swedish.

After the questionnaire evaluation, one child had to be excluded due to extensive contact with the Danish language through his Danish father. Furthermore, four children had to be excluded because their parents indicated that their children could read and write "many" words or "almost everything", leaving 21 children for the analysis. The children ranged in age from 4.0 to 6.5 ($\mu = 5.2$, $\sigma = 0.9$).

2.2 STIMULUS MATERIAL

Visual stimuli were employed as targets. The visual stimulus material consisted of 200 pictures from a picture database developed at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands. The auditory stimuli consisted of 50 Swedish-Danish cognate nouns that had different phonetic distances to their Swedish counterparts. The auditory stimuli ranged in distance from *identical* (0 percent distance) to *very distant* (100 percent distance) and were selected in the following way.

To make sure that the target pictures that should be matched to the auditory stimuli were clearly labelled with a cognate word by four-year-old Swedish children in their native language, and in order to find appropriate labels for the target pictures, the compilation of the material started by selecting pictures in the following way. 212 pictures were shown to five Swedish and five Danish four-year-old children in a pre-test. The children were asked to label the pictures spontaneously, i.e., to name the depicted object with one single word. If several labels were given by a child, only the first label was used for the calculation of the labelling consistency of every picture. Only pictures that were labelled with cognates and that had a labelling consistency of at least 80 percent (i.e. that were labelled similarly by at least four out of five children) were included in the set of visual stimuli. This was the case for 53 pictures. The pictures' labels formed the auditory set of stimuli.

The Danish and Swedish labellings were used as stimuli and were produced by female native speakers that had grown up in and still lived in Odense and Växjö, respectively. The recordings took place in sound atenuated rooms at the Centre for Child Language in Odense and in a radio studio in Växjö. Three of the stimuli were used for a demo version that was shown in advance to every child, leaving 50 stimuli for the word recognition experiment.

To determine the phonetic distance between every pair of cognates, the pronunciation of our two native speakers were transcribed. The mean distance of these 50 stimuli is 51 percent with a standard deviation of 24 percent, which can be seen in Figure 3, so the stimuli represent a broad range of evenly distributed phonetic distances. The Danish-Swedish cognate pair *bil-bil* (/biːˀl/ - /biːl/; ‘car’) constitutes one end of the continuum with a distance of 0 percent and the cognate pairs abe-apa (/æːbə/-/ɑpa/; ‘monkey’, ‘ape’), *stjärna-stjerne* (/sd̥jæɐ̯nə/-/ɧæːɳa/; ‘star’) and *öga-øje* (/ʌjə/-/øːga/; ‘eye’) constitute the other end of the continuum with a distance of 100 percent. Examples of cognates between these extremes are *katt/kat* (/kʰæd̥/-/kʰatː/; ‘cat’, 33 percent), *hund/hund* (/hunˀ/-/hɵnd/; ‘dog’, 50 percent) and *nyckel/nøgle* (‘key’, 67 percent). The Levenshtein distance continuum of the stimulus material is illustrated in Figure 4.

Additionally to the auditory stimuli and the 50 pre-tested pictures, 150 distracter pictures from the same picture database were chosen. Three pictures were randomly assigned to every pre-tested stimulus-picture pair, resulting in a set of four pictures per stimulus. Auditory or visual similarities between the pictures in one trial and the stimulus were avoided, to make sure that the task itself was equally difficult across trials. Since our task was a multiple-choice task, we assume that the neighbourhood effect was minimised or excluded from this experiment.

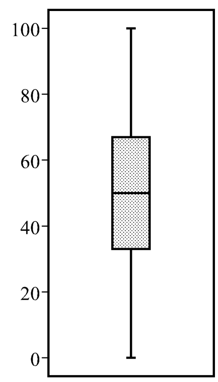

Figure 3: Box plot of Levenshtein distances for all 50 employed stimuli.

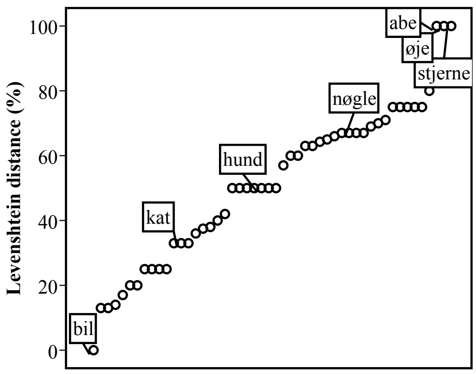

Figure 4: The Levenshtein distance continuum of the 50 stimuli.

2.3 PROCEDURE

The testing session consisted of a stimulus-response experiment (followed by a short interview with every child. The experiment was programmed and run in E-Prime 2.0. Before the experiment started, the children were familiarised with the task through a training session, which included the auditory presentation of two Swedish stimuli preceding one stimulus in Danish, and the visual presentation of four pictures per stimulus on a touch screen (LG L1510SF). The children were instructed to point to the picture that corresponded to the stimulus they heard. After the training session, the experiment started, which in the same way involved the choice of one out of four targets per stimulus in a picture-pointing task.

During the training session and the experiment, the child sat in front of the touch screen wearing ear phones. The stimulus material was presented randomised, but the same four pictures were assigned to a specific stimulus across subjects. Every stimulus was presented twice with an inter-stimulus interval of 3 000 ms. The four pictures remained on the screen until the child pointed to one of the pictures, or for 10 000 ms. One session lasted between one and a half and four minutes, depending on how quickly the child made a decision in every trial.

3. RESULTS

All 21 children were able to complete the experiment. The children recognised on average 63 percent of the stimuli (*Mdn* = 0.68, σ = 0.16). A Kolmogorov-Smirnov test revealed that the recognition percentages per word that were obtained in the experiment, *D*(50) = 0.12, *p* = .09, were normally distributed. We therefore assume that the included stimuli represent an appropriate range of difficulties to the listeners.

3.1 LEVENSHTEIN DISTANCE

Figure 5 displays a scatterplot of the word recognition scores plotted against the calculated Levenshtein distance for every auditory stimulus. It can be seen that there are no stimuli that were not decoded by any child, but four stimuli that were decoded correctly by all children (from left to right in the scatterplot: *lastbil* 'truck', *hest* 'horse', *ko* 'cow', *sko* 'shoe').

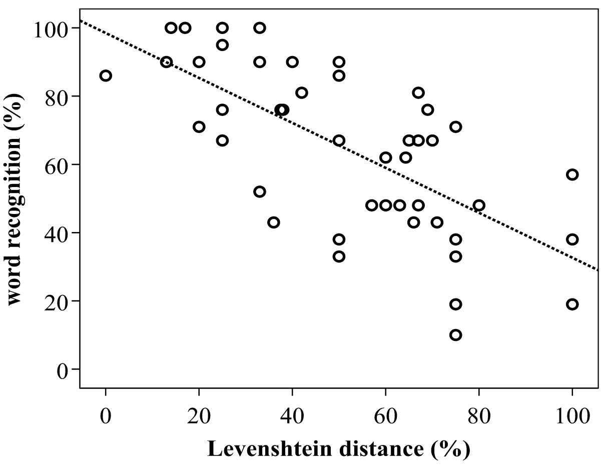

Figure 5: Scatterplot of word recognition scores plotted against calculated phonetic distances.

Our data confirms the negative correlation found by Beijering et al. (2008) and Kürschner et al. (2008). That means that the greater the calculated distance between the test word and the native cognate, the lower generally the recognition score for that specific pair of cognates. A correlation analysis between average word recognition scores per word and Levenshtein distance as confirmed this relation and resulted in the correlation coefficient $r = -.62$, ($p < .001$).

However, in order to find out which other factors influence non-native word recognition in pre-schoolers, we scrutinised the role of the remaining four factors discussed in sections 1.2 to 1.5 for word recognition scores. These features will be presented in detail below.

3.2 Tonemes

We were interested in finding out whether word recognition was enhanced or impaired for stimuli that correspond to words with toneme 1 or toneme 2 in Swedish, the participants' native language. An independent *t*-test revealed that the recognition scores were related to what toneme the counterpart of the stimuli had in Swedish. Danish cognates of words that have toneme 1 in Swedish were significantly easier to decode for Swedish-speaking subjects ($\mu = 0.73$) than Danish cognates of words that have toneme 2 ($\mu = 0.55$) in Swedish ($t\ (51) = -2.77$, $p = .008$). This finding is consistent with Gooskens & Kürschner's (2010) results. It is likely that Swedish listeners interpret the Danish tonal contour as corresponding to the Swedish unmarked toneme 1. When listeners expect the marked pitch contour of toneme 2, this expectation is not met when the Danish stimulus is presented. Apparently, this causes more problems for the listeners than if the marked contour is not expected and not found, as is the case for Danish cognates of words that have toneme 1.

3.3 Word Length

Kürschner et al. (2008) report that Swedish adults have fewer problems decoding long Danish words than short Danish words ($r = .25$, $p < .001$). We expected the same tendency in our participants, if this result is independent of the neighbourhood effect.

A correlation analysis between word recognition and the number of syllables per word resulted in the correlation coefficient r = -.26, p = .07, indicating that word length and word recognition are not significantly related to each other. Furthermore, the observed trend points in the opposite direction than expected, as indicated by the negative correlation coefficient. That means that the participants in our study have more problems recognising long words than short words. However, since monosyllabic words always carry toneme 1, which are easier to decode (see section 3.2), we need to control for that factor. If the potentially disrupting variable 'toneme' is controlled for in a partial correlation, the resulting coefficient is r = -.09 (p = .95). This indicates that there is no effect of word length on word recognition in our participants. It might be the case that the effect of word length found in other studies is mainly due to the neighbourhood effect.

3.4 DIFFERENCE IN SYLLABLE NUMBER

Kürschner et al. (2008) found a significant correlation between word recognition and the difference in syllable number between the presented stimuli and the corresponding words in the participants' native language (r = -.25, $p < .001$). We coded our stimuli trinarily, indicating whether the two cognates have the same number of syllables (difference is '0'), or whether Danish had more ('1') or fewer syllables ('-1'). A point-biserial correlation analysis of this factor showed that there was no significant correlation in our data (r_{pb} = .01, p = .94). However, this might be due to the fact that 44 of our stimuli (88 percent) did not differ in syllable number which renders insufficient statistical power. A reason for that, in turn, is likely that our stimuli were chosen on the basis of frequency and early acquisition in childhood. These prerequisites are usually met by a disproportionately high number of monosyllabic words. In our data, 28 Danish stimuli, i.e. 56 percent of the material, were monosyllabic and most of them (namely 26) corresponded to monosyllabic words in Swedish. A partial correlation controlling for word length revealed that, in contrast to Kürschner et al.'s (2008) data, there is no significant correlation between the difference in syllable number and word recognition (r = -.12, p = .43) in our data.

3.5 UNKNOWN SOUNDS

Some non-native sounds seem to fall within corresponding phoneme boundaries for a specific group of listeners, whereas others are more difficult to categorise. We hypothesised that words including the Danish sound [ð] were particularly difficult to decode for Swedish listeners and aimed at identifying which Danish sounds are disproportionally more difficult to categorise for Swedish listeners. We therefore split up our transcriptions segment by segment in a matrix and correlated the presence or absence of any Danish sound with word recognition scores (see an example in Table 2).

X-SAMPA	a	b	d	e	f	g	h	i	j	k	l	..	Word recognition (%)
[sbajˀl]	1	1	0	0	0	0	0	0	1	0	1	..	41
[biˀl]	0	1	0	0	0	0	0	1	0	0	1	..	91
[gafəl]	1	0	0	0	1	1	0	0	0	0	1	..	95

Table 2: Extract of the matrix of all transcribed phonemes.

Three sounds showed a significant negative correlation. The schwa appeared 18 times in 17 of the stimuli (the word *rutsjebane* 'slide' contains two schwas) and seems to be difficult to match to the corresponding native sound for the Swedish listeners (r_{pb} = -.33, p = .02). Another unknown sound was [ɛ], which appeared 18 times in 16 words (r_{pb} = -.29, p = .04). Finally, there were four words containing the Danish approximant [ð]. This typical Danish sound also correlated negatively with word recognition (r_{pb} = -.28, p = .05), indicating that it caused problems for Swedish listeners. It can be assumed that words containing one or several instances of the sounds [ə], [ɛ] or [ð] are especially difficult to recognise for Swedish children. We therefore recoded our stimuli indicating whether none ('0'), one ('1'), or several ('2', '3' etc) of these three sounds were present for inclusion in a multiple regression analysis for the following analysis.

3.5 MULTIPLE REGRESSION ANALYSIS

To summarise, we found that phonetic distance, tonemes, and unknown sounds correlate significantly with word recognition. To investigate how much these factors contribute to word recognition in relation to each other, we ran a multiple regression analysis in the enter mode with phonetic distance, tonemes, and unknown sounds as the three independent factors. As Kürschner et al. (2008) reported that phonetic distances have the greatest impact on word recognition it was entered in block 1, with tonemes and unknown sounds in block 2. The results are displayed in Table 3.

The analysis revealed that predicting word recognition in pre-schoolers by means of these three factors was slightly more accurate (R = .66, p > .001) than using phonetic distance alone (R = .62, p > .001). To predict word recognition among speakers of a closely related variety, these factors need to be considered. However, their impact was not large.

		B	*SE B*	*Beta*
Step 1	**Constant**	0.99	0.06	
	Levenshtein distance	-0.61	0.11	-.62**
Step 2	**Constant**	0.97	.06	
	Levenshtein distance	-0.52	.12	-.54**
	Toneme	.00	.07	.00
	Unknown Sounds	-.06	.03	-.24*

Note R^2 = .39**; ΔR^2 = .05, p < .001. * p < .01, ** p = .001 (one-tailed).

Table 3: Results from a linear regression analysis with average word recognition scores per word as dependent variable and phonetic distance, toneme, and 'unknown sounds' as independent variables.

4. DISCUSSION

The correlation coefficient between word recognition and phonetic distance reported by Kürschner et al. (2008) for adult Danes confronted with spoken Swedish was -.27 (N = 384, p < .001). The fact that our correlation coefficient is significantly higher might be due to the fact that we presented Danish stimuli to Swedes, whereas Kürschner et al. (2008) presented Swedish stimuli to Danes. However, it might also indicate that the Levenshtein algorithm is indeed more suitable for predicting word recognition in illiterate participants that have no knowledge of other foreign languages. We assume that some words might have been easier to decode for the participants in Kürschner et al.'s (2008) investigation than what was solely predicted by the phonetic distance because the listeners could make use of their orthographic and broader L2 knowledge. Since our subjects were illiterate and naïve with respect to foreign languages, this was no option for them.

Furthermore, as the task in Kürschner et al.'s (2008) experiment was to translate the presented stimulus and type it, it is likely that some stimuli were partly wrongly decoded because a similar word exists in the subjects' native language ('neighbourhood effect'). In our experiment, we used a multiple-choice task where the children could only choose between four different answers, which reduced the neighbourhood effect to a minimum. Thirdly, in contrast to the stimuli employed by Kürschner et al. (2008), who presented highly frequent nouns from formal and informal corpora, only highly frequent nouns from informal corpora were presented in our experiment, which might make a difference to the subjects. We can conclude that phonetic distances account for 42 percent of the variance in the word recognition of spoken Danish among Swedish pre-schoolers without any L2 background.

In a multiple regression analysis we were able to explain 44 percent of the variance in word recognition by means of the Levenshtein distance and the number of unknown sounds. Again, this is a higher proportion than Kürschner et al. (2008) found ($R^2_{Nagelkerke}$ = .21). This might partly be influenced by the statistics employed, since they employed a logistic regression. It might also be a result of the fact that word recognition in illiterate pre-schoolers, i.e. in naïve participants, depends on fewer variables than in adults, and therefore phonetic distance plays a greater role for word recognition. However, in order to gain results that are more accurately comparable between children and adults, more experiments need to be carried out. Investigating word recognition in Danish speaking pre-schoolers confronted with spoken Swedish can shed more light on the role of linguistic factors for word recognition. A replication of our study with adult speakers of Danish confronted with spoken Swedish is highly desirable. Also, the role of orthography should be investigated more thoroughly in the future. Finally, the two factors considered in our regression analysis explain less than half of the variance of word recognition in preschoolers. Future research needs to define possible further factors that influence isolated word recognition, both in adult listeners, and in children.

ACKNOWLEDGEMENTS

We would like to 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 on the experimental design by Renée van Bezooijen and Sebastian Kürschner.

REFERENCES

Beijering, Karin, Charlotte Gooskens & Wilbert Heeringa. 2008. Predicting word recognition and perceived linguistic distances by means of the Levenshtein algorithm. *Linguistics in the Netherlands*, 13-24.

Best, Catherine T. 1994. The emergence of native-language phonological influences in infants: a perceptual assimilation model. In: Howard C. Nusbaum & Judith Goodman (eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, 1-30. Cambridge, MA: MIT.

Bolognesi, Robert & Wilbert Heeringa. 2002. De invloed van dominante talen op het lexicon en de fonologie van Sardische dialecten. *Gramma/TTT* 9, 45-84.

Doetjes, Gerard & Charlotte Gooskens. 2009. Skriftsprogets rolle i den dansk-svenske talesprogsforståelse. *Språk och stil* 19, 105-123.

Elert, Claes-Christian. 1972. Tonality in Swedish: Rules and a list of minimal pairs. In: Firchow, Evelyn S., Kaaren Grimstad, Nils Hasselmo & Wayne A O'Neil. (eds.), *Studies for Einar Haugen*. The Hague: Mouton, 151–173.

Flege, James E. 1995. Second language speech learning: Theory, findings and problems. In: Strange, Winifred (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues*, 233–277. Baltimore: York Press.

Gooskens, Charlotte & Wilbert Heeringa. 2006. The Relative Contribution of Pronunciation, Lexical and Prosodic Differences to the Perceived Distances between Norwegian dialects. *Literary and Linguistic Computing* 21:4, 477-492.

Gooskens, Charlotte & Sebastian Kürschner. 2010. Hvilken indflydelse har danske stød og svenske ordaccenter på den dansk-svenske ordforståelse? In: Falk, Cecilia, Andreas Nord & Rune Palm (eds.), *Svenskans beskrivning 30*, 82-92.

Grønnum, Nina. 1998. *Fonetik og Fonologi. Almen og dansk*. Copenhagen: Akademisk forlag.

Heeringa, Wilbert. 2004. *Measuring dialect pronunciation differences using Levenshtein distance*. Groningen Dissertations in Linguistics.

Kessler, Brett, 1995. Computational dialectology in Irish Gaelic. *Proceedings of the 7th Conference of the European Chapter of the Association for Computational Linguistics*, 60-67.

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, 83–100.

Lahiri, Aditi, Allison Wetterlin & Elisabet Jönsson-Steiner. 2005. Lexical specification of tone in North Germanic. *Nordic Journal of Linguistics* 28, 61-96.

Meyer, Ernst. A. 1959. *Die Intonation im Schwedischen II. Die norrländischen Mundarten*. Stockholm: Studies in Scandinavian philology.

Nerbonne, John & Christine Siedle. 2005. Dialektklassifikation auf der Grundlage aggregierter Ausspracheunterschiede. *Zeitschrift für Dialektologie und Linguistik* 72, 129-147.

Nerbonne, John & Wilbert Heeringa. 2010. Measuring Dialect Differences. In: Schmidt, Jürgen Erich & Peter Auer (eds.), *Language and Space: Theories and Methods* in series *Handbooks of Linguistics and Communication Science*, 550-567. Berlin: Mouton De Gruyter.

Perre, Laetitia & Johannes C. Ziegler. 2008. On-line activation of orthography in spoken word recognition. *Brain Research* 1188, 132 – 138.

Perre, Laetitia, Katherine Midgley & Johannes C. Ziegler. 2009. When beef primes reef more than leaf: Orthographic information affects phonological priming in spoken word recognition. *Psychophysiology* 46:4, 739-746.

Perridon, Harry. 2003. On the origin of the Scandinavian word accents. Talk presented at the First International Conference on Franconian Tone, University of Leiden, 13–14 June 2003.

Scharpff, Peter J. & Vincent van Heuven. 1988. Effects of pause insertion on the intelligibility of low quality speech. In: Ainsworth, W. A. & J. N. Holmes (eds.), *Proceedings of the 7th FASE/Speech-88 Symposium*, Edinburgh, 261-268.

Schüppert, Anja. 2003. *Der Einfluss der Merkmale der Erstsprache auf den Zweitspracherwerb des Schwedischen: Eine vergleichende Studie anhand von serbisch, bosnisch- und albanischen Lernern*. Unpublished Master thesis, Humboldt-Universität zu Berlin.

Schüppert, Anja & Charlotte Gooskens. fc. The role of extra-linguistic factors for receptive bilingualism: Evidence from Danish and Swedish pre-schoolers. *International Journal of Bilingualism* 16:3.

Svensson, Bengt. 1978. On tonal accent in Hungarian Swedish: Some experimental results. In: Gårding, Eva, Gösta Bruce & Robert Bannert (eds.), *Nordic prosody*, 279-289. Lund: Dept. of Linguistics, Lund University.

Wiener, F. M. & George A. Miller. 1946. Some characteristics of human speech. Transmission and reception of sounds under combat conditions. Summary technical report of Division 17, National defense Research Committee, Washington, 58-68.